# Learning about Women's Competence:
# the Dynamic Response of Political Parties to Gender Quotas in
# South Korea[*]

Jay Euijung Lee[†]        Martina Zanella[‡]

April 14, 2023

## Abstract

Although gender quotas in politics are one of the most common affirmative action policies worldwide, the merits of these policies remain an object of debate. Opponents of quotas are concerned with an equity-efficiency tradeoff – quotas lower the average quality of politicians. This paper explores this tradeoff by studying the effect of gender quotas on parties' selection of candidates in South Korean municipal council elections. South Korea provides us with a rare opportunity to observe how highly male-dominated political parties react to gender quotas: (i) we can exploit the discontinuity in the intensity of the quota's effect at given cut-offs of council size; (ii) gender quotas were implemented in only one of two independent election arms, leaving space for adjustment in the other arm. We find that political parties initially counteract the quota by putting forth fewer female candidates in the unaffected arm. However, this pattern gradually reverses over time. The evidence is consistent with efficiency gains from the quota. Parties initially selected a suboptimally low number of women due to biased beliefs regarding their ability, but they slowly revise their beliefs after exposure.

# 1 Introduction

Gender quotas in politics are currently used by 130 countries of the world (International IDEA, 2021). The rationale behind quotas is that there would be too few women in politics otherwise. Given a particular notion of the optimal female share as one that maximizes the average ability of politicians, "too few women" can mean two very different things: i) we have attained the optimal number of women, just not equal representation, or ii) we are not at the optimum. The case of the former constitutes a standard equity-efficiency tradeoff – quotas reserve for women the seats that would have been filled by more qualified men. Yet, Holzer and Neumark (2000) review that the empirical case against affirmative action on the grounds of efficiency is weak at best, and more recent empirical studies find that quotas either maintain or improve the overall quality of political representatives (Bagues and Campa, 2021; Murray, 2010; Besley et al., 2017; Baltrunaite et al., 2014; Weeks and Baldez, 2015)

If quotas are needed because there are efficiency gains to be had with more women, then why hasn't the optimal number of women been attained in the first place? In this paper, we shed light on this question by studying how highly male-dominated political parties in South Korea change their selection of political candidates in response to the introduction of gender quotas. Our findings put forward a lack of information and biased beliefs about the competence of women as a force behind their low representation.

South Korean municipal councils are a particularly apt setting to study this question for several reasons. The institutional setup of the quota in South Korea allows us to detect what is typically unobservable: the preference of political parties for female candidates, or lack thereof. The gender quota regulates only one of the two separate arms through which councilors get elected. We can thus study how parties strategically respond in the unconstrained arm. Furthermore, with a 3% female share prior to the quota, South Korean municipal councils represent an environment potentially susceptible to a lack of information on women – a common feature of highly gender-imbalanced environments where quotas are needed the most. This also sets the paper apart from previous studies of gender quotas where female shares were already above 30% before quotas. Lastly, in a highly male-dominated environment where the quota may act as a big shock to the status quo, long-term strategies might differ from immediate responses. We therefore track parties over four election cycles post-quota. This is made possible by a novel dataset containing information on the universe of candidates and elected councilors for South Korean municipal councils since their inception in 1995 - 55,040 candidates and 7 election cycles - that we assembled by scraping the website of the National Election Commission. It is indeed the evolution of parties' responses over time that helps us uncover the reasons behind the initial under-representation of women.

For our identification strategy, we employ a regression discontinuity design that exploits the cross-sectional variation in the intensity of the quota emerging from differences in council size. In South Korea's mixed electoral system, one group of councilors is elected by plurality vote in the municipality's constituent wards ("ward warm") while the other group is elected by party-list proportional representation ("PR arm"). The gender quota stipulates that all odd-number candidates in the party list for the PR arm be female.[1] The number of PR seats increases as a step function of a municipality's council size, creating discontinuities in the intensity of the quota's bite at certain cutoffs of council size. We use this cross-sectional variation in the number of seats reserved for proportional representation as a measure of the intensity of exposure to the gender quota to study its effect on how political parties select candidates in the ward election arm.

We find that the presence of an unconstrained arm allowed parties to counteract the introduction of the quota. In the treated municipalities affected more intensely by the quota, parties initially put forth fewer female candidates in the unconstrained ward arm. The reduction in the number of female candidates is especially pronounced when the probability of winning is higher – in favorable ballot positions, in wards where a party has a stronghold, and among the two largest parties. Hence, although the quota was successful in fostering the election of women through the constrained PR arm, fewer female councilors got elected through the unconstrained ward arm in the first cycle after the introduction of the quota.

However, the pattern gradually reverses over time. Over the following three election cycles, parties in the treated municipalities gradually increase the number of female candidates in the unconstrained ward arm. By the last election, these municipalities had entirely reversed their initial reaction and, in fact, had a *greater* number of female ward candidates than control municipalities. These patterns are echoed in the election outcomes. Treated municipalities elect significantly more female councilors and significantly fewer male councilors from election cycle 5 onward.

What is driving this initial counteraction and gradual change in the response to quotas? The evidence is consistent with the predictions of a simple model of learning where parties have imperfect information and biased beliefs regarding women's ability as political leaders. As a consequence, they initially select a suboptimally low number of women, but they gradually learn about their competence through exposure to female councilors. Four pieces of evidence, guided by the model predictions, support this learning story.

First, the change in candidates' selection is related to the acquisition of new information about women as politicians. The model predicts that the process of belief updating is faster the greater the number of newly elected female councilors. We find that the quota effectively promoted the

---

[1]Parties, in particular party leaders, determine the set of candidates running for election. The case is obvious for the PR arm because one cannot be elected without being a member of a party in that arm. However, also in the ward arm, it is the parties that nominate the candidates to run for election. It is possible that a candidate runs as an independent, but very rarely he or she will win.

election of a higher number of rookie female councilors in treated municipalities, exposing parties to new information about women's competence as political leaders.

Second, the shift towards female candidates occurs at a faster rate when elected women are more competent, in line with the prediction that beliefs on the ability of women update faster the better women are found to be. Using education as a proxy for competence (Baltrunaite et al., 2014; Bagues and Campa, 2021), we compare the evolution in the choice of candidates for parties in municipalities where the first female PR councilors post-quota (cycle 4) were above- and below-median level of education. In the cycles following the exposure to these differentially able women, we find stronger increases in the number of female ward candidates in treated municipalities in the above-median group.

Third, experiencing a woman in action is pivotal for parties to learn about her competence, and parties with more biased priors learn more. In a *party*-level analysis, we compare parties in close PR elections that marginally won a female councilor against those that marginally lost. As only the winners are able to see their female candidates elected by random chance, this comparison gives us the causal effect of experiencing those women in action. We also distinguish parties by their ex-ante gender bias, using the fact that the gender of the even-number positions in the party list is unrestricted. We find that winning the election significantly improves a woman's chances to rerun in the next election cycle. However, the effect is completely driven by parties that placed a man in the second PR seat, so it is not due to gained experience or visibility. Rather, this evidence is consistent with the model prediction that belief updating is faster the larger the bias in the prior.

Lastly, the evidence points towards parties not only learning about the competence of the elected women but also updating their beliefs regarding women as a group. The increase in female ward candidates comes from not just incumbent women, but also rookie women who have zero councilor experience and on whom parties lack information. This suggests that parties' beliefs are changing about women as a whole as a decreasing gender gap in the prior leads to an increasing number of new women relative to men.

Even if there were absolutely no learning, three potential alternative explanations could be consistent with a faster increase in the number of women in treated municipalities. First, the supply of available competent female candidates might grow faster in treated municipalities compared to control municipalities. Second, parties might be responding to a faster change in voter preferences for females in treated municipalities. Third, parties may have taste-based discrimination against women, and grow out of their distaste for women faster in treated municipalities. None of these alternative mechanisms find support in the data.

From a policy perspective, much can be learned from the South Korean experience. This paper shows that quotas might deliver undesirable results when biased beliefs against female politicians are prevalent. With females holding only 3% of seats before the quota and more than 60% of the

population agreeing with the statement that men make better political leaders than women (see Figure A2), South Korea would have been such a context. Indeed, we observe parties and their leaders initially counteracting the quota.

Quotas can still be effective in the long run, however, when designed appropriately. The quota imposed that a minimum number of qualified female councilors got elected, thereby opening up the possibility for learning to take place. This is different from other types of quotas that do not ensure women end up elected, such as quotas on the minimum share of women in candidate lists. Such quotas have been found to be limited in increasing female representation to appreciable levels (Bagues and Campa, 2021; Dahlerup and Freidenvall, 2013).

South Korean politics is not the only context characterized by strong male dominance and conservative gender attitudes. The South Korean experience can be informative for many other countries currently characterized by very low female representation in politics, including Brazil, Japan, Hungary, Iran, Malaysia, Moldavia, Lebanon, Mali, and Nigeria.[2] Moreover, affirmative action policies are currently discussed also in other settings such as company boards (e.g. Global Gender Gap Report, 2021), where the incumbents are similarly if not more, male-dominated and attitudes are equally male-friendly (see Figure A3 in Appendix Section A).

Our paper contributes to a large and growing literature on the consequences of political gender quotas, studied in numerous contexts and with mixed results (see Hessami and da Fonseca, 2020 for a review).[3] However, while several studies look at final outcomes, only a few focus on the mechanisms that lead to these results, and even fewer study the reactions of the male-dominated pre-existing parties and their leaders (e.g. Esteve-Volart and Bagues, 2012; Casas-Arce and Saiz, 2015; Besley et al. (2017)). We contribute to this literature as we are able to pin down the strategic responses of parties to the introduction of quotas and study them in an unusually rich way. The voting system with the quota only applying to the proportional representation arm implies a whole other arm is unconstrained. Furthermore, the unconstrained arm of the election system is the way through which around 85% of councilors are elected, and therefore constitutes the more consequential arm. Such a structure of gender quota greatly expands the degree of freedom in which parties can respond, relative to quotas that reserve seats for women (Chattopadhyay and Duflo, 2004; Clayton, 2015), alternate between male and female candidate lists (Besley et al., 2017), or mandate a minimum share of women in candidate lists (De Paola et al., 2010; Esteve-Volart and

---

[2]See Figure A1 in the Appendix for data on female representation in national parliaments and attitudes.

[3]A sizable body of papers study the impact of quotas - and more generally female representation - on women's political presence and empowerment (e.g. Baltrunaite et al., 2019; Baskaran and Hessami, 2018; Bhalotra et al., 2018; Brollo and Troiano, 2016; Casas-Arce and Saiz, 2015; Beaman et al., 2009;), competence/quality of future candidates and elected politicians (for example, Bagues and Campa, 2021; Bagues and Campa, 2020; Besley et al., 2017; Baltrunaite et al., 2014), and allocation of public expenditure (Lippmann, 2022; Bagues and Campa, 2021; Braga and Scervini, 2017; Rigon and Tanzi, 2012; Ferreira and Gyourko, 2014; Chattopadhyay and Duflo, 2004; Clots-Figueras, 2012; etc.).

Bagues, 2012; Baltrunaite et al., 2014). The unique nature of South Korea's election system is what allows us to disentangle the immediate and follow-up responses, uncovering the story of learning about female competence.

The remainder of the paper is organized as follows. Section 2 provides a background on the institutional setting of South Korea's municipal council elections. We then describe the data in Section 3. Section 4 lays out our empirical strategy, and Section 5 discusses the results. In Sections 6 and 7, we present a model and we discuss the pieces of evidence that point towards learning as an explanation for the results. Finally, Section 8 concludes.

# 2 Institutional setting

## 2.1 The role of municipal councils

There are 226 municipal councils in South Korea. Municipal councils represent the legislative branch that works with municipal governments, the executive branch, to oversee local matters. Councils have several legally defined responsibilities, which include reviewing and approving the spending of municipal governments, adopting and revising local bills, monitoring the municipal governments' administrative functions, and examining petitions submitted by residents. Municipal governments administer around a third of South Korea's total public expenditure (Ministry of the Interior and Safety, 2018).

## 2.2 Electoral rules and gender quotas

Municipal councils were established during the mid-1990s, and from then, elections have taken place every four years. Our sample covers seven elections, with 2018 being the last election year. Up to the third election in 2002, all councilors were directly elected through plurality vote in single-member constituent wards. It was extremely rare to find candidates affiliated with a political party.

However, major reforms were made to the electoral rules from the fourth election in 2006. They are summarized in Table I. First, a parallel voting system was introduced. Candidates could be elected through two alternative arms[4]: a party-list proportional representation arm (PR arm) and a plurality vote arm (Ward arm). At least 10% of the councilors needed to be elected through party-list proportional representation. Hence, the number of councilors elected through the PR arm increased as a step function of the total council size: 1 for councils with up to 10 seats, 2 for those with 11 to 20 seats, 3 for those with 21 to 30 seats, and so on.

The remaining seats were reserved for plurality voting in multi-member constituent wards.

---

[4]The same person could only be listed as candidate for the party in one of the two arms.

Each constituency elected between 2 and 4 councilors, and therefore multiple candidates from the same party could run in the same constituency.[5] Figure I illustrates what the ballot papers look like for the two arms of the municipal council elections.

Second, a gender quota was put in place: all odd-number candidates in the party list for the proportional-representation arm needed to be female. As the numbers of seats reserved for proportional representation are small, most elected councilors turned out to be the first candidates in the lists, and therefore female. As a consequence, the introduction of quotas sharply increased the proportion of female councilors. Municipal councils were severely male-dominated prior to the reform, with only 3% of councilors being female. Due to the introduction of quotas in 2005, female representation in municipal councils reached more than 30% in the last election in 2018. Figure II illustrates how the female ratio developed in municipal councils over time. The most striking feature is the sharp rise in the female ratio immediately after the reform.

Last, subsidies were offered to parties based on the female ratio among the parties' candidates nationwide. However, it is unlikely that the subsidies affected much of the political parties strategies, particularly at the municipality level. The scale of the subsidies has been criticized for being too low to effectively expand female nomination (Jin, 2018; Kim et al., 2003; Lee, 2003). Indeed, they account for only around 5 to 6% of the total value of election subsidies (National Election Commission, 2018). Therefore, the presence of the subsidies are unlikely to have impacted political parties' selection of candidates.

Amendments to electoral rules continued between the 2006 and 2010 elections. It was stipulated that in either the municipal council elections or the higher-up provincial council elections, there must be at least one female candidate in each general election district. As there are around 250 general election districts, compared to 226 municipalities, a general election district approximately compares to a municipality.[6] Legislative Impact Analysis Reports indicate that most parties chose to satisfy this rule in the municipal council elections, due to the larger number of candidates (Lee, 2019). Selecting which ward to place the female candidate would have been a strategic concern for the political parties.

## 2.3 Background behind the adoption of gender quotas

If some parties had led the move for the reform against opposition from other parties, then we should recognize that parties' strategic responses to the quota might be very heterogeneous in nature. Thus, here we discuss the background behind the adoption of the quota.

---

[5]The maximum number of candidates a party could nominate for a ward equalled the preset number of seats for that ward.

[6]General election districts are divided depending on population size and local representativeness. A large municipality may contain five general election districts, and up to five small municipalities may comprise a general election district.

Before gender quotas were adopted in the municipal council elections, they were adopted first in the general election for the National Assembly in 2004. The adoption was influenced by increasing demands by women's organizations to raise female representation in politics, which at the time was dramatically behind the international average.[7] As females constitute half the voters, it was in the interest of political parties to put gender quotas forward amongst their election pledges. Moreover, Jeon (2013) argues that the adoption of the quota was also a political tactic. Political parties wanted to increase the size of the National Assembly back to what it was before the size cut during the Asian Financial Crisis, and the fact that the majority of the added seats will go to females, with the quota, made for a good excuse to expand the Assembly.

Once the quota was adopted in the general election, it became the natural next step to introduce it in the regional elections. The gender quota in the municipal council election was passed in the National Assembly, led by both major parties. Some argue that there was political motivation behind it, too (Kim, 2005). One new element in the reform was the party nomination system – a ward candidate must be nominated by their party in order to run with the party affiliation – but it was disputed as a ploy to deepen party influence. Political parties used the quota to justify the party nomination system, since the gender quota was embedded in the proportional representation arm where party nomination was essential.

To sum, it is difficult to say that there was a major division among political parties in their support of the gender quota when it was passed.

# 3  Data

Two sources of data are used. First, data related to the execution of the elections are collected by web scraping the website of the National Election Commission. The website posts detailed data on all past elections, including population, candidate information, and vote outcomes. Second, to examine the consequence of the municipal councils' legislative activities, we use the data on municipal governments' expenditures from the Local Finance Disclosure System of the Ministry of the Interior and Safety.

## 3.1  Population

Because ward divisions are centrally determined based on population size, population data is published. The number of residents is available by ward, voting eligibility, gender, and citizenship status. Moreover, the data includes the number of households by ward. This data is used to perform

---

[7]See Cho and Kim (2010) for a summary of the major activities of women's organizations.

balancing checks in order to validate the identification strategy, which relies on the assumption that municipalities locally around the PR seat thresholds are similar.

## 3.2 Candidates

Various background characteristics of all candidates are also made publicly available by the National Election Commission. These are election arm (ward or PR) classification, election district name, candidate number, party affiliation, name, gender, date of birth, age, occupation, education, and pertinent work experience. Whether a candidate is favored by his or her party is revealed by the election arm and candidate number. Typically, candidates that are deemed less competitive are placed on the PR election arm, and the candidate numbers directly translate to the position on the ballot, in which higher positions attract more votes.

Figure III illustrates how the female share among candidates have been increasing continually, even when not stipulated by the quota. In particular, plot [b] shows that more females are running in wards as the sole candidates of their parties, and plot [c] shows that more females are taking the highest ballot positions even when multiple same-party candidates are running. Plot [d], on the other hand, shows that more females are taking the even-number party list slots, which would not happen with a strong preference for men.[8]

## 3.3 Votes

The website of the National Election Commission also includes vote counts by ward. These vote counts enable us to see in which wards parties have their strongholds. Therefore, we can categorize wards into safe and contestable ones in the perspective of the political parties. Parties would then allocate their favored and less favored candidates to different wards accordingly.

Moreover, we can learn by which margin the winners won. In the regression discontinuity identification strategy, we rely on the assumption that close victories result in sharp changes in the composition of councilors by party, in an environment where parties enjoy similar degrees of popularity from the voters.

Electoral outcomes determine the gender ratio of the elected councilors. Table II provides descriptive statistics on the gender composition of councils by election cycle. The table also depicts how the reform in 2005 introduced the PR arm as well as the gender quota in that arm.

---

[8]In the latest general election of 2020, where the same gender quota on the PR arm applies, almost all PR candidates in positions 2, 4, and 6 are male.

## 3.4 Municipal budget

Municipal budget data is used to perform balancing checks, to show that municipalities locally around the PR seat thresholds are similar in terms of economic scale and council performance. The budget of a municipal government reflects the economic prosperity of the municipality, as around a half is sourced from local tax and non-tax revenue. In addition, data is available on the share of the municipality's expenditures spent on running the municipal council (2002-2020). There have been numerous accusations in the past of councilors appropriating large sums of the local budget for their private use (Local Decentralization Bureau – Election and Local Council Division, 2019). For instance, they would go on international policy-research trips where the itinerary largely consists of sightseeing. Another example is of councilors ordering member pins made of pure gold. As such, a measure of the performance of a council is the frugality of its operation costs. Newspapers have traditionally included it in their assessments of councils (Jang, 2008).

# 4 Empirical Strategy

## 4.1 Regression discontinuity design around the number of PR seats

To get at the causal effect of the gender quota, we make use of the fact that the gender quota affects municipalities at different intensities depending on the number of PR seats in the council. To remind ourselves, the number of PR seats is important as the gender quota only applied to the PR arm, and it stipulated that all odd-number candidates in the party list for the proportional-representation arm needed to be female. Since, at least 10% of the councilors needed to be elected through party-list proportional representation, the number of councilors elected through the PR arm increased as a step function of the total council size, which is pre-determined centrally by the National Election Commission based on population size and regional representativeness. The step function is depicted by the navy dots in Figure IV.

This provides a discontinuity in the number of PR seats at given thresholds of council size. For all councils with size smaller or equal to 10 seats, at least one councilor must be elected through proportional-representation. For councils with a number of seats higher than 10 but smaller or equal to 20, at least two councilors must be elected through proportional-representation, and so on and so forth.

We exploit this discontinuous change in the number of PR seats as a function of council size in a regression discontinuity design that compares the characteristics of candidates in municipalities on each side of the thresholds, while controlling for council size. In order to account for the fact

that there is not just one but many thresholds (10, 20, 30), we categorize councils into bins based on the proximity to thresholds, as illustrated in Figure IV.

This strategy identifies the effect of an additional PR councilor, rather than an additional *female* PR councilor. However, in this setting, an additional PR councilor strongly implies an increase in the number of female PR councilors. While the gender quota does not necessitate that the second PR councilor be male, in practice almost all PR councilors end up being female.[9] This is due to PR candidates even in even-number positions frequently being female, and elected PR councilors frequently being the number-1 candidates of multiple parties.[10] Evidence that this is the case is provided in Section 4.2.

The regression discontinuity specification is given by:

$$Y_{cbt} = \alpha_b + \alpha_t + \sum_{s=4}^{7} \beta_s \times Treat_{cbt} + f(x_{cbt}) + \epsilon_{cbt} \tag{1}$$

where $Y_{cbt}$ denotes the outcome variable for municipal council $c$ belonging to bin $b$ in election cycle $t$. As we are interested in characterizing parties' strategies in the selection of candidates, the outcomes we consider are the number of ward and PR candidates and councilors by gender. The running variable is $x_{cbt} \equiv (\text{council size})_{cbt} - \text{threshold}_b$, with $\text{threshold}_b \in \{10, 20, 30\}$. In addition, $Treat_{cbt} \equiv \mathbb{1}(x_{cbt} \geq 0)$, signifying an additional PR councilor. Therefore, $\beta_s$ estimates the effect of having an additional PR councilor, pooling all the bins together, in election cycle $s$. Moreover, the baseline function form of $f$ is linear, and we do not allow for the effect of $x_{cbt}$ to differ to the left and right of the threshold. The reason for this choice is that making $f$ quadratic or allowing for differential trends on either side of the threshold barely makes a difference, as it can be seen in Table A5, and Table A6 in Appendix.

A factor to note is that when the outcome variable relates to ward elections, we change the running variable to $\tilde{x}_{cbt} \equiv (\text{N. of ward councilors})_{cbt} - (\text{N. of ward councilors at the threshold})_b$ for ease of interpretation. To illustrate why this is important, let's consider councils just above and below the threshold of 10 councilors, 11 and 10 councils respectively. When $Treat = 0$, one councilor must be elected through proportional representation, which implies that 9 remaining councilors are elected through the ward arm. At the same time, when $Treat = 1$, two councilors must be elected through proportional representation, which implies that 9 remaining councilors are elected through the ward arm. This is not taken into account when we control for a running variable based on council size. As council size differs by one unit above and below the threshold, controlling for council size would effectively induce us to compare a council where $Treat = 1$ with one fewer ward councilor than the one with $Treat = 0$, therefore making more difficult

---

[9]Table II shows that among PR councilors, 87% to 98% are female each election.

[10]Appendix Table A1 shows that it is relatively rare to find multiple PR seats getting allocated to the same party.

to interpret the sign of the coefficient on $Treat$. Redefining the running variable addresses this problem.[11]

Besides equation (1), we consider a second specification based on treatment status at election cycle 4:

$$Y_{cbt} = \alpha_b + \alpha_t + \sum_{s=4}^{7} \beta_s \times (\textit{Treat at cycle } 4)_{cb} + f(x_{cb4}) + X'_{cbt}\gamma + \epsilon_{cbt} \qquad (2)$$

where $(\textit{Treat at cycle } 4)_{cb} \equiv Treat_{cb4}$, and $X_{cbt}$ denote control variables such as council size or the number of ward seats.

The difference between equations (1) and (2) is that they estimate the effects of contemporaneous treatment and initial treatment, respectively. This distinction is important if councils change size and, as a consequence, treatment status across election cycles. In practice, it makes barely any difference which specification we use, because the treatment status changes after election cycle 4 for only 3.7% of the councils.

For the analysis, we settle on equation (2) as our main specification for two reasons. First, the initial treatment assignment is more exogenous. Upon the first treatment, treated and control municipalities may evolve on different paths, causing council size and treatment status to change, making treatment and control group no longer balanced. The second reason has to do with the possibility to compare the estimated treatment effects across election cycles and identify the long-term effect of quotas. Measuring the effect of the initial treatment assignment maintains the same composition of treated municipalities across election cycles. On the other hand, defining treatment status by exploiting contemporaneous treatment assignment wouldn't allow us to attribute variations in treatment effect over time to differences between immediate and follow-up effects for the same councils. As a matter of fact, differences might also be due to a small number of councils changing their treatment status over time. With equation (2), this second channel is shut down.

The standard errors are clustered by municipality. This is motivated by two considerations. First, the variation of the treatment assignment is at the level of the municipality. Second, parties formulate strategies chiefly within a municipality, rather than moving around candidates across municipalities. In fact, there are many factors that tie down a candidate to a certain municipality to be nominated in. A candidate is legally required to have been a resident of the municipality they are running in for at least 60 days prior to the election. In addition, as municipal councilors deal with local grass-roots matters, a candidate familiar with the municipality will win more votes *ceteris paribus*. Hence, a candidate usually runs in the municipality they have a connection with, such as their birthplace, long-term residence, or place of education. Moreover, the final say of a party's

---

[11]Changing the running variable this way does not change much else. In fact, the coefficients $\hat{\psi}_0$ and $\hat{\psi}_1$ stay the same, as well as the R-squared value.

nomination lies on the head of the municipal branch of the party, so a candidate typically serves the local activities of the party in the municipality they desire to run in for a long time before getting nominated. Finally, once a candidate is nominated in a municipality, they put on a campaign and become known to the residents. So if they were to run again, they would not start over at a new location. For all these factors, rarely do parties move around candidates across municipalities for strategic reasons.

## 4.2   Did the quota bite?

Because we are interested in the consequence of the change in the gender composition of councilors brought about by the quota, it is important to verify that there is a change in the number of *female* PR councilors at the discontinuity thresholds.

Table III reports the results of equation (2) with the number of female PR councilors as the outcome variable, separately for each bin. Two patterned can be noticed. First, having an additional PR councilor at cycle 4 significantly increases the number of female PR councilors over all the cycles at bins 1 and 2. However, there is no such effect at bin 3, where we also have very few observations. The regression results of Table III are echoed by Figure V, which shows that the average number of female PR councilors sharply increase at the thresholds of bins 1 and 2, but not at bin 3. Therefore, in the reduced-form results that follow, we restrict the sample to bins 1 and 2.

Second, for both bins 1 and 2 (columns 1 and 2 of Table III) the effect of the treatment at cycle 4 on the number of female PR councilors remains constant across election cycles. This is not surprising as the vast majority (96.3%) of the initially treated municipalities continue to get treated each cycle. However, it is important as it confirms that if we see an evolution of the treatment effect on party strategies over time, this could be safely attributed to the initial treatment leading treatment and control groups on different paths.

As a way to buttress the validity of the regression discontinuity design, Appendix Section C.1 formally tests and confirms that as council size increases, there is a change in the number of female PR councilors only at the thresholds and at no other point.

## 4.3   Validity of the regression discontinuity design

The critical identifying assumption behind the identification strategy is the smoothness of the relationship between the outcome variable and councils size apart from the discontinuity of interest. This section provides evidence in support of this assumption.

*Balance Tests*   The first piece of evidence concerns the absence of discontinuities in pre-determined characteristics at the threshold in cycle 4. The sample consists of councils at election cycle 4, and

the results of estimating equation (2) are presented in Table IV.

Panels A, B, and C confirm that the population characteristics are indeed balanced. In particular, the voting age population by gender is no different, alleviating the concern that the preference for female councilors among voters may be different between the treated and control municipalities. Furthermore, education and labor force participation by gender are not systematically different.[12] In Panel D, columns (1) and (2) refer to the vote share received by each main party in the previous election's PR arm. Columns (3) and (4) show that the initial treatment group is balanced in terms of economic prosperity and council performance. Columns (5) and (6) demonstrate that the structure of the ward election arm is balanced, as there is no difference in the number or size of wards between the treatment and control municipalities.

*Bunching*    The second piece of analysis examines the density of council size at the threshold, looking for bunching in its distribution. The concern here is that councils might be able to manipulate their constituent areas to affect their council's size and, therefore, influence their treatment status.

Figure VI displays the histogram of the frequency of municipalities by council size. Visually, it is hard to say there is bunching around the thresholds of 10 and 20. Due to the coarseness of the council size variable, even the discrete version of the McCrary (2008) density test proposed by Frandsen (2017) does not perform well; while no bunching is rejected at the thresholds of 10 and 20, it is also rejected for randomly selected cutoffs of council size.

Nonetheless, the evidence from the previous balance checks and existing electoral rules support the hypothesis that municipalities do not manipulate their council size around the threshold. First, municipalities are found not to be systematically different below and above the threshold, which we would expect to find if manipulation was possible. Second, electoral rules make it difficult to gerrymander in practice. The division of election constituencies is determined by the Municipal Council Election Committee, which municipal councilors or party members are not allowed to be on.[13] The committee determines the council size based on population, administrative districts, topography, transportation, and other conditions. The committee also cannot split the smallest administrative district and make it a part of another ward. In sum, strict rules prevent the membership of interested individuals on the committee and circumscribe how the election constituencies are drawn up.

*Placebo Test*    Section 5.4 presents the results of a placebo test where the discontinuity in council size had no effect on the gender composition of candidates and elected councilors in the three

---

[12]The data for panels B and C are at the province level (16 provinces), so the results there should be taken with more caution.

[13]The committee consists of up to 11 members appointed by the district mayor among the individuals nominated by the media, legal and academic communities, civic groups, the district council, and District Election Committee.

election cycles before the introduction of the quota.

# 5    Main Results

## 5.1    Did the reform have the intended effect? The evolution of councils' gender composition

The results of equation (2) on the number of councilors elected in each municipality are reported in Table VI. Two interesting patterns emerge.

First, in the first cycle after the introduction of the quota - cycle 4 - treated municipalities display an overall higher number of elected female councilors (columns 7-9), but the effect is not significantly different from zero. Had the quota not been effective in fostering an increase in the number of elected female councilors? Columns 5 and 6 provide evidence that this is not the case. A higher number of women gets elected in treated councils through the PR arm due to the introduction of the quota. In particular, in treated councils, 0.76 additional women are elected for every additional PR seat. However, in treated councils fewer women and significantly more men are elected in cycle 4 through the ward arm, the arm unaffected by the quota (columns 1-3).

Second, this pattern is not constant over time. Starting from election cycle 5 onward, treated municipalities display significantly more female councilors and significantly fewer male councilors (columns 7-9). As column 5 and 6 confirm that the same number of additional women is elected through the PR arm as a result of the quota, this reversal cannot be attributed to the intensity of the quota effect changing over time. On the contrary, the evidence indicates that this can be traced back to a change in the composition of councilors elected through the unconstrained arm unaffected by the quota. As displayed in Columns 1-3, fewer men and more women are elected through the ward arm starting from cycle 5. A pattern that becomes stronger across election cycles.

## 5.2    Parties' strategies and candidates' selection

What is driving these results on elected councilors? Why did a fewer number of women get elected in cycle 4 through the unconstrained ward arm? Why does this pattern gradually reverse over time? Two potential not exclusive explanations exist. On the one hand, voters might have initially displayed a preference for male candidates, becoming more female-friendly in their endorsement of female candidates over time. On the other hand, parties might have initially "counteracted" the quota by placing fewer female candidates in the unconstrained arm, gradually reversing their behavior across election cycles. Parties in treated municipalities have a higher probability of seeing their female candidates elected through the PR arm due to the higher number of available PR seats

in the council. As a consequence, they might have initially strategically decided to put forth *fewer* female candidates in the unconstrained ward arm, gradually changing their strategies over time.

The evidence points towards parties' strategic responses being the main explanation. In Section 7.5, we show that the evidence is not consistent with voters having an initial preference for men, nor with preferences for females in treated municipalities changing faster over time. If the quota induces a change in voters' attitudes towards women, we expect the gender gap in votes obtained by candidates in treated municipalities to close over time compared to the gender gap of candidates with comparable characteristics, placed in the same ballot position in control municipalities. We find no evidence of such a pattern. On the other hand, we find evidence of parties in treated municipalities putting forth fewer female candidates in the unconstrained ward arm in cycle 4, especially in positions in the ballot list with high probability of election. Furthermore, the evidence points towards the same parties in treated municipalities changing their strategies over time.

Table VI displays the results of equation (2) on the gender composition of the candidates put forth in each municipality in each arm. The most interesting result is captured by columns 1 and 2. In response to the treatment at cycle 4, parties initially put up more male ward candidates but gradually decrease the number across election cycles. Eventually, at election cycle 7, the parties in the treated municipal councils put forth fewer male candidates than those in untreated councils. As for female ward candidates, the opposite pattern holds: the coefficient sign changes from negative (albeit statistically insignificant) in cycle 4 to positive from cycle 5 onward.

## 5.3   Focusing on candidates likely to get elected

Table VI provides evidence on the composition of the overall pool of ward candidates put forth in each municipality in each election cycle. However, considering the municipality as a whole, pooling all participating parties, obstructs the study of dynamic changes in party strategies when many small parties emerge and soon disappear. Moreover, changes in the composition of ward candidates may not affect the composition of political bodies if they are driven by candidates in positions that have no hope in getting elected.

Thus, we next turn our focus to candidates from the two main parties and in ballot positions characterized by a high probability of election. The two main parties – the Conservative party and the Progressive party – dominate South Korean local elections, producing at least 74% of ward councilors and 82% of PR councilors every election (Appendix Table A2). To determine positions likely to win, we define two concepts: (i) useful positions: either the first or the only candidate of the party, based on the observation that within-party, candidates in high-up positions on the ballot win more votes; (ii) safe wards: wards where a party has a stronghold, i.e. the party won with over 10 percentage points higher vote share than the next popular party in the previous election cycle.

Table VII shows that when we restrict our attention to candidates for whom election is probable, not only do we see the same patterns, but the patterns are even stronger. Main parties in treated municipalities put forth a higher number of male candidates and a smaller number of female candidates in cycle 4 (Columns 1-3), especially in pivotal positions in the ballot, where the probability of election is higher (Columns 4-6). However, the initial counteraction in the unrestricted ward arm is reversed already from cycle 5.

As political parties can choose how many candidates to champion[14], Columns 3 and 6 report how the number of female ward candidates compares between treatment and control municipalities, when we control for the total number of ward candidates. One may argue that the total number of ward candidates is a "bad control" in the regression because it is an outcome of the treatment (Angrist and Pischke, 2009). However, as the total number of ward candidates equals the sum of the number of male and female ward candidates, Columns 3 and 6 can inform us regarding whether the additional woman (man) is placed in the ballot list as a substitute of men (women) or in addition to men (women). We can clearly see that *substitution* between the two genders occurs.

Table VIII uses a different way of gauging the election probability of candidates. It shows how parties select candidates for different types of wards. We categorize wards based on whether the party had a stronghold in the previous election, in which case we call the ward "safe."[15] These are wards where we can assume the party candidates have a very high probability of being elected. Since whether a ward is safe is dependent on the party at hand, the regressions in Table VIII are at the ward×party level. We can see in Column 5 and 7 that parties had a strong preference for placing male candidates in safe wards in cycle 4, especially in useful seats. The preference for men, however, disappears from the cycle 5 onward. Furthermore, we can see in columns 6 and 8 that, from cycle 5, parties in treated municipalities start placing more women among candidates in unsafe wards. This is what is driving the overall increase the number of female candidates we observe at the municipality level. Therefore, although the number of female ward candidates increases faster over time in treatment than in control municipalities, the gains for women still remain bounded to the less-preferred wards with lower likelihood of election.

## 5.4 Placebo test

In the last sections we showed that treated municipalities, i.e. municipalities above the threshold where the quota intensity is stronger as an additional councilor is elected through the regulated PR arm, put forth a different number of female and male candidates compared to municipalities below

---

[14]The maximum number of ward candidates for a party is the total number of ward seats in the council, but there is no minimum.

[15]A party is considered to have a stronghold in a ward if the party wins the greatest vote share in the PR arm in the ward, and it got over 10 percentage points more vote share than the next popular party. For cycle 4, the margin of victory is calculated based on district elections as there is no previous PR election.

the threshold where this arm is less important. Before going into the potential mechanisms that can explain the effects we observe, we report the result of a placebo test where we check that the thresholds are meaningful only after and not before the reform to the election system.

We test whether the number of male and female ward candidates changed at the threshold before and after the reform. Before the introduction of the quota in 2006, there was no proportional representation arm and all the candidates were elected through plurality voting (as in the ward arm after the introduction of the quota). Hence this test allows us to assess whether the probability of getting an additional PR seat upon the reform is correlated with other factors that affect the number of male and female candidates. Table IX shows that up to election cycle 3, the effect of being past the threshold is not statistically significantly distinguishable from zero. It is at election cycle 4 that the treatment induces an effect, as expected. This test provides supportive evidence that we are estimating the effect of the introduction of the quota, and the results are not driven by treated municipalities being different from control municipalities ex-ante.

# 6   A simple model of party learning

The evidence provided so far indicates that parties reacted to the introduction of the quota in the PR arm by reducing the number of women among ward candidates immediately after the reform. The substitution away from women was stronger in ballot positions and wards where candidates had higher chances at getting elected. However, from cycle 5 onward, parties in municipalities above the threshold started placing more women among ward candidates.

We propose that these results can be reconciled with the quota allowing parties to learn about women's competence through exposure. In this section, we lay out a simple model of learning. Parties start with downward biased beliefs about women's competence, and consequently select a suboptimally low number of women. However, they gradually update their beliefs toward the truth as the policy forces them to be exposed to female politicians. Guided by the predictions of the model, we present in section 7 the empirical evidence consistent with learning being the key mechanism behind the observed reversal in candidates' selection and rule out alternative explanations.

## 6.1   Setting

Parties $p = 1, ..., P$ participate in a municipal council election in election cycle $t$. A party selects from a group of prospective candidates those to put up for election. For the selected candidates, the party also allocates them to different candidate positions, which are ranked by some historic probability of election. Prospective candidates differ solely in their gender $g \in \{M, F\}$ and their

competence $a_i$ (Besley et al. 2017):

$$a_i \sim N(\mu_g, \sigma^2)$$

where $\mu_g$ is the mean ability of gender g.

Parties have imperfect information regarding the competence of a new candidate. The true competence of the prospective candidate is unknown, but parties can observe a signal of competence:

$$s_i = a_i - \mu_g + \epsilon_i, \quad \epsilon_i \sim N(0, \sigma_s^2)$$

$E(s_i) = 0$, therefore the signal is informative about the *relative* competence of $i$ within gender $g$.

## 6.2 Party's problem

A party has only one objective: maximize the number of elected councilors from the party. We assume that voters vote for parties as opposed to individual candidates[16], and that they choose a party based on inherent party ideology and the ability and gender of the candidates from the party. Therefore, for electoral success, party $p$ chooses candidates among the pool of available prospective candidates to maximize

$$E(a_{p,t}) + h(f_{p,t})$$

where $a_{p,t}$ denotes the average ability of councilors from party $p$ in election cycle $t$, and $f_{p,t}$ is the share of female councilors from party $p$. $h(.)$ is a strictly concave function which attains its maximum at $f_t^* \in (0,1)$ and denotes voter preferences for the gender composition of councilors. Put simply, the party wants to maximize the expected average ability of councilors while achieving some gender ratio close to $f_t^*$.

That voters care about councilor competence is a critical feature of the model. It is in line with the evidence on election probabilities of candidates before the introduction of the quota, where voters indeed reward candidates' competence. More details can be found in Appendix D.1.

## 6.3 Party's beliefs at candidate selection

As parties do not have perfect information regarding the ability of new candidates, they also hold imperfect information regarding the mean ability of the group $\mu_g$. A party's prior belief about the

---

[16]This assumption is made for simplification and is based on the observation that even in the ward arm, voters tend to vote for the candidate from their preferred party. To see this, we first predicted the vote share of a candidate in the ward arm based on the popularity of his/her party – measured by the party's vote share in the PR arm – and the historical tendency of voters to vote more for candidates higher up on the ballot paper. The correlation between the actual vote share and the predicted one is 0.90.

value of $\mu_g$, before the election in election cycle $t$, follows a normal distribution with mean $\tilde{\mu}_{g,t}$, variance $\tilde{\sigma}^2_{g,t}$.

For a prospective candidate $i$ with signal $s_i$, the party expects $i$'s ability to be

$$\tilde{E}(a_i|s_i, g, t) = \tilde{\mu}_{g,t} + \frac{\sigma^2}{\sigma_s^2 + \sigma^2} s_i$$

where $\tilde{E}$ indicates expectation taken over the prior distribution.

The party selects candidates taking into account the value of $\tilde{E}(a_i|s_i, g, t)$. Within gender, it chooses the prospective candidate with the highest value of $s_i$ first, then moves on to the one with the next highest value of $s_i$, etc.

Because the signals provide information on only the *relative* ability within gender, these signals are not informative about the value of $\mu_g$. Hence, the party cannot find out about the average ability of each gender $g$ from the set of signals of the prospective candidates.

## 6.4 Party learning

Once candidate $i$ is elected and serves as councilor, the party learns about his/her competence. We assume for simplicity that learning is complete and the true competence of the candidate $\mathbf{a} = \{a_i\}$ is revealed.[17]

Given the observed abilities, the party also learns about the average ability of each gender $\mu_g$. The party makes an inference about the value of $\mu_g$ via maximum likelihood. In doing so, parties use the fact that these councilors are positively selected out of the pool of prospective candidates, and they had previously received initial signals $\mathbf{s} = \{s_i\}$.

Conditional on $s_i$, the distribution of $a_i$ is:

$$a_i|s_i \sim N\left(\mu_g + \frac{\sigma^2}{\sigma_s^2 + \sigma^2} s_i, \frac{\sigma^2 \sigma_s^2}{\sigma_s^2 + \sigma^2}\right) \tag{3}$$

Let's define $c = \frac{\sigma^2}{\sigma_s^2 + \sigma^2}$, $m(s_i) = \mu_g + \frac{\sigma^2}{\sigma_s^2 + \sigma^2} s_i = \mu_g + c s_i$, $\bar{\sigma}^2 = \frac{\sigma^2 \sigma_s^2}{\sigma_s^2 + \sigma^2}$.

Then the likelihood function is

$$\begin{aligned}
\mathcal{L}(\mu_g) &= P(\mathbf{a}|\mathbf{s}; \mu_g) \\
&= \Pi_{i=1}^n f(a_i|s_i; \mu_g) \quad \text{where } f : \text{Gaussian probability density function} \\
&= \frac{1}{(\sqrt{2\pi}\bar{\sigma})^n} \exp\left(-\frac{1}{2\bar{\sigma}^2} \sum_{i=1}^n \left(a_i - m(s_i)\right)^2\right)
\end{aligned}$$

---

[17]This assumption could be relaxed. An extension of the model with imperfect belief updating regarding councilors' competence can be found in Appendix D.2.

The maximum likelihood estimator is

$$\hat{\mu}_g = \frac{1}{n} \sum_{i=1}^{n} (a_i - cs_i) \sim N\left(\mu_g, \frac{1}{n}\bar{\sigma}^2\right) \tag{4}$$

Call $V = Var(\hat{\mu}_g)$. Then, the posterior distribution about the value of $\mu_g$ is normal with mean $\tilde{\mu}_{g,t+1}$ and variance $\tilde{\sigma}_{g,t+1}$, which are weighted averages of the prior and the maximum likelihood estimator:

$$\tilde{\mu}_{g,t+1} = \frac{V\tilde{\mu}_{g,t} + \tilde{\sigma}_{g,t}^2 \hat{\mu}_g}{V + \tilde{\sigma}_{g,t}^2}$$

$$\tilde{\sigma}_{g,t+1}^2 = \frac{V\tilde{\sigma}_{g,t}^2}{V + \tilde{\sigma}_{g,t}^2}$$

The updating speed of party beliefs about the value of $\mu_g$ is given by

$$\tilde{\mu}_{g,t+1} - \tilde{\mu}_{g,t} = \frac{\tilde{\sigma}_{g,t}^2}{V + \tilde{\sigma}_{g,t}^2}(\hat{\mu}_g - \tilde{\mu}_{g,t})$$

## 6.5 Comparative statistics

***Speed of updating***: Updating is faster

1. The larger the number of female councilors:

$$\frac{\partial(\tilde{\mu}_{g,t+1} - \tilde{\mu}_{g,t})}{\partial n} > 0$$

2. The higher the ability of female councilors:[18]

$$\frac{\partial(\tilde{\mu}_{g,t+1} - \tilde{\mu}_{g,t})}{\partial a_i} = \frac{\partial(\tilde{\mu}_{g,t+1} - \tilde{\mu}_{g,t})}{\partial\hat{\mu}_g}\frac{\partial\hat{\mu}_g}{\partial a_i} > 0$$

3. The larger the (downward) bias in the prior belief:

$$\frac{\partial(\tilde{\mu}_{g,t+1} - \tilde{\mu}_{g,t})}{\partial\tilde{\mu}_{g,t}} < 0$$

***Selection of candidates***: A party selects more male candidates

---

[18]The exact comparative statistics that match the empirical evidence of the learning mechanism in Section 7.2 is slightly different, although capturing the same idea. It is: Updating is faster the higher the signal of female councilors (unconditional on the ability)

$$\frac{d(\tilde{\mu}_{g,t+1} - \tilde{\mu}_{g,t})}{ds_i} = \frac{\partial(\tilde{\mu}_{g,t+1} - \tilde{\mu}_{g,t})}{\partial\hat{\mu}_g}\left[\frac{\partial\hat{\mu}_g}{\partial s_i} + \frac{\partial\hat{\mu}_g}{\partial a_i}\frac{da_i}{ds_i}\right] > 0 \tag{5}$$

1. The larger the difference in the male and female prior mean ($\tilde{\mu}_{m,t} - \tilde{\mu}_{f,t}$):

$$\frac{\partial\big[\tilde{E}(a_i|s_i, M, t) - \tilde{E}(a_i|s_i, F, t)\big]}{\partial\big(\tilde{\mu}_{m,t} - \tilde{\mu}_{f,t}\big)} > 0$$

(numerator: gap in the perceived ability of men and women with the same signal $s_i$)

# 7   Evidence on learning about women's competence

The model formalizes the idea that the change in candidates' selection could be explained by parties learning about women's competence by being forced to experience female councilors after the introduction of the quota. With females holding only 3% of seats before the quota and more than 60% of the population agreeing with the statement that men make better political leaders than women do (see Figure A2), it is quite likely that party leaders might have started with imperfect information and biased beliefs regarding the competence of women as politicians.

In this section, we present the empirical evidence that supports this hypothesis. We start by checking that the model predictions are confirmed in the data in sections 7.1-7.4. Together, these sections paint a convincing picture of the learning mechanism. Then we rule out three alternative mechanisms. Firstly, the increase in the number of new women cannot be completely reconciled with the supply of available qualified female candidates growing faster in treated municipalities compared to control municipalities. Secondly, the evidence is not consistent with parties responding to a change in voters' preferences for females candidates. Thirdly, initial taste-based discrimination and parties growing out of their distaste for women faster in treated municipalities is not a mechanism supported by the evidence.

## 7.1   Learning is faster with more new information on women

The model predicts that the process of belief updating is faster the greater the number of newly elected female councilors. In this section we check if this prediction is empirically met by looking at the effect of being above the threshold on the number of elected rookie and incumbent female councilors. A candidate or a councilor is defined as incumbent if they have been elected in at least one of the previous election cycles.

Table X shows that the quota was effective in promoting the election of a higher number of rookie female councilors in treated municipalities, exposing parties to new information about women's competence as political leaders. Comparing Columns 2 and 5, we can clearly see that the additional women elected through the PR arm due to the stronger intensity of quota in treated municipalities are predominantly rookie women. In fact, up until cycle 5, these additional women

22

are solely rookie women. As a consequence, despite parties' counteraction of the quota resulting in a lower number of female rookie ward councilors (Column 4) in cycle 4, the *total* number of female rookie councilors in treated municipalities is higher overall (Column 6), even if not significantly so. Hence, the quota was successful in forcing parties to experience new female politicians in action.

## 7.2 Learning is faster the better the women experienced

The second prediction of the model indicates that if parties are learning about the competence of females, then this learning process takes place at a faster rate if they are exposed to more competent women. To provide evidence in this direction, we perform a heterogeneity analysis based on the level of education of the first elected PR women of cycle 4 and track the evolution after cycle 4. As shown in the previous section, the quota promoted the election of a higher number of rookie female councilors through the PR arm in treated municipalities. We now check whether the competence of these PR women matters.

Following an approach commonly used in the literature, we exploit information on the education of the candidates as a proxy for competence[19]. We divide municipalities into two groups by whether the average years of schooling of the PR women elected in the municipality in cycle 4 are above or below the median of all municipalities. We next examine if the evolution over time of the treatment effect (effect of an additional PR seat, and hence an additional woman) on the selection of ward candidates differ in the two groups.

Table XI displays the results of equation (2) for the sample of councils in which the female PR councilors elected in cycle 4 have below-median and above-median years of schooling, respectively. To ensure that any difference in the evolution of the treatment effect over time cannot be attributed to ex-ante differences in the overall pool of women available, Columns 2, 4, 6, and 8 add controls for the average education level of all female candidates affiliated with the two main parties in cycle 4 and for the total number of ward female candidates affiliated to two main parties in the municipality in cycle 4. The analysis is again restricted to the main two parties, which are the only parties that we can follow over time.

In line with the prediction of the model, we find that the shift towards female ward candidates is more apparent and stronger when the first elected PR women are more educated (Column 7). This is confirmed also when we control for ex-ante characteristics of the pool of women in the two groups of municipalities (Column 8). This evidence points to the story of learning about the competence of females over time, and not one of a growing taste for women that would occur irrelevantly of ability.

---

[19]For example, Baltrunaite et al., 2014

## 7.3 Parties with greater gender bias learn more about elected women

The previous sections show evidence consistent with the story of learning through exposure; the gradual increase in female candidates is stronger where there is more new information about women's competence. Up to now, the learning story was documented at the municipality level. This section further corroborates the story by unpacking the dynamic learning process at the level of parties. In doing so, we highlight how parties with different ex-ante favorability towards women respond differently to the election of women.

To conduct this *party*-level analysis, we compare the strategies of parties that marginally won or lost the election of their first – and therefore female – PR candidate in the previous election cycle. This comparison gives us the causal effect of experiencing a woman from one's own party in action. In close electoral races in which the outcome of the election is uncertain, the winner is typically determined by factors that are beyond the control of parties and candidates, so which party wins the seat can be considered random (Lee, 2008).

We take marginal parties to be the *two* parties that either marginally won or lost a PR seat for the municipality. In order to differentiate marginal winners from losers, we measure how far off the vote share received by a party was, from the share it needed to win that seat. For party $p$ in municipal council $c$ at election cycle $t$, this value is given by $v_{cpt} \equiv (\text{vote share})_{cpt} - \bar{v}_{cpt}$, where $\bar{v}$ denotes the verdict-determining vote share. As the simplest example of $\bar{v}$, when two parties are competing for one PR seat, $\bar{v} = 0.5$ for both parties. The precise way we compute $\bar{v}$ for all possible contest scenarios is detailed in Appendix Section E.1. Figure VII shows the distribution of the vote shares received by the two marginal parties competing for the last PR seat in a municipality. The histogram shows that there are plenty of parties that received a vote share close to the share needed to win that seat.

To focus on the effect of experiencing a *woman* from within the party, the marginal parties in our sample are restricted to those whose marginal candidate was the number-1 candidate in the party list.[20] Thus, we compare parties that marginally got their first (female) PR candidate elected in the previous election cycle with parties that marginally failed by random chance.

We employ a regression discontinuity design of the following form:

$$Y_{cpt} = \alpha_t + \beta \times Winner_{cp,t-1} + f(v_{cp,t-1}) + X'_{cpt}\gamma + \epsilon_{cpt} \tag{6}$$

where $Winner_{cp,t-1} \equiv \mathbb{1}(v_{cp,t-1} \geq 0)$. $f(v_{cp,t-1})$ is linear and allows for different slopes to the left and right of the cutoff $v_{cp,t-1} = 0$. $X_{cpt}$ represents the control variables, including the number

---

[20]To clarify, these include parties in municipalities with more than 1 PR seat. For example, if there are 2 PR seats in the council and the contest is whether the most popular party wins both seats or the second-most popular party wins one, too, then only the second-most popular party is in the sample. Its marginal candidate is the first PR candidate on the party list, while the most popular party's is the second.

of ward seats and the total council size for the *contemporaneous* election, i.e. election cycle $t$. A further factor to note is that the sample includes only the two major parties in South Korea, in order to track the parties over time. $X_{cpt}$ also includes a dummy that indicates which of the two major parties party $p$ is.

Critically, we distinguish how the $Winner$ effect varies between two different types of parties: those that placed men as the number-2 candidate in the party list in the previous election, and those that did not. This is important for two reasons. First, as parties can place a candidate of any gender in even positions in their lists, the parties that place women in even positions can be expected to have more gender-equal attitudes ex-ante. Hence, by comparing the effect of marginally winning a woman for parties that placed a woman or a man in the second PR position, we can test whether the evidence is consistent with the prediction of the model that, if learning takes place, belief updating should be faster the larger is the (downward) bias in prior beliefs. Second, the comparison allows us to net out other potential reasons why party strategies change when parties win a seat in general, that affect both types of parties.

Table XII reports the result of specification (6) where the outcome variable is whether the number-1 PR candidate in cycle $t-1$ is one of the ward candidates in cycle $t$. The outcome variable tells us whether the PR candidate gets moved "up" to the ward arm where she has to win elections on her own account, and where her selection is no longer legally imposed. The results are provided for all parties (Panel A), parties for which the second PR candidate was a man (Panel B), and parties where the second PR candidate was a woman (Panel C).[21]

The evidence is consistent with parties learning about a woman's competence by experiencing her in action. Columns 1-4 of Panel A show that the probability that the number-1 PR candidate in cycle $t-1$ runs in the ward arm in cycle $t$ is significantly higher for parties where she actually won. Comparing Panel B to C, we can see that the effect is completely driven by parties that placed a man in the second position on the party list. Hence, this cannot be reconciled with elected women being selected again as candidates because she has acquired experience as a councilor or visibility among voters, as the same is true for the PR woman elected from parties that placed a woman in the second position. Rather, the result shows that more gender-equal parties move women to the ward arm whether or not they were elected previously in the PR arm, whereas less gender-equal parties only do so once she has proven herself after election. We take this result to be consistent with the third prediction of the model: parties with larger (downward) bias in prior beliefs about the competence of women respond more – update their beliefs more – to new information about a woman's capability that gets revealed once she serves as councilor.[22]

---

[21]The effect of marginally winning a PR woman is estimated also using different functional forms for $f(v_{cp,t-1})$. Results are very consistent and can be found in Table A8 in Appendix.

[22]To be precise, the third model prediction is about the updating speed about the value of $\mu_g$, i.e. learning about the competence of women as a whole, as opposed to learning about individual councilors. But the same prediction

## 7.4 Learning takes place for women as a whole

The evidence discussed so far is consistent with parties starting with imperfect information and biased beliefs regarding women as political leaders and subsequently learning about the competence of elected women. The previous section, in particular, focused on how parties respond to new information about the capabilities of individual women that get revealed after they get elected. We argue in this section that parties not only learn about the elected women but also about the group of women as a whole.

We return to the specification of (2) to analyze the effect of being above the threshold on the number and gender composition of incumbent and rookie ward candidates. Table XIII shows the results for the number of incumbents (columns 1-3) and rookies (columns 4-6) respectively. Columns 1 and 3 present the results for male candidates, columns 2 and 4 present the results for female candidates, while the results in columns 3 and 6 can be interpreted as the change in the share of women keeping the total number of candidates constant.

We can clearly see that parties being able to access a larger pool of known and verified experienced female candidates is one of the explanations for the different evolution in the composition of ward candidates in treated councils. As shown in Columns 1 to 3, parties put forth a higher number of incumbent women among ward candidates over time.

However, this cannot be the only mechanism in place. The evidence is consistent with parties learning about the competence of women as a group. Columns 4-6 of Table XIII show that parties in treated municipalities also start putting forth an increasing number of rookie women over time. These are women that the party has not experienced in action and for whom information regarding competence is imprecise. Hence, this points towards a change in the prior regarding the competence of women as a whole. In line with the last prediction of the model, the smaller is the difference in the party's prior regarding the competence of men and women, the greater the female share is among selected candidates[23]

## 7.5 Alternative mechanisms

*Change in the supply of qualified women*    Even if there is absolutely no learning, the patterns that we see in the number of female candidates and councilors could be consistent with the initial shortage and gradual growth in the supply of qualified and experienced women. We provide evidence that this mechanism cannot be the whole story.

---

would hold for learning about individual councilors, if we relax the model simplification that true ability of councilors is fully revealed after election.

[23]In the next section we provide evidence that there is no differential evolution in the gender gap in education in municipalities above and below the threshold, which might be an alternative mechanism.

First of all, for the treatment effect of fewer female ward candidates in cycle 4 to be explained by a shortage of women, we should find that parties in treatment municipalities have greater trouble finding female candidates. We test this hypothesis. A party is defined as unconstrained if the number of female candidates in its party list is greater to the number of women it needs to list as candidates due to quotas.[24] Table XIV shows that parties below the threshold are not more unconstrained in the number of women they can list as candidates, particularly in cycle 4.

Secondly, if the only reason for the treatment effect of rising number of female ward candidates over cycles 5 through 7 was the greater availability of experienced female candidates, then there should be no shift in the female share among rookie candidates with zero councilor experience. This stands contrary to the results in Table X. In addition, the number of female councilors in cycle 4 was still too small to account for all the rise in the female candidates over time.

Lastly, the quota did not improve the relative quality of female candidates faster in treated municipalities, as measured by the gender gap in education. We estimate an individual-level version of equation (2):

$$
\begin{aligned}
Y_{icbt} = \alpha_b + \alpha_t + \sum_{s=4}^{7} \beta_s \cdot Treat_{cb4} + \sum_{s=4}^{7} \pi_s \cdot (Treat_{cb4} \times Female_{icbt}) \\
+ \sum_{s=4}^{7} \kappa_s \cdot Female_{icbt} + f(x_{cb4}) + X'_{cbt}\gamma + X'_{icbt}\delta + \epsilon_{cbt} \quad (7)
\end{aligned}
$$

where $Treat_{cb4}$ is treatment defined based on treatment status at election cycle 4, $X_{cbt}$ denotes municipality-level control variables such as the contemporaneous number of seats, $\alpha_b$ and $\alpha_t$ are bin and cycles fixed effects, while $X_{icpt}$ indicates individual-level controls, including the party of affiliation, whether one's occupation is related to politics, whether one was previously elected, and age and age squared. The outcomes we consider are i) a categorical variable for the level of education,[25] ii) years of schooling, iii) whether the candidate holds a bachelor degree or more, iv) whether the candidate has received tertiary education in the top 20 universities in Korea.[26]

Table XV shows that the treatment does not affect the gender gap in candidates' education. The coefficient of the interaction between the treatment dummy and the female dummy is never statistically significantly different from zero. We do observe an increase in the education levels of both men and women over time, but it does not occur differentially between treatment and control municipalities.

---

[24]For example, the number of women a party needs to include in the party list is 1 if the number of PR seats for the municipality is 1 or 2, and 2 if the number of PR seats is 3 or 4. In election cycle 4, the quota, while legislated, did not have an enforcement system, so a small number of parties listed fewer women than stipulated.

[25]Ranging from 0 to 12 based on final education degree enrolled in and degree completion status.

[26]A university belongs to the top 20 if it was in the top 20 in the University Rankings in any year between 1995 and 2019 according to JoongAng Ilbo, one of the three biggest newspapers in Korea.

***Change in voter preference for women***   The observed gradual increase in the number of fe-
male candidates could be explained by parties responding to a change in voters' attitudes towards
women. Voters might be the ones learning about women's competence, or might simply become
more favorable towards women after experiencing female councilors. Hence, parties could nom-
inate female candidates for reasons not directly related to them learning about their competence,
but to strategically adapt their strategies to follow voters' preferences.

In order to test whether this is the case, we follow Esteve-Volart and Bagues (2012) and we
compare the gender gap in votes received by candidates in treated and control municipalities over
time. We estimate equation (7) with the vote share a candidate obtained in the ward election as the
outcome variable. Columns (1) to (5) of Table XVI display the estimates of the specification as we
gradually add controls. Columns (1) includes only the number of ward seats for the *contempora-
neous* election. Column (2) includes also the position on the ballot list and the party of affiliation,
while in Column (3) we additionally control for individual characteristics. Lastly, Columns (4)
includes a control for the number of ward candidates in the municipality, and Column (5) includes
list fixed effects. A list is defined over a given election cycle, for a given municipality, and for a
given political party.

If the quota induced a change in voters' attitudes towards women, female ward candidates
would get higher vote shares relative to males with comparable characteristics and ballot positions
in treated compared to control municipalities. The only exception is the unlikely event in which
parties are able to preempt voter preferences perfectly and put forth just the right number of females
in the right position every election cycle.[27] Looking at Columns (1) - (5), we can see that this is not
the case. There is no evidence of a treatment effect on the gender gap in vote share or a different
evolution of the gender gap over time in treated compared to control municipalities.

***Change in parties' taste for women***   There is an additional mechanism that would deliver observa-
tionally equivalent results. Parties in treated councils could have initially selected a lower number
of female ward candidates due to taste-based discrimination. The quota could then have then been
successful in forcing parties in treated councils to experience a higher number of women, and as a
consequence grow out of their distaste for them faster than control municipalities.

However, Table XIII shows that there is no counteraction for incumbent ward candidates
(columns 2, 3) in cycle 4, but only for rookie ward candidates (columns 5, 6). This suggests
that general distaste for women is probably not one of the reasons why we see parties initially
putting forth fewer women among ward candidates. Lack of information regarding their compe-
tence seems to be a much more plausible explanation.

Moreover, the results of the heterogeneity analysis based on the education of the first PR elected

---

[27]If this was the case, we would see no gender gap in vote share. However, XVI shows that this is not the case.

women displayed in Table XI also supports this hypothesis. The shift towards female candidates is more apparent and stronger when the first women are more educated. This points again to a story of learning about the competence of females over time, and not one of a growing taste for women that would occur irrelevantly of ability.

# 8   Conclusion

This paper highlights that with time, affirmative action policies can still be effective despite an initial backlash, as long as the policies are not completely undone. Moreover, such is the case even in settings where the target group consists of a very small minority among the incumbents. Through exposure to the minority group, the policies provide incumbents with an opportunity to learn about the competence of the minority group. Once the learning takes off, the policy itself might be unneeded.

Gender quotas are needed not for equal representation's sake, but also from an efficiency standpoint. If informational failure leads to suboptimally low numbers of female politicians, quotas play a role in rectifying this source of inefficiency.

Although gender quotas in parliaments have been adopted broadly worldwide, there are still many countries that have none in place, such as Egypt, India, Liberia, Mauritius, Sao Tome and Principe, Sierra Leone, and Sri Lanka. Unsurprisingly, these countries also suffer from low levels of female representation in national parliaments. The South Korean setting of this paper is unique in that it studies the effect of a gender quota in the legislative body from a starting point of practically zero women. Therefore, this paper is informative about the effect of gender quotas where they are most needed.

What remains to be crystallized is exactly which aspect of women's competence parties are learning about. As of yet, we do not know whether it is their election campaign skills, their loyalty to the party, their keenness as legislators, or their ability to meet the demands of the electorate. Further evidence is needed in this direction.

This paper is a part of a bigger agenda that attempts to study how a gender quota might trigger a gradual process of learning in favor of women. To tackle the precise mechanisms through which the learning takes place, we plan to study in future work the specific interactions among councilors recorded in the transcripts of council meetings.

# References

Angrist, J. and J.-S. Pischke (2009). *Mostly Harmless Econometrics: An Empiricist's Companion* (1 ed.). Princeton University Press.

Bagues, M. and P. Campa (2020). Women and Power: Unpopular, Unwilling, or Held Back? A Comment. *Journal of Political Economy 128*(5), 2010–2016.

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

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

Baltrunaite, A., A. Casarico, P. Profeta, and G. Savio (2019). Let the voters choose women. *Journal of Public Economics 180*, 104085.

Baskaran, T. and Z. Hessami (2018, August). Does the election of a female leader clear the way for more women in politics? *American Economic Journal: Economic Policy 10*(3), 95–121.

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

Besley, T., O. Folke, T. Persson, and J. Rickne (2017). Gender quotas and the crisis of the mediocre man: Theory and evidence from sweden. *American Economic Review 107*(8), 2204–42.

Bhalotra, S., I. Clots-Figueras, and L. Iyer (2018). Pathbreakers? women's electoral success and future political participation. *The Economic Journal 128*(613), 1844–1878.

Braga, M. and F. Scervini (2017). The performance of politicians: The effect of gender quotas. *European Journal of Political Economy 46*, 1 – 14.

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

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

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

Cho, H.-O. and E. H. Kim (2010). A Study of Historical Context and Institutionalization Process of Gender Quota System and Women's Political Empowerment Movement in Korea 2000-2010. *Journal of Korean Social Trend and Perspective* (79), 110–139.

Clayton, A. (2015). Women's political engagement under quota-mandated female representation: Evidence from a randomized policy experiment. *Comparative Political Studies 48*(3), 333–369.

Clots-Figueras, I. (2012). Are female leaders good for education? evidence from india. *American Economic Journal: Applied Economics 4*(1), 212–244.

Dahlerup, D. and L. Freidenvall (2013). Electoral gender quota systems and their implementation in europe. Technical report, The European Parliament, the Committee on Women's Rights and Gender Equality.

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.

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.

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

Global Gender Gap Report (2021). Technical report, World Economic Forum.

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

Holzer, H. and D. Neumark (2000, September). Assessing affirmative action. *Journal of Economic Literature 38*(3), 483–568.

International IDEA (2021). Gender quotas database. Global Database of Gender Quotas in parliaments worldwide by International Institute for Democracy and Electoral Assistance, Inter-Parliamentary Union and Stockholm University.

Jang, G. S. (2008). Busan Provincial Council gets "Poor Report Card". *"No-cut" News*.

Jeon, J. Y. (2013). Political Dynamics of Adopting Gender Quota in Korea. *Journal of Korean Politics 22*(1), 29–52.

Jin, J. W. (2018). Female nomination subsidy monopolized by two major parties. *Women News*.

Kim, S. H. (2005). Local Council Elections - "Reformed" or "Deformed"? *Pressian News*.

Kim, W. H., M. J. Kim, H. C. Lee, and H. K. Kim (2003). *Research on measures to expand female candidate nomination by political parties*. Korea Women's Development Institute.

Lee, D. S. (2008). Randomized experiments from non-random selection in u.s. house elections. *Journal of Econometrics 142*(2), 675 – 697. The regression discontinuity design: Theory and applications.

Lee, J. J. (2019). *The Legislative Impact Analysis of Gender Quota in Local Council District Elections*. Number 44. National Assembly Research Service.

Lee, M. S. (2003). Measures to expand female political representation in preparation for the 17th general election. Panel discussion in the Meeting of Female Members of the National Assembly and Municipal Councils To Increase the Participation of Women in Politics.

Lim, H. K. (2018). Women's political representation and the current state of Gyeonggido: Focused on local elections in 2018. (102). Gyeonggido Family and Women Research Institute [Issue Analysis].

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

Local Decentralization Bureau – Election and Local Council Division (2019). Ministry of the Interior and Safety to Block "Self-Assessment" of Overseas Training for Local Councils and to Redeem Unjust Spending. Ministry of the Interior and Safety.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics 142*(2), 698–714.

Ministry of the Interior and Safety (2018). Municipal councilor's activities guide. Publication Registration Number: 11-1741000-000101-15.

Murray, R. (2010). Second among unequals? a study of whether france's "quota women" are up to the job. *Politics amp; Gender 6*(1), 93–118.

National Election Commission (2018). Election subsidy payments by year (2001-2018). National Election Commission Data Resource Archive.

Rigon, M. and G. M. Tanzi (2012, April). Does gender matter for public spending? Empirical evidence from Italian municipalities. Temi di discussione (Economic working papers) 862, Bank of Italy, Economic Research and International Relations Area.

Weeks, A. C. and L. Baldez (2015). Quotas and qualifications: the impact of gender quota laws on the qualifications of legislators in the italian parliament. *European Political Science Review 7*(1), 119–144.

# 9 Tables

<div align="center">

**Table I:**
**AMENDMENTS TO LEGISLATION ON MUNICIPAL COUNCIL ELECTIONS**

</div>

| First applicable election year | Amendment |
|---|---|
| 2006 | [PR] Proportional representation introduced<br>[W] Single-member plurality vote → Multi-member plurality vote<br>[PR] Odd-number candidates in party lists must be female (not enforced)<br>[W] Subsidies to parties for nominating female candidates |
| 2010 | [PR] Odd-number candidates in party lists must be female (enforced)<br>[W] At least one female candidate per general election district |

*Notes*: Adapted from Lim (2018). [PR] indicates rules relating to proportional representation councilors and [W] to ward councilors.

<div align="center">

**Table II:**
**DESCRIPTIVE STATISTICS ON MUNICIPAL COUNCILS' GENDER COMPOSITION**

</div>

| | Election cycle (year) | | | | | | |
|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| | (1995) | (1998) | (2002) | (2006) | (2010) | (2014) | (2018) |
| **Total number of councilors** | | | | | | | |
| Min. | 7 | 7 | 7 | 7 | 7 | 7 | 7 |
| Mean | 19.9 | 15.0 | 15.0 | 12.6 | 12.6 | 12.8 | 12.8 |
| Max. | 50 | 40 | 41 | 36 | 34 | 43 | 44 |
| **Number of PR councilors** | | | | | | | |
| Min. | - | - | - | 1 | 1 | 1 | 1 |
| Mean | - | - | - | 1.63 | 1.63 | 1.67 | 1.70 |
| Max. | - | - | - | 4 | 4 | 5 | 5 |
| **Gender ratio** | | | | | | | |
| Min. | 0 | 0 | 0 | 0 | 0.06 | 0.08 | 0.10 |
| Mean | 0.01 | 0.01 | 0.02 | 0.15 | 0.21 | 0.25 | 0.29 |
| Max. | 0.43 | 0.22 | 0.28 | 0.46 | 0.57 | 0.86 | 0.64 |
| **Gender ratio among PR councilors** | | | | | | | |
| Min. | - | - | - | 0* | 0* | 0.50 | 0.50 |
| Mean | - | - | - | 0.87 | 0.96 | 0.97 | 0.98 |
| Max. | - | - | - | 1 | 1 | 1 | 1 |
| **Minimum number of women required** | | | | | | | |
| Min. | - | - | - | 0 | 1 | 1 | 1 |
| Mean | - | - | - | 1.12 | 1.12 | 1.13 | 1.13 |
| Max. | - | - | - | 2 | 2 | 3 | 3 |

*Notes*: *Gender quotas were introduced in 2005. However, during the election of 2006, they remained merely a strong recommendation, so it was still legal to place a male in slot 1 of party lists. Most municipalities complied, but 14 of them had no female PR councilors. In election year 2010, the minimum of the gender ratio among PR councilors is 0 because in one council the elected woman was invalidated for being a member of multiple parties.

**Table III:**
**THE EFFECT OF AN ADDITIONAL PR SEAT AT CYCLE 4 ON THE NUMBER OF FEMALE PR COUNCILORS**

| | Number of female PR councilors | | |
| --- | --- | --- | --- |
| | Bin 1 | Bin 2 | Bin 3 |
| | (1) | (2) | (3) |
| Treat at cycle 4× Cycle 4 | 0.84*** | 0.52*** | -0.23 |
| | (9.35) | (2.82) | (-0.42) |
| Treat at cycle 4 × Cycle 5 | 0.84*** | 0.32* | 0.44 |
| | (12.28) | (1.71) | (1.21) |
| Treat at cycle 4 × Cycle 6 | 0.77*** | 0.58*** | 0.04 |
| | (10.18) | (3.33) | (0.10) |
| Treat at cycle 4 × Cycle 7 | 0.77*** | 0.57*** | 0.16 |
| | (9.61) | (3.68) | (0.40) |
| Running variable form | council | council | council |
| $N$ | 670 | 198 | 33 |

*Notes*: This table reports the results of regressing (2), separately for each bin, with the number of female PR councilors as the outcome variable. t statistics from standard errors clustered by municipality in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

## Table IV:
## BALANCE TESTS ON PRE-DETERMINED CHARACTERISTICS

| | Panel A: Population characteristics | | | | | | |
|---|---|---|---|---|---|---|---|
| | Total population | | Voting age population | | | Households | |
| | Total (1) | Foreign (2) | Total (3) | Male (4) | Female (5) | Total (6) | Foreign (7) |
| Treat at cycle 4 | -23.97 (-0.78) | 0.01 (0.78) | -17.22 (-0.76) | -7.84 (-0.70) | -9.38 (-0.82) | -5.59 (-0.50) | 0.01 (0.83) |
| Running variable form | council | council | council | council | council | council | council |
| N | 219 | 219 | 219 | 219 | 219 | 219 | 219 |

| | Panel B: Time use and labor force participation | | | | | |
|---|---|---|---|---|---|---|
| | House work (hours) | | Employed | | Unemployed | |
| | Male (1) | Female (2) | Male (3) | Female (4) | Male (5) | Female (6) |
| Treat at cycle 4 | 1.30 (0.93) | -0.31 (-0.09) | 86940.60 (0.36) | 42556.87 (0.30) | 18867.18 (0.15) | 54934.05 (0.22) |
| Running variable form | council | council | council | council | council | council |
| N | 219 | 219 | 219 | 219 | 219 | 219 |

| | Panel C: Education | | | | | |
|---|---|---|---|---|---|---|
| | Elementary School or less | | Middle School | | High School | |
| | Male (1) | Female (2) | Male (3) | Female (4) | Male (5) | Female (6) |
| Treat at cycle 4 | 8835.58 (0.84) | 4324.18 (0.28) | 5113.39 (0.34) | 2762.45 (0.23) | 35035.82 (0.38) | 24315.10 (0.41) |
| Running variable form | council | council | council | council | council | council |
| N | 219 | 219 | 219 | 219 | 219 | 219 |

| | Tech. University | | University | | Graduate Studies | |
|---|---|---|---|---|---|---|
| | Male (1) | Female (2) | Male (3) | Female (4) | Male (5) | Female (6) |
| Treat at cycle 4 | 10063.70 (0.33) | 3714.14 (0.17) | 21383.83 (0.26) | 6028.32 (0.15) | 6508.27 (0.32) | 1412.68 (0.18) |
| Running variable form | council | council | council | council | council | council |
| N | 219 | 219 | 219 | 219 | 219 | 219 |

| | Panel D: Political leaning, economic, and ward division characteristics | | | | | |
|---|---|---|---|---|---|---|
| | Past vote share by party | | Budget | | Ward characteristics | |
| | Conservative (1) | Progressive (2) | Total (3) | Council expenses (4) | Num of wards (5) | Seats per ward (6) |
| Treat at cycle 4 | -0.02 (-0.26) | -0.00 (-0.12) | 54.16 (0.63) | 0.02 (0.36) | -0.23 (-1.30) | 0.17 (1.48) |
| Running variable form | council | council | council | council | ward | ward |
| N | 219 | 219 | 219 | 219 | 219 | 219 |

*Notes*: The regression specification follows equation (2), and the sample consists of bins 1 and 2 at election cycle 4. Panel A: number of residents, residents of voting age, and households by gender or citizenship status (source: National Election Commission). Panel B: province-level information on the number of hours spent doing unpaid domestic or care-giving services for household and family members by gender (2004 Statistics of Korea Time Use Survey), the number of employed and unemployed individuals (2005 South Korean Census). Panel C: province-level information on the number of individuals by education and gender (2005 South Korean Census). Panel D: vote share of each main party in the previous election's PR arm, and number and size of wards in each municipality (National Election Commission), municipal government expenditures (Local Finance Disclosure System of the Ministry of the Interior and Safety). t statistics from standard errors clustered by municipality in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

**Table V:**
**THE EFFECT OF BEING PAST THE THRESHOLD ON THE NUMBER OF COUNCILORS**

| | All political parties | | | | | | | | |
| | Ward | | | PR | | | All | | |
| | Male (1) | Female (2) | Female (3) | Male (4) | Female (5) | Female (6) | Male (7) | Female (8) | Female (9) |
|---|---|---|---|---|---|---|---|---|---|
| Treat at cycle 4 x Cycle 4 | 0.45* | -0.34 | -0.36* | 0.09 | 0.76*** | 0.76*** | -0.29 | 0.29 | 0.29 |
| | (0.26) | (0.22) | (0.22) | (0.08) | (0.09) | (0.08) | (0.33) | (0.28) | (0.28) |
| Treat at cycle 4 x Cycle 5 | -0.13 | 0.31 | 0.28 | 0.10 | 0.71*** | 0.71*** | -0.87** | 0.88*** | 0.88*** |
| | (0.29) | (0.23) | (0.23) | (0.06) | (0.08) | (0.08) | (0.35) | (0.29) | (0.28) |
| Treat at cycle 4 x Cycle 6 | -0.22 | 0.52* | 0.47* | 0.08 | 0.71*** | 0.70*** | -0.97** | 1.09*** | 1.06*** |
| | (0.33) | (0.27) | (0.26) | (0.06) | (0.08) | (0.08) | (0.38) | (0.32) | (0.31) |
| Treat at cycle 4 x Cycle 7 | -0.28 | 0.82*** | 0.73** | 0.04 | 0.72*** | 0.69*** | -1.08*** | 1.41*** | 1.33*** |
| | (0.36) | (0.29) | (0.29) | (0.06) | (0.08) | (0.08) | (0.41) | (0.35) | (0.33) |
| N. of ward councilors | | | 0.17* | | | | | | |
| | | | (0.09) | | | | | | |
| N. of councilors | | | | | | 0.09*** | | | 0.24*** |
| | | | | | | (0.02) | | | (0.09) |
| Running variable form | ward | ward | ward | council | council | council | council | council | council |
| $N$ | 868 | 868 | 868 | 868 | 868 | 868 | 868 | 868 | 868 |

*Notes*: The regression specification is given by equation (2). The sample includes only bins 1 and 2. t statistics from standard errors clustered by municipality in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

**Table VI:**
**THE EFFECT OF BEING PAST THE THRESHOLD ON THE NUMBER OF CANDIDATES**

| | All political parties | | | | | |
| | Ward | | PR | | All | |
| | Male (1) | Female (2) | Male (3) | Female (4) | Male (5) | Female (6) |
|---|---|---|---|---|---|---|
| Treat at cycle 4 x Cycle 4 | 3.70*** | -0.24 | 0.94*** | 0.96*** | 1.89*** | 0.51 |
| | (1.16) | (0.35) | (0.22) | (0.22) | (0.28) | (0.49) |
| Treat at cycle 4 x Cycle 5 | 0.56 | 0.49 | 0.63*** | 1.20*** | 1.83*** | 1.51*** |
| | (0.91) | (0.36) | (0.17) | (0.22) | (0.28) | (0.49) |
| Treat at cycle 4 x Cycle 6 | -1.39* | 0.91** | 0.25* | 1.02*** | 1.28*** | 1.76*** |
| | (0.84) | (0.42) | (0.15) | (0.22) | (0.25) | (0.54) |
| Treat at cycle 4 x Cycle 7 | -2.23** | 1.10** | 0.21 | 1.27*** | 1.49*** | 2.22*** |
| | (1.00) | (0.44) | (0.16) | (0.22) | (0.24) | (0.58) |
| Running variable form | ward | ward | council | council | council | council |
| $N$ | 868 | 868 | 868 | 868 | 868 | 868 |

*Notes*: The regression specification is given by equation (2). The sample includes only bins 1 and 2. t statistics from standard errors clustered by municipality in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

**THE EFFECT OF BEING PAST THE THRESHOLD ON THE NUMBER OF CANDIDATES LIKELY TO BE ELECTED**

| | Main political parties | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| | All ward candidates | | | Useful positions | | |
| | Male (1) | Female (2) | Female (3) | Male (4) | Female (5) | Female (6) |
| Treat at cycle 4 x Cycle 4 | 1.43* | -0.28 | -0.43* | 0.80* | -0.39** | -0.47*** |
| | (0.76) | (0.23) | (0.22) | (0.47) | (0.18) | (0.17) |
| Treat at cycle 4 x Cycle 5 | 0.39 | 0.48* | 0.36* | -0.25 | 0.42** | 0.38** |
| | (0.65) | (0.25) | (0.22) | (0.48) | (0.18) | (0.17) |
| Treat at cycle 4 x Cycle 6 | 0.57 | 0.76** | 0.59** | 0.07 | 0.45* | 0.35* |
| | (0.64) | (0.29) | (0.27) | (0.45) | (0.23) | (0.20) |
| Treat at cycle 4 x Cycle 7 | -0.54 | 1.17*** | 1.09*** | -0.13 | 0.68*** | 0.58** |
| | (0.65) | (0.32) | (0.31) | (0.46) | (0.25) | (0.23) |
| N. of ward candidates | | | 0.13*** | | | |
| | | | (0.02) | | | |
| N. of ward candidates in useful positions | | | | | | 0.19*** |
| | | | | | | (0.02) |
| Running variable form | ward | ward | ward | ward | ward | ward |
| $N$ | 867 | 867 | 867 | 867 | 867 | 867 |

*Notes*: The regression specification is given by equation (2). The sample includes only bins 1 and 2 and is restricted to the two main parties. The number of observations is 867 instead of 868 since in one municipality main parties only have proportional candidates. Useful positions refer to candidates in the high-up positions on the ballot for the party in a ward (position 1 if the ward elects 1-2 councilors, and positions 1 and 2 if the ward elects 3-4 councilors). t statistics from standard errors clustered by municipality in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

## Table VIII:
## THE EFFECT OF BEING PAST THE THRESHOLD AT ELECTION CYCLE 4
## ON THE NUMBER OF WARD CANDIDATES IN SAFE AND UNSAFE WARDS

| | Main political parties, (ward×party)-level regressions | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | All ward positions | | | | | | Useful ward positions | |
| | Male | Female | All | Female | Female Safe | Female Unsafe | Female Safe | Female Unsafe |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Treat at cycle 4 × Cycle 4 | 0.06 | 0.00 | 0.06 | -0.00 | -0.01 | 0.00 | -0.05** | -0.01 |
| | (0.68) | (0.11) | (0.74) | (-0.13) | (-0.40) | (0.09) | (-2.02) | (-0.21) |
| Treat at cycle 4 × Cycle 5 | -0.05 | 0.07** | 0.01 | 0.06** | 0.04 | 0.10*** | 0.01 | 0.10*** |
| | (-0.62) | (2.34) | (0.18) | (2.26) | (1.12) | (2.60) | (0.33) | (3.15) |
| Treat at cycle 4 × Cycle 6 | -0.03 | 0.06* | 0.03 | 0.06* | -0.04 | 0.10*** | -0.01 | 0.06** |
| | (-0.33) | (1.95) | (0.39) | (1.76) | (-0.74) | (2.64) | (-0.18) | (2.05) |
| Treat at cycle 4 × Cycle 7 | -0.08 | 0.07** | -0.01 | 0.07** | 0.04 | 0.09** | -0.01 | 0.05 |
| | (-0.96) | (2.18) | (-0.15) | (2.16) | (0.70) | (2.55) | (-0.21) | (1.55) |
| N. ward candidates | | | | 0.11*** | 0.10*** | 0.13*** | 0.09*** | 0.10*** |
| | | | | (10.17) | (7.05) | (9.10) | (4.55) | (4.22) |
| Running variable form | ward | ward | ward | ward | ward | ward | ward | ward |
| $N$ | 6035 | 6035 | 6035 | 6035 | 2337 | 3698 | 2337 | 3698 |

*Notes*: The regression specification is given by equation (2). The sample includes only bins 1 and 2 and is restricted to the two main parties. The level of observation is party ward. A ward is considered as safe for a party if the party won the greatest vote share in the PR arm in the ward in the previous election cycle, and it got over 10 percentage points more vote share than the next popular party. Unsafe wards are all the others. t statistics from standard errors clustered by municipality in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

**Table IX:**
**PLACEBO TEST - THE EFFECT OF BEING PAST THE THRESHOLD BEFORE THE REFORM**

|  | Number of ward candidates | | | |
|---|---|---|---|---|
|  | Male | Female | Male | Female |
|  | (1) | (2) | (3) | (4) |
| Treat × Cycle 1 | 0.92 | 0.01 | 0.03 | 0.04 |
|  | (1.32) | (0.26) | (1.33) | (0.31) |
| Treat × Cycle 2 | 0.40 | -0.05 | -0.61 | 0.03 |
|  | (1.12) | (0.24) | (1.01) | (0.30) |
| Treat × Cycle 3 | 1.06 | 0.06 | 0.01 | 0.15 |
|  | (1.02) | (0.27) | (0.90) | (0.32) |
| Treat × Cycle 4 | 3.22*** | 0.72** | 2.93** | 0.59** |
|  | (1.21) | (0.28) | (1.15) | (0.27) |
| Treat × Cycle 5 |  |  | -0.71 | 1.29*** |
|  |  |  | (0.83) | (0.31) |
| Treat × Cycle 6 |  |  | -2.29*** | 1.58*** |
|  |  |  | (0.79) | (0.42) |
| Treat × Cycle 7 |  |  | -2.79*** | 1.58*** |
|  |  |  | (0.98) | (0.44) |
| Running variable form | ward | ward | ward | ward |
| $N$ | 899 | 899 | 1577 | 1577 |

*Notes*: The running variable and treatment status are defined contemporaneously. While municipality size and divisions remained almost unchanged from cycle 4 onward, they changed dramatically during the first three election cycles. Thus, it would be inaccurate to define treatment for the first three cycles using cycle 4 municipality characteristics. The regression specification is given by equation (1). t-statistics from standard errors clustered by municipality in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

**Table X:**

**THE EFFECT OF BEING PAST THE THRESHOLD ON ROOKIE AND INCUMBENT FEMALE COUNCILORS**

| | Main political parties | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| | Incumbent Councilors | | | Rookie Councilors | | |
| | Ward Female | PR Female | All Female | Ward Female | PR Female | All Female |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Treat at cycle 4 x Cycle 4 | -0.10 | 0.00 | -0.16 | -0.24* | 0.44*** | 0.19 |
| | (-0.89) | (0.07) | (-1.20) | (-1.91) | (4.00) | (0.95) |
| Treat at cycle 4 x Cycle 5 | 0.14 | 0.02 | 0.09 | 0.11 | 0.41*** | 0.50** |
| | (1.01) | (0.75) | (0.58) | (0.84) | (4.06) | (2.52) |
| Treat at cycle 4 x Cycle 6 | 0.41** | -0.00 | 0.34* | 0.07 | 0.68*** | 0.74*** |
| | (2.33) | (-0.25) | (1.79) | (0.49) | (6.83) | (3.62) |
| Treat at cycle 4 x Cycle 7 | 0.39** | -0.00 | 0.32* | 0.38** | 0.63*** | 1.00*** |
| | (2.28) | (-0.19) | (1.76) | (2.15) | (6.39) | (4.45) |
| Running variable form | ward | council | council | ward | council | council |
| $N$ | 867 | 865 | 868 | 867 | 865 | 868 |

*Notes*: The regression specification is given by equation (2). The sample includes only bins 1 and 2 and is restricted to the two main parties. The number of observations is 867 instead of 868 in columns 1 and 4 since in one municipality main parties only have proportional candidates, and 865 in columns 2 and 5 as in 3 municipality main parties only have ward candidates. A candidate is defined as incumbent if they have been elected in at least one of the previous election cycles. Rookies are the remaining candidates. t statistics from standard errors clustered by municipality in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

**Table XI:**
**THE EFFECT OF BEING PAST THE THRESHOLD BY YEARS OF SCHOOLING OF**
**THE FIRST FEMALE PR COUNCILORS**

| | **Main political parties** | | | | | | | |
| | Ward candidates | | | | | | | |
| | Below Median Education | | | | Above Median Education | | | |
| | Male | Male | Female | Female | Male | Male | Female | Female |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Treat at cycle 4 x Cycle 5 | 0.30 | 0.19 | -0.18 | -0.01 | -0.31 | -0.03 | 0.87** | 0.92*** |
| | (0.29) | (0.19) | (-0.37) | (-0.03) | (-0.32) | (-0.03) | (2.27) | (2.64) |
| Treat at cycle 4 x Cycle 6 | 0.33 | 0.22 | 0.19 | 0.35 | -0.06 | 0.22 | 1.06** | 1.11*** |
| | (0.36) | (0.24) | (0.38) | (0.82) | (-0.06) | (0.22) | (2.37) | (2.69) |
| Treat at cycle 4 x Cycle 7 | -0.80 | -0.91 | 0.70 | 0.86* | -0.72 | -0.43 | 1.13** | 1.18*** |
| | (-0.79) | (-0.91) | (1.28) | (1.77) | (-0.77) | (-0.45) | (2.53) | (3.03) |
| Running variable form | ward | ward | ward | ward | ward | ward | ward | ward |
| $N$ | 275 | 275 | 275 | 275 | 266 | 266 | 266 | 266 |
| Cycle 4 female candidates avg. education | No | Yes | No | Yes | No | Yes | No | Yes |
| N. ward cycle 4 female candidates | No | Yes | No | Yes | No | Yes | No | Yes |

*Notes*: Columns 1, 3, 5, 7 display the results of equation (2) for the sample of councils in which the female PR councilors elected in cycle 4 have below-median and above-median years of schooling, respectively. Columns 2, 4, 6, and 8, include a control for the average education level of all female candidates affiliated with the two main parties in cycle 4 and for the total number of ward female candidates affiliated to two main parties in the municipality in cycle 4. The sample includes only bins 1 and 2. We divide municipalities into two groups by whether the average years of schooling of the PR women elected in the municipality in cycle 4 are above or below the median. The sample of analysis includes only cycles 5, 6, and 7 and municipalities with at least one PR elected woman among main party candidates in cycle 4 for which we have information on education. t statistics from standard errors clustered by municipality in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

## Table XII:
## THE EFFECT OF MARGINALLY WINNING A PR WOMAN
## IN THE PREVIOUS ELECTION

| | **Main political parties** | | | |
|---|---|---|---|---|
| | $\mathbb{1}$(Number-1 PR candidate in $t-1$ is a ward candidate in $t$) | | | |
| Bandwidth ($|v_{cpt}|$) | 0.20 | 0.15 | 0.10 | 0.05 |
| | (1) | (2) | (3) | (4) |
| **Panel A: All parties** | | | | |
| $Winner_{t-1}$ | 0.41*** | 0.42*** | 0.35*** | 0.42*** |
| | (5.51) | (5.36) | (3.76) | (3.37) |
| $N$ | 414 | 313 | 216 | 114 |
| **Panel B: 2nd PR candidate = Man** | | | | |
| $Winner_{t-1}$ | 0.44*** | 0.49*** | 0.40*** | 0.45*** |
| | (5.60) | (5.73) | (3.96) | (3.24) |
| $N$ | 308 | 241 | 165 | 90 |
| **Panel C: 2nd PR candidate = Woman** | | | | |
| $Winner_{t-1}$ | 0.35* | 0.21 | 0.07 | -0.21 |
| | (1.87) | (1.05) | (0.32) | (-0.57) |
| $N$ | 106 | 72 | 51 | 24 |

*Notes*: The table reports the result of specification (6) on the probability that the number-1 PR candidate in cycle $t-1$ is one of the ward candidates in cycle $t$. The results are provided for all parties (Panel A), parties for which the second PR candidate in cycle $t-1$ was a man (Panel B), and parties where the second PR candidate was a woman (Panel C). Each column shows the coefficient of the $Winner_{t-1}$ dummy considering different bandwidths. The analysis is restricted to the two main parties. The standard errors are clustered at the municipality $\times$ party level. t statistics from standard errors clustered by municipality$\times$party in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

**Table XIII:**
**EFFECT OF BEING ABOVE THE THRESHOLD FOR INCUMBENT AND ROOKIE WARD CANDIDATES**

| | **Main political parties** | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| | Incumbent Ward Candidates | | | Rookie Ward Candidates | | |
| | Male | Female | Female | Male | Female | Female |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Treat at cycle 4 x Cycle 4 | 0.54 | -0.00 | -0.07 | 0.89 | -0.28* | -0.35** |
| | (1.26) | (-0.00) | (-0.60) | (1.43) | (-1.67) | (-2.30) |
| Treat at cycle 4 x Cycle 5 | 0.94** | 0.47*** | 0.27* | -0.55 | 0.01 | 0.07 |
| | (2.47) | (2.83) | (1.69) | (-0.91) | (0.05) | (0.50) |
| Treat at cycle 4 x Cycle 6 | 0.78** | 0.61*** | 0.42** | -0.21 | 0.15 | 0.15 |
| | (2.06) | (2.88) | (2.20) | (-0.43) | (0.83) | (0.95) |
| Treat at cycle 4 x Cycle 7 | -0.13 | 0.79*** | 0.70*** | -0.41 | 0.38* | 0.38* |
| | (-0.32) | (4.02) | (3.68) | (-0.78) | (1.82) | (1.87) |
| N. ward incumbent candidates | | | 0.14*** | | | |
| | | | (8.15) | | | |
| N. ward rookie candidates | | | | | | 0.12*** |
| | | | | | | (8.75) |
| Running variable form | ward | ward | ward | ward | ward | ward |
| $N$ | 867 | 867 | 867 | 867 | 867 | 867 |

*Notes*: The regression specification is given by equation (2). The sample includes only bins 1 and 2 and is restricted to the two main parties. The number of observations is 867 instead of 868 since in one municipality main parties only have proportional candidates. A candidate is defined as incumbent if they have been elected in at least one of the previous election cycles. Rookies are the remaining candidates. t statistics from standard errors clustered by municipality in parentheses *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$

**Table XIV:**
**PROBABILITY OF BEING CONSTRAINED IN**
**THE NUMBER OF FEMALE CANDIDATES**

|  | Main political parties (party x municipality) | |
| --- | --- | --- |
|  | Pr(constrained) | |
|  | All parties | Ward candidates>0 |
|  | (1) | (2) |
| Treat at cycle 4 x Cycle 4 | 0.088 | 0.093 |
|  | (0.075) | (0.074) |
| Treat at cycle 4 x Cycle 5 | -0.089 | -0.119* |
|  | (0.073) | (0.071) |
| Treat at cycle 4 x Cycle 6 | -0.020 | -0.041 |
|  | (0.065) | (0.064) |
| Treat at cycle 4 x Cycle 7 | -0.072 | -0.086 |
|  | (0.068) | (0.069) |
| Running variable from | council | council |
| $N$ | 1557 | 1520 |

*Notes*: The regression specification is given by equation (2). The unit of analysis is party by municipality. The sample includes only bins 1 and 2, and only the two main parties. A party is defined as constrained if the number of female candidates in the party's list is smaller or equal to the number of women the party is obliged to place in its list due to quotas. In column (2) the sample is restricted to only the main parties in municipalities where the parties have at least one ward candidate. t statistics from standard errors clustered by municipality in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

## Table XV:
## THE EFFECT OF BEING ABOVE THE THRESHOLD ON THE GENDER GAP IN EDUCATION

|  | Main political parties | | | |
|---|---|---|---|---|
|  | Education level (1) | Years of schooling (2) | Bachelor or more (3) | Attended top 20 uni (4) |
| Female | -0.119 | -0.224 | 0.002 | -0.006 |
|  | (0.110) | (0.183) | (0.027) | (0.015) |
| Female × Cycle 5 | 0.089 | 0.141 | -0.020 | 0.021 |
|  | (0.119) | (0.194) | (0.032) | (0.017) |
| Female × Cycle 6 | -0.015 | 0.010 | -0.003 | -0.016 |
|  | (0.141) | (0.230) | (0.038) | (0.019) |
| Female × Cycle 7 | 0.372*** | 0.566*** | 0.002 | -0.006 |
|  | (0.127) | (0.209) | (0.034) | (0.020) |
|  |  |  |  |  |
| Treat at cycle 4 | -0.150 | -0.262 | -0.017 | 0.059** |
|  | (0.152) | (0.247) | (0.034) | (0.026) |
| Treat at cycle 4 × Cycle 5 | -0.054 | -0.085 | -0.036 | -0.006 |
|  | (0.086) | (0.140) | (0.022) | (0.012) |
| Treat at cycle 4 × Cycle 6 | -0.108 | -0.165 | -0.018 | -0.029** |
|  | (0.094) | (0.154) | (0.025) | (0.014) |
| Treat at cycle 4 × Cycle 7 | -0.063 | -0.117 | -0.028 | -0.022 |
|  | (0.099) | (0.160) | (0.027) | (0.017) |
|  |  |  |  |  |
| Treat at cycle 4 × Female | 0.127 | 0.235 | -0.020 | -0.033 |
|  | (0.171) | (0.282) | (0.043) | (0.025) |
| Treat at cycle 4 × Female × Cycle 5 | 0.073 | 0.064 | 0.065 | -0.019 |
|  | (0.171) | (0.283) | (0.052) | (0.026) |
| Treat at cycle 4 × Female × Cycle 6 | 0.102 | 0.099 | 0.017 | 0.032 |
|  | (0.202) | (0.324) | (0.055) | (0.029) |
| Treat at cycle 4 × Female × Cycle 7 | -0.150 | -0.283 | 0.011 | 0.035 |
|  | (0.201) | (0.327) | (0.055) | (0.031) |
|  |  |  |  |  |
| Cycle 5 | 0.595*** | 0.992*** | 0.167*** | -0.005 |
|  | (0.052) | (0.083) | (0.015) | (0.008) |
| Cycle 6 | 1.080*** | 1.763*** | 0.272*** | 0.026*** |
|  | (0.058) | (0.093) | (0.016) | (0.010) |
| Cycle 7 | 1.244*** | 2.075*** | 0.348*** | 0.019* |
|  | (0.058) | (0.093) | (0.016) | (0.011) |
| $N$ | 13235 | 13235 | 13235 | 13235 |
| Running variable form | council | council | council | council |
| Age polynomials | Yes | Yes | Yes | Yes |
| Party affiliation | Yes | Yes | Yes | Yes |
| Rookie vs Incumbent | Yes | Yes | Yes | Yes |
| Political Experience | Yes | Yes | Yes | Yes |

*Notes*: The regression specification is given by equation (7). The analysis is performed at the individual level. The sample includes only bins 1 and 2, and only candidates from the two main parties for which we have education information. The outcome variables are the candidates' level of education (0-12), years of schooling (0-22), a dummy equal to one if the candidate has a bachelor degree or more, and a dummy equal to one if the candidate has received tertiary education from a top 20 university in Korea. t statistics from standard errors clustered by municipality in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

## Table XVI:
## EVOLUTION OF VOTER PREFERENCES

| | Main political parties | | | | |
| | Candidate's vote share in the ward election arm (%) | | | | |
| | No controls (1) | Ballot position (2) | Individual characteristics (3) | N. ward candidates (4) | List fixed effects (5) |
| --- | --- | --- | --- | --- | --- |
| Female | 0.478 | -0.708 | -0.040 | 0.457 | -2.000** |
| | (1.061) | (0.972) | (1.016) | (1.020) | (0.867) |
| Cycle 5 x Female | -0.677 | -1.593 | -2.657** | -3.093*** | -0.996 |
| | (1.317) | (1.130) | (1.194) | (1.171) | (0.993) |
| Cycle 6 x Female | -1.457 | -1.743* | -2.941*** | -3.473*** | -1.452* |
| | (1.169) | (0.996) | (1.031) | (1.036) | (0.873) |
| Cycle 7 x Female | 1.688 | 1.612 | 0.136 | -0.680 | -0.101 |
| | (1.222) | (1.154) | (1.191) | (1.164) | (0.972) |
| | | | | | |
| Treat at cycle 4 | -2.124** | -1.982** | -2.166*** | -0.824 | |
| | (1.014) | (0.796) | (0.801) | (0.671) | |
| Treat at cycle 4 x Cycle 5 | 0.063 | 0.187 | 0.225 | -0.472 | |
| | (0.610) | (0.545) | (0.542) | (0.555) | |
| Treat at cycle 4 x Cycle 6 | 0.828 | 0.883 | 0.896 | -0.456 | |
| | (0.722) | (0.623) | (0.627) | (0.667) | |
| Treat at cycle 4 x Cycle 7 | 1.256 | 1.147* | 1.177* | -0.360 | |
| | (0.782) | (0.632) | (0.630) | (0.624) | |
| | | | | | |
| Treat at cycle 4 x Female | -1.213 | 0.428 | -0.187 | -0.465 | 0.408 |
| | (1.867) | (1.540) | (1.616) | (1.525) | (1.147) |
| Treat at cycle 4 x Cycle 5 x Female | 2.100 | -0.595 | -0.324 | -0.207 | -1.139 |
| | (2.086) | (1.649) | (1.736) | (1.663) | (1.368) |
| Treat at cycle 4 x Cycle 6 x Female | 1.579 | -0.435 | 0.204 | 0.571 | -0.091 |
| | (1.880) | (1.527) | (1.598) | (1.531) | (1.294) |
| Treat at cycle 4 x Cycle 7 x Female | 0.599 | -0.941 | -0.147 | 0.115 | 0.187 |
| | (2.088) | (1.722) | (1.750) | (1.684) | (1.417) |
| | | | | | |
| N | 11246 | 11246 | 10791 | 10791 | 10719 |
| Running variable form | ward | ward | ward | ward | ward |
| Ballot Position | No | Yes | Yes | Yes | Yes |
| Schooling (0-22) | No | No | Yes | Yes | Yes |
| Rookie vs Incumbent | No | No | Yes | Yes | Yes |
| Political Experience | No | No | Yes | Yes | Yes |
| Age polynomials | No | No | Yes | Yes | Yes |
| N. ward candidates | No | No | No | Yes | Yes |
| List Fixed Effects | No | No | No | No | Yes |

* p<.10, ** p<.05, *** p<.01

*Notes*: The regression specification is given by equation (7). The analysis is performed at the individual level. The sample includes only bins 1 and 2, and only candidates from the two main parties. t statistics from standard errors clustered by municipality×election cycle in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

# 10 Figures

**Figure I: Ballot papers in municipal council elections**



*Notes*: This figure illustrates the ballot papers that a voter residing in ward X of municipality A receives for the municipal council elections. The one on the left is used to vote for ward councilors and the one on the right is used to vote for PR councilors. The red ticks indicate how the voter might vote.

**Figure II: Proportion of females in municipal councils, nationwide average**
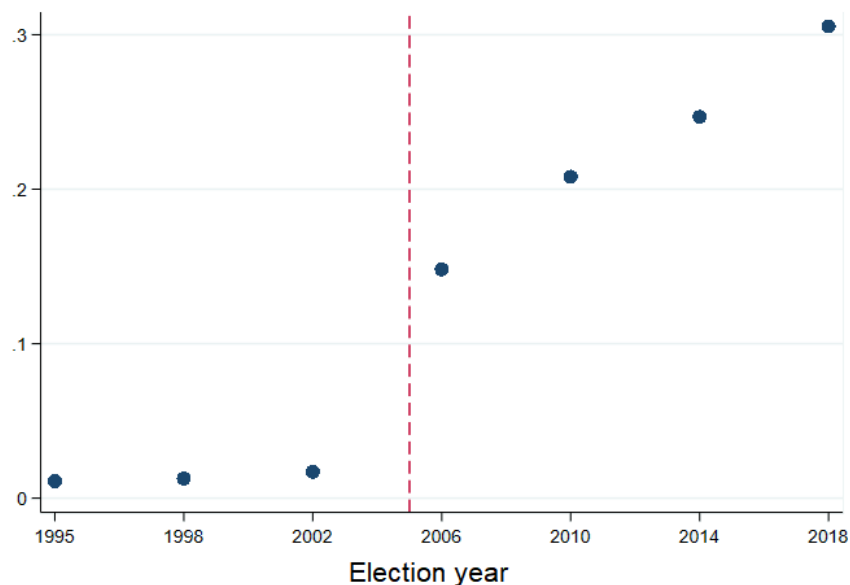


*Notes*: This figure illustrates the nationwide average of the gender ratio in municipal councils, for every election cycle since their emergence. The red dotted line indicates the year of the major reform that instituted the gender quota.

48

**Figure III: The share of females among non-quota candidates**
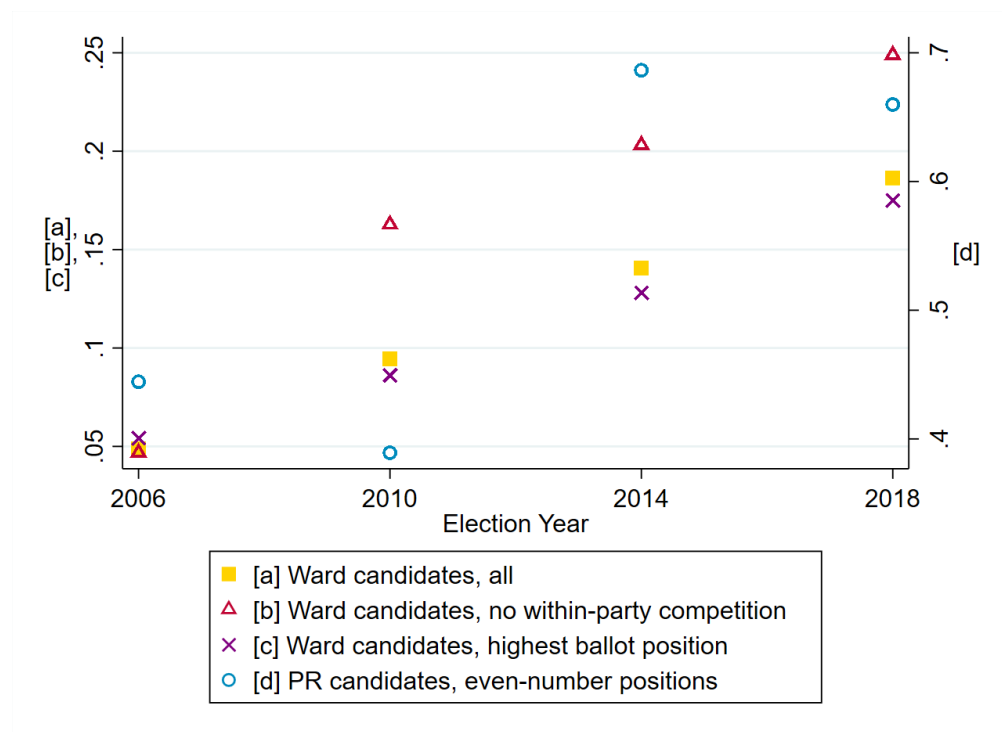


*Notes*: This figure plots the share of females among [a] all ward candidates, [b] ward candidates with no within-party competition, [c] ward candidates that have within-party competition but is ranked the highest, and [d] PR candidates in even-number party list positions. The left-hand vertical axis corresponds to [a], [b], and [c], whereas the right-hand one corresponds to [d].

**Figure IV: Councils by bins around each threshold**
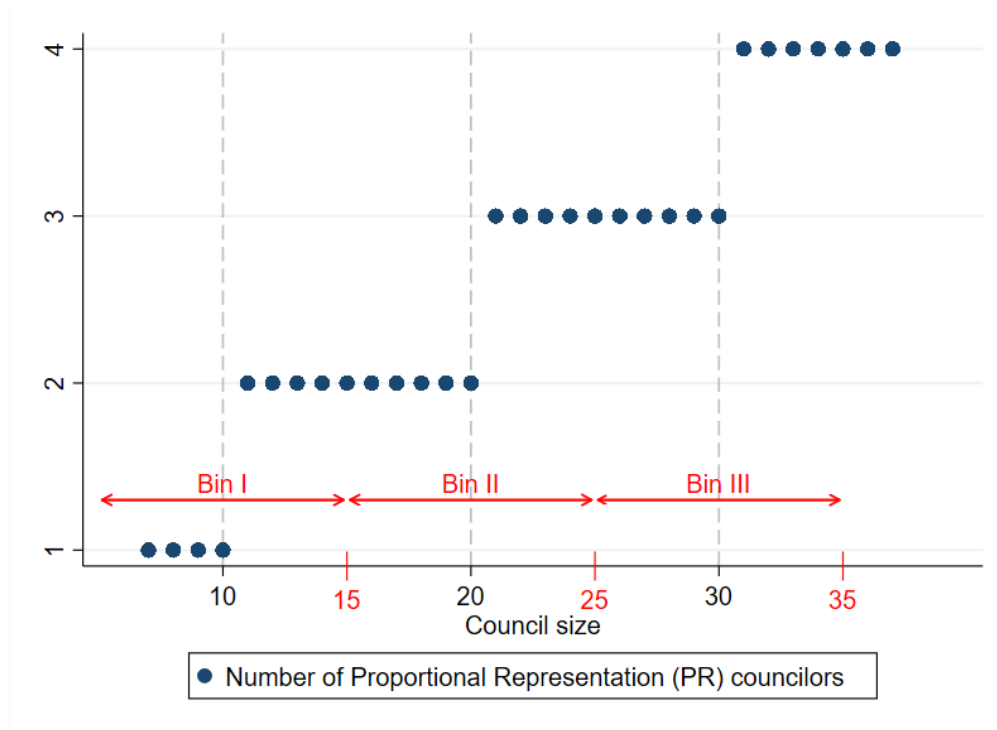


*Notes*: This figure depicts how the number of seats reserved for the proportional representation arm increases as a step function of the total number of councilors in a municipality. There are municipalities that do not correspond to the step function, because they are formed by the union of multiple municipalities *after* the election took place. The municipalities pre- and post-union are all excluded from the sample as outliers and are not shown in this figure. Moreover, the figure depicts how a council is categorized into a bin based on its most proximate threshold.

**Figure V: The average number of female PR councilors by council size**
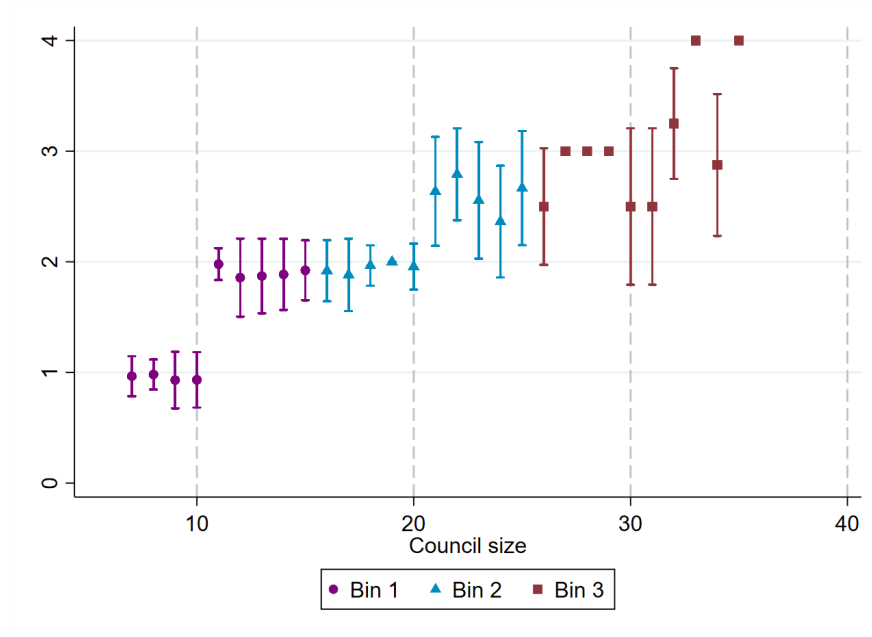


*Notes*: The error bars indicate standard deviation of the number of female PR councilors by council size. Where the erroe bars are missing, there is only one municipality for that council size. Therefore, we can tell that there are only a small number of municipal councils belonging to bin 3.

**Figure VI: Histogram of council size**
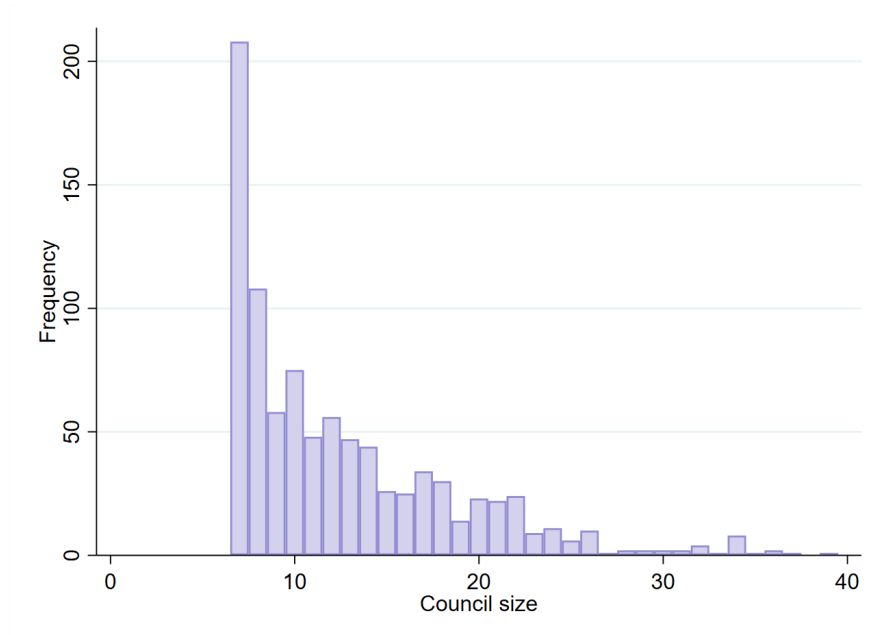


*Notes*: The sample includes all municipal councils of election cycles 4, 5, 6, and 7.

51

**Figure VII: Marginal winners and losers of the last PR seat**
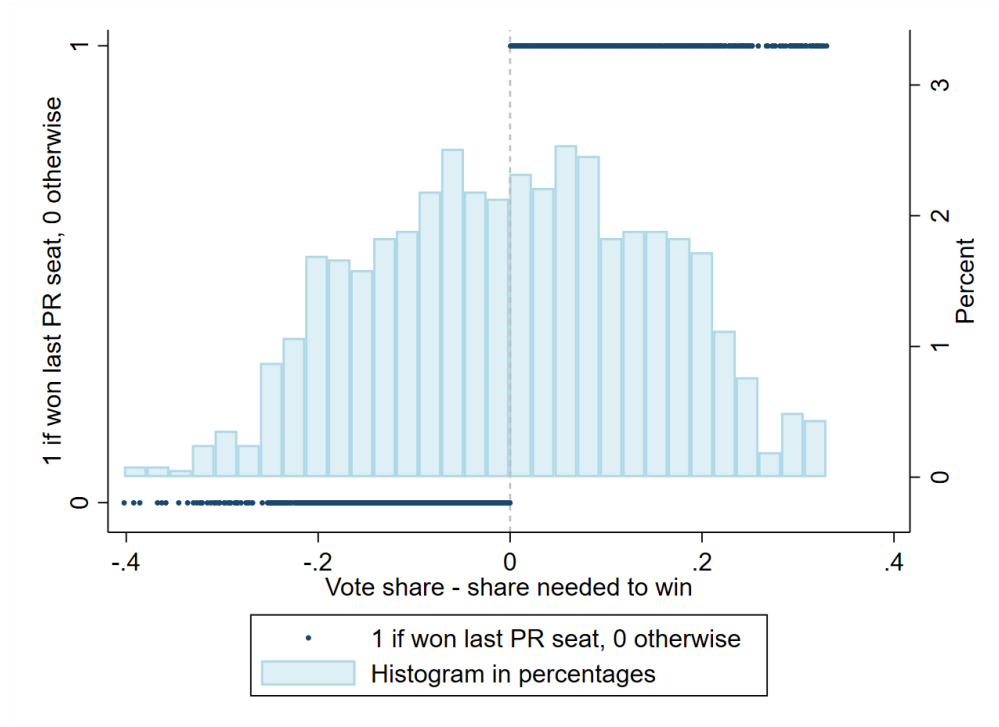


*Notes*: This figure shows the distribution of the vote shares received by the two marginal parties competing for the last PR seat in a municipality. The vote share is computed to be the share of votes received among qualifying parties, i.e. parties that received more than 5% of the raw votes in the PR election arm.

# Appendix A    Figures

**Figure A1:**
**FEMALE SHARE IN NATIONAL PARLIAMENTS AND ATTITUDES TOWARDS WOMEN (2017-2022)**
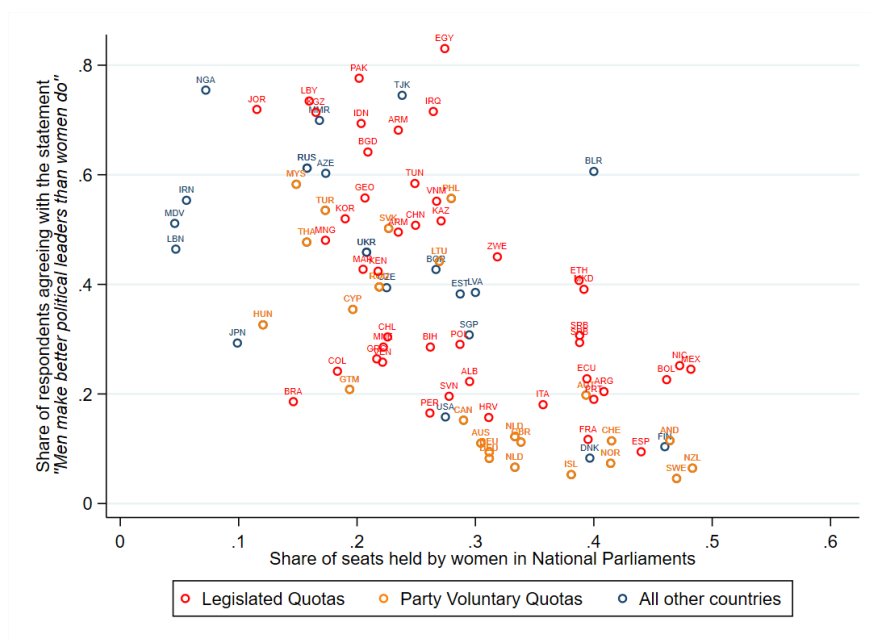


*Notes*: Sources:(i) Attitudes towards women - World Values Survey, 2017-2021  European Values Survey, 2017-2022; (ii) Share of seats held by women in national parliaments - World Bank Gender Statistics, latest data available between 2021-2018; (iii) Gender quotas in national parliaments - International IDEA Gender Quotas Database, last updated June 2022

## Figure A2:
## FEMALE SHARE IN NATIONAL PARLIAMENTS AND ATTITUDES TOWARDS WOMEN (1995-2004)
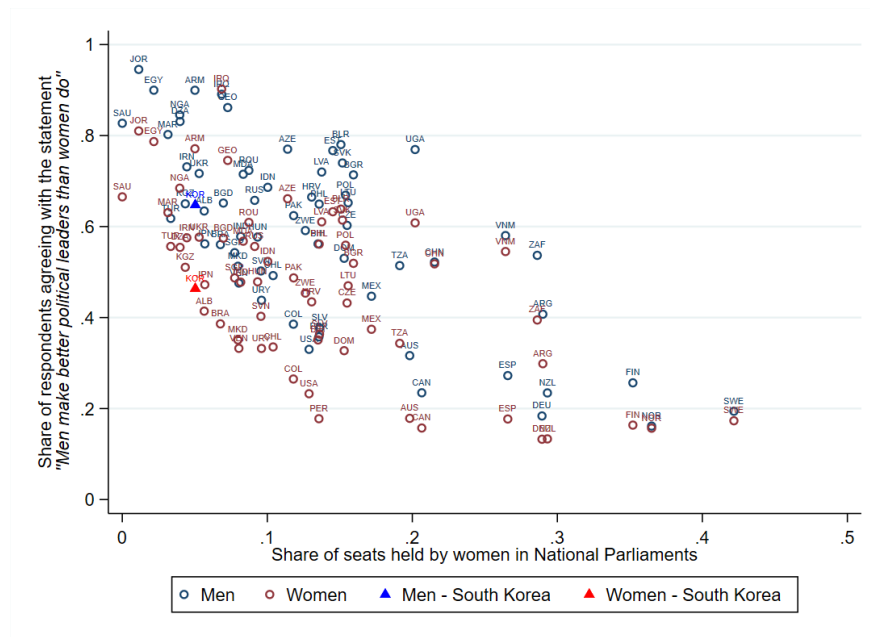


*Notes*: Sources - Attitudes towards women: World Values Survey waves 3 (1995-1998) and 4 (1999-2004); Share of seats held by women in national parliaments: World Bank Gender Statistics, average years 1995-2004.

## Figure A3:
## FEMALE REPRESENTATION ON BOARDS IN LISTED COMPANIES AND ATTITUDES TOWARDS WOMEN (2017-2022)
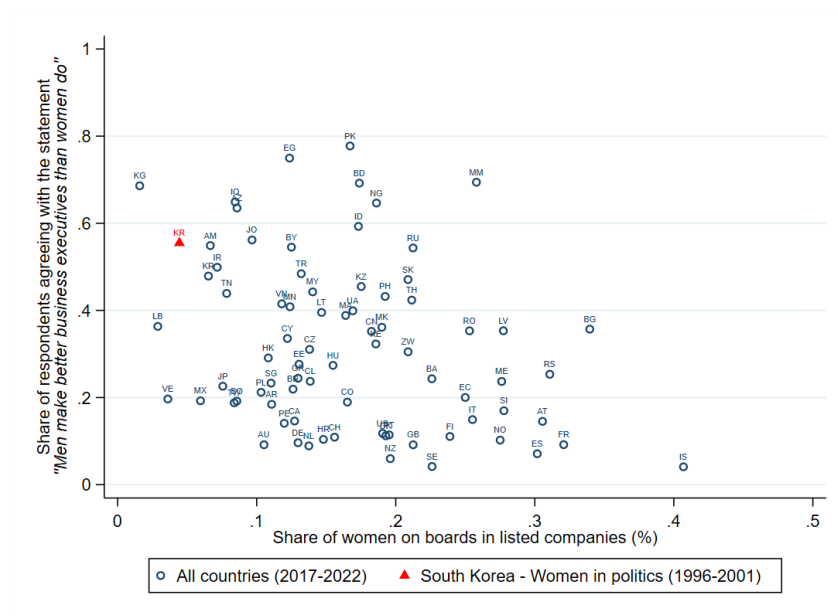


*Notes*: Sources: All countries - Attitudes: World Values Survey, wave 7 (2017-2022); Share of women on boards in listed companies: ORBIS, 2022. South Korea - Attitudes: World Values Survey, waves 3 (1996) and 4 (2001); Share of women in national parliaments: World Bank Gender Statistics, average years 1997-2001.

# Appendix B   Tables

<div align="center">

**Table A1:**
**THE ALLOCATION OF PROPORTIONAL REPRESENTATION SEATS ACROSS PARTIES**

</div>

| | **1 PR seat** | | **2 PR seats** | | **3 PR seats** | |
| --- | --- | --- | --- | --- | --- | --- |
| | N. | Percent. | N. | Percent. | N. | Percent. |
| **Election Cycle 4** | | | | | | |
| 1 Party | 117 | 100% | 15 | 17.86% | 0 | 0 |
| 2 Parties | 0 | 0 | 69 | 82.14% | 15 | 83.33% |
| 3 Parties | 0 | 0 | 0 | 0 | 3 | 16.67% |
| | | | | | | |
| **Election Cycle 5** | | | | | | |
| 1 Party | 117 | 100% | 5 | 6.02% | 0 | 0 |
| 2 Parties | 0 | 0 | 78 | 93.98% | 13 | 72.22% |
| 3 Parties | 0 | 0 | 0 | 0 | 7 | 27.78% |
| | | | | | | |
| **Election Cycle 6** | | | | | | |
| 1 Party | 110 | 100% | 18 | 20.22% | 0 | 0 |
| 2 Parties | 0 | 0 | 71 | 79.78% | 17 | 100% |
| 3 Parties | 0 | 0 | 0 | 0 | 0 | 0% |
| | | | | | | |
| **Election Cycle 7** | | | | | | |
| 1 Party | 105 | 100% | 9 | 9.89% | 1 | 4.35% |
| 2 Parties | 0 | 0 | 82 | 90.11% | 18 | 94.74% |
| 3 Parties | 0 | 0 | 0 | 0 | 1 | 5.26% |
| | | | | | | |
| **Total** | **449** | | **347** | | **72** | |

*Note: "Councils by the number of PR seats" spans the three PR seat column groups.*

*Notes*: The sample is restricted to bins 1 and 2, i.e. to municipal councils with up to 25 councilors.

**CANDIDATES AND COUNCILORS' PARTY AFFILIATION**

| | | Candidates | | | | Councilors | | | |
| | | Ward | | PR | | Ward | | PR | |
| | N | Mean | Std | Mean | Std | Mean | Std | Mean | Std |
|---|---|---|---|---|---|---|---|---|---|
| **Election Cycle 1** | | | | | | | | | |
| Independent | 226 | 1 | 0 | | | 1 | 0 | | |
| | | | | | | | | | |
| **Election Cycle 2** | | | | | | | | | |
| Independent | 228 | 1 | 0 | | | 1 | 0 | | |
| | | | | | | | | | |
| **Election Cycle 3** | | | | | | | | | |
| Independent | 228 | 1 | 0 | | | 1 | 0 | | |
| | | | | | | | | | |
| **Election Cycle 4** | | | | | | | | | |
| Independent | 230 | 0.41 | 0.15 | 0 | 0 | 0.11 | 0.14 | 0 | 0 |
| Progressive party | 230 | 0.16 | 0.08 | 0.29 | 0.17 | 0.20 | 0.18 | 0.18 | 0.25 |
| Conservative party | 230 | 0.25 | 0.13 | 0.47 | 0.30 | 0.54 | 0.31 | 0.64 | 0.38 |
| | | | | | | | | | |
| **Election Cycle 5** | | | | | | | | | |
| Independent | 228 | 0.32 | 017 | 0 | 0 | 0.14 | 0.16 | 0 | 0 |
| Progressive party | 228 | 0.21 | 0.16 | 0.31 | 0.29 | 0.33 | 0.27 | 0.41 | 0.39 |
| Conservative party | 228 | 0.33 | 0.18 | 0.49 | 0.31 | 0.43 | 0.25 | 0.44 | 0.40 |
| | | | | | | | | | |
| **Election Cycle 6** | | | | | | | | | |
| Independent | 227 | 0.34 | 0.16 | 0 | 0 | 0.13 | 0.15 | 0 | 0 |
| Progressive party | 227 | 0.24 | 0.16 | 0.38 | 0.27 | 0.37 | 0.25 | 0.40 | 0.37 |
| Conservative party | 227 | 0.36 | 0.18 | 0.55 | 0.30 | 0.49 | 0.26 | 0.59 | 0.38 |
| | | | | | | | | | |
| **Election Cycle 7** | | | | | | | | | |
| Independent | 226 | 0.20 | 0.16 | 0 | 0 | 0.09 | 0.14 | 0 | 0 |
| Progressive party | 226 | 0.32 | 0.12 | 0.43 | 0.20 | 0.54 | 0.21 | 0.66 | 0.33 |
| Conservative party | 226 | 0.29 | 0.16 | 0.38 | 0.25 | 0.35 | 0.20 | 0.33 | 0.33 |

# Appendix C   Identification

## C.1   Confirming that the number of female PR councilors changes only at the thresholds

In order to buttress the regression discontinuity design, we test whether there is a change in the number of female PR councilors as council size increases, at points *other* than the thresholds. We regress, for each value of $x \in \{-4, -3, ..., 3, 4\}$ i.e. distance from the threshold,

$$(\text{number of female PR councilors})_{cbt} = \beta \times TreatOne_{cbt} + \delta_b + \gamma_t + \epsilon_{cbt} \tag{8}$$

$$\text{where } TreatOne_{cbt} = \begin{cases} 1, & \text{if (council size)}_{cbt} = x \\ 0, & \text{if (council size)}_{cbt} = x - 1 \end{cases}$$

Equation (8), therefore, estimates the change in the number of female PR councilors when the council size increases by 1, for all points around the threshold. Table A3 reports the results. It confirms that there is a positive effect only at the threshold.

**Table A3:**
**THE EFFECT OF AN INCREASE IN COUNCIL SIZE ON THE NUMBER OF FEMALE PR COUNCILORS**

| | $x$ value | | | | | | | | |
| | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|---|---|---|---|
| Coefficient ($\hat{\beta}$) | -0.03 | 0.03 | -0.03 | -0.01 | 0.92*** | -0.01 | -0.03 | -0.03 | 0.09 |
| Standard error | (-0.36) | (1.32) | (-0.99) | (-0.20) | (15.00) | (-0.14) | (-0.36) | (-0.35) | (1.04) |
| $N$ | 267 | 380 | 210 | 170 | 168 | 150 | 136 | 111 | 87 |

*Notes*: This table reports the results of regression equation (8). The sample includes only bins 1 and 2. t statistics from standard errors clustered by municipality in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

## C.2 Robustness to bandwidth choice

**Table A4:**
**ROBUSTNESS: THE EFFECT OF BEING PAST THE THRESHOLD ON THE NUMBER OF CANDIDATES AND COUNCILORS, FOR VARIOUS BANDWIDTHS**

| | Candidates | | | | Councilors | | | | | |
| | Ward | | PR | | Ward | | PR | | All | |
| | Male (1) | Female (2) | Male (3) | Female (4) | Male (5) | Female (6) | Male (7) | Female (8) | Male (9) | Female (10) |
|---|---|---|---|---|---|---|---|---|---|---|
| ***Panel A*: distance$^{\dagger} \leq 4$** | | | | | | | | | | |
| Treat | 0.37 | 0.31 | 0.58*** | 1.23*** | -0.15 | 0.15 | 0.09** | 0.91*** | -0.06 | 1.06*** |
| | (0.46) | (0.94) | (4.64) | (8.29) | (-0.74) | (0.74) | (1.97) | (19.58) | (-0.26) | (4.78) |
| $N$ | 868 | 868 | 868 | 868 | 868 | 868 | 868 | 868 | 868 | 868 |
| | | | | | | | | | | |
| ***Panel B*: distance $\leq 3$** | | | | | | | | | | |
| Treat | 0.14 | 0.33 | 0.51*** | 1.24*** | -0.13 | 0.13 | 0.08 | 0.92*** | -0.04 | 1.04*** |
| | (0.16) | (0.99) | (3.77) | (7.95) | (-0.59) | (0.59) | (1.57) | (17.77) | (-0.19) | (4.48) |
| $N$ | 811 | 811 | 811 | 811 | 811 | 811 | 811 | 811 | 811 | 811 |
| | | | | | | | | | | |
| ***Panel C*: distance $\leq 2$** | | | | | | | | | | |
| Treat | 0.89 | 0.37 | 0.54*** | 1.29*** | -0.13 | 0.13 | 0.08 | 0.92*** | -0.06 | 1.06*** |
| | (0.99) | (1.07) | (3.59) | (7.50) | (-0.59) | (0.59) | (1.33) | (16.07) | (-0.23) | (4.16) |
| $N$ | 514 | 514 | 514 | 514 | 514 | 514 | 514 | 514 | 514 | 514 |
| | | | | | | | | | | |
| ***Panel D*: distance $\leq 1$** | | | | | | | | | | |
| Treat | 0.34 | 0.44 | 0.61*** | 1.25*** | -0.24 | 0.24 | 0.09 | 0.91*** | -0.15 | 1.15*** |
| | (0.34) | (1.17) | (3.55) | (6.66) | (-0.94) | (0.94) | (1.36) | (14.03) | (-0.53) | (4.02) |
| $N$ | 320 | 320 | 320 | 320 | 320 | 320 | 320 | 320 | 320 | 320 |
| | | | | | | | | | | |
| ***Panel E*: distance $= 0$** | | | | | | | | | | |
| Treat | 0.38 | 0.42 | 0.58*** | 1.23*** | -0.24 | 0.24 | 0.08 | 0.92*** | -0.16 | 1.16*** |
| | (0.37) | (1.10) | (3.26) | (6.53) | (-0.91) | (0.91) | (1.25) | (15.00) | (-0.55) | (3.99) |
| $N$ | 168 | 168 | 168 | 168 | 168 | 168 | 168 | 168 | 168 | 168 |

*Notes*: This table reports the results of regression equation (1). The sample includes only bins 1 and 2.
$^{\dagger}$ Distance refers to the distance to the threshold. To illustrate, the council sizes for which distance equals 0 are 10, 11, 20, and 21, while the council sizes for which distance equals 1 are 9, 12, 19, and 22. t statistics from standard errors clustered by municipality in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

## C.3   Robustness to choice of $f(x)$

**Table A5:**
**ROBUSTNESS: THE EFFECT OF BEING PAST THE THRESHOLD ($f(x) = x \times$ Treat)**

| | All political parties | | | | | | | |
| | Candidates | | | | Councilors | | | |
| | Ward | | PR | | Ward | | PR | |
| | Male | Female | Male | Female | Male | Female | Male | Female |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Treat at cycle 4 x Cycle 4 | 3.90*** | -0.45 | 0.95*** | 0.93*** | 0.51** | -0.51** | 0.09 | 0.77*** |
| | (1.22) | (0.38) | (0.22) | (0.21) | (0.23) | (0.23) | (0.08) | (0.09) |
| Treat at cycle 4 x Cycle 5 | 0.75 | 0.29 | 0.65*** | 1.18*** | -0.13 | 0.13 | 0.10 | 0.71*** |
| | (1.00) | (0.39) | (0.17) | (0.22) | (0.24) | (0.24) | (0.07) | (0.08) |
| Treat at cycle 4 x Cycle 6 | -1.20 | 0.70 | 0.27* | 1.00*** | -0.32 | 0.32 | 0.08 | 0.70*** |
| | (0.96) | (0.43) | (0.15) | (0.22) | (0.27) | (0.27) | (0.06) | (0.08) |
| Treat at cycle 4 x Cycle 7 | -2.04* | 0.89* | 0.23 | 1.25*** | -0.58* | 0.58* | 0.04 | 0.70*** |
| | (1.11) | (0.47) | (0.16) | (0.21) | (0.30) | (0.30) | (0.06) | (0.08) |
| Running variable form | ward | ward | council | council | ward | ward | council | council |
| $N$ | 868 | 868 | 868 | 868 | 868 | 868 | 868 | 868 |

*Notes*: This table reports the results of regression equation (1), with $f(x) = x \times Treat$. The sample includes only bins 1 and 2. t statistics from standard errors clustered by municipality in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

**Table A6:**
**ROBUSTNESS: THE EFFECT OF BEING PAST THE THRESHOLD ($f(x) = x + x^2$)**

| | All political parties | | | | | | | |
| | Candidates | | | | Councilors | | | |
| | Ward | | PR | | Ward | | PR | |
| | Male | Female | Male | Female | Male | Female | Male | Female |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Treat at cycle 4 x Cycle 4 | 3.90*** | -0.39 | 0.98*** | 0.88*** | 0.46** | -0.46** | 0.10 | 0.77*** |
| | (1.20) | (0.37) | (0.23) | (0.21) | (0.22) | (0.22) | (0.08) | (0.09) |
| Treat at cycle 4 x Cycle 5 | 0.75 | 0.35 | 0.67*** | 1.12*** | -0.18 | 0.18 | 0.10 | 0.71*** |
| | (0.97) | (0.38) | (0.18) | (0.23) | (0.23) | (0.23) | (0.07) | (0.08) |
| Treat at cycle 4 x Cycle 6 | -1.20 | 0.76* | 0.29* | 0.95*** | -0.37 | 0.37 | 0.09 | 0.70*** |
| | (0.93) | (0.43) | (0.16) | (0.22) | (0.27) | (0.27) | (0.07) | (0.08) |
| Treat at cycle 4 x Cycle 7 | -2.03* | 0.95** | 0.25 | 1.19*** | -0.63** | 0.63** | 0.04 | 0.70*** |
| | (1.08) | (0.46) | (0.17) | (0.21) | (0.30) | (0.30) | (0.06) | (0.08) |
| Running variable form | ward | ward | council | council | ward | ward | council | council |
| $N$ | 868 | 868 | 868 | 868 | 868 | 868 | 868 | 868 |

*Notes*: This table reports the results of regression equation (1), with $f(x) = x + x^2$. The sample includes only bins 1 and 2.  t statistics from standard errors clustered by municipality in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

# Appendix D  Model

## D.1  Validity of the assumptions on the party's choice of candidates

In this section, we provide evidence of the validity of the model assumptions by studying election probabilities for the candidates elected before the introduction of the quota. Municipal elections were introduced in 1995, so we can exploit three rounds of elections and information for a total of 18,636 male candidates[28]. Before the introduction of the quota, all councilors were directly elected through plurality vote in single-member constituent wards with no party affiliation. Hence, we can study how much voters value different characteristics without the potential confounding factors such as unobserved preferences for parties, differences in parties strategic behaviors across wards, or the influence of ballot list position.

Table A7 displays the results of this analysis. The probability of being elected and the candidates' vote share are regressed on municipality $\times$ election cycle fixed effects, and candidates' characteristics - age, education, a dummy for whether the candidate works in politics, and a dummy for whether the candidate was ever elected before. The table shows that an incumbency premium exists, older candidates have a higher probability of being elected (probably capturing social networks and the influence of the candidate in the community), and candidates with political experience get more votes. More relevantly for our modeling assumption, voters value the education of candidates. We can see that having a high school degree is associated with a 2.3 percentage points higher probability of being elected and 1 percentage point higher vote share compared to an elementary school certificate (omitted category), all else constant. In fact, the coefficients on the education degrees show that election outcomes improve monotonically with higher levels of education.

---

[28]We restrict the analysis to male candidates as women were rare and very selected.

**Table A7:**
**WHAT DO VOTERS CARE ABOUT?**

| | Pr(Elected) (1) | Vote Share (2) |
|---|---|---|
| Middle School | 0.008 | 0.005 |
| | (0.015) | (0.004) |
| High School | 0.023* | 0.010*** |
| | (0.013) | (0.004) |
| Undergraduate Degree | 0.044*** | 0.014*** |
| | (0.014) | (0.004) |
| Graduate Degree | 0.062*** | 0.023*** |
| | (0.020) | (0.006) |
| Incumbent | 0.156*** | 0.058*** |
| | (0.010) | (0.003) |
| Age | 0.049*** | 0.020*** |
| | (0.005) | (0.001) |
| Age squared | -0.000*** | -0.000*** |
| | (0.000) | (0.000) |
| Politician | 0.033** | 0.013*** |
| | (0.013) | (0.004) |
| $N$ | 18442 | 17401 |
| Locality $\times$ Cycle FEs | Yes | Yes |

*Notes*: The probability of being elected (Column 1) and the candidate's vote share (Column 2) are regressed on municipality $\times$ election cycle fixed effects, and candidates' characteristics - age, education, a dummy for whether the candidate works in politics, and a dummy for whether the candidate was ever elected before. The sample consists of the universe of male candidates for the three municipal elections before the introduction of the quota in 1995, 1998, and 2002 (19480 male candidates), for which we could retrieve occupation, education (95%), and vote share information (89%). t-statistics from standard errors clustered by municipality in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

## D.2 Extension: if the exact ability of councilors is not revealed

What if the exact ability of councilors is not revealed while they serve their term? Rather, for councilor $i$, a party receives a second signal of ability that is highly informative about the *absolute* ability of $i$:

$$v_i \sim N(a_i, \sigma_v^2)$$

where $\sigma_v^2$ is a small number. Moreover, say the precision of $v_i$ is inversely related to the closeness of the interaction between councilor $i$ and a party. For instance,

$$\sigma_v^2 = \begin{cases} \sigma_1^2 & \text{if } i \text{ belongs to own party} \\ \sigma_2^2 & \text{else} \end{cases}$$

with $\sigma_1^2 < \sigma_2^2$.

Once the values of the second signals of ability of councilors, $\mathbf{v} = \{v_i\}$, are revealed, the party

makes an inference about the value of $\mu_g$ via maximum likelihood as before.

Conditional on $s_i$, the distribution of $v_i$ is

$$v_i|s_i \sim N\left(\mu_g + \frac{\sigma^2}{\sigma_s^2 + \sigma^2}s_i, \sigma_v^2 + \frac{\sigma^2\sigma_s^2}{\sigma_s^2 + \sigma^2}\right) \tag{9}$$

Let's call $\bar{\sigma}_1^2 = \sigma_1^2 + \frac{\sigma^2\sigma_s^2}{\sigma_s^2+\sigma^2}$ and $\bar{\sigma}_2^2 = \sigma_2^2 + \frac{\sigma^2\sigma_s^2}{\sigma_s^2+\sigma^2}$. $N_1$ denotes the set of own-party councilors of size $n_1$, and $N_2$ the set of other councilors of size $n_2$.

The likelihood function is

$$
\begin{aligned}
\mathcal{L}(\mu_g) &= P(\mathbf{v}|\mathbf{s};\mu_g) \\
&= \Pi_{i\in N_1}f(v_i|s_i;\mu_g,\sigma_1^2) \times \Pi_{i\in N_2}f(v_i|s_i;\mu_g,\sigma_2^2) \\
&= \frac{1}{(\sqrt{2\pi}\bar{\sigma}_1)^{n_1}}\exp\left(-\frac{1}{2\bar{\sigma}_1^2}\sum_{i\in N_1}\left(v_i - m(s_i)\right)^2\right) \times \frac{1}{(\sqrt{2\pi}\bar{\sigma}_2)^{n_2}}\exp\left(-\frac{1}{2\bar{\sigma}_2^2}\sum_{i\in N_2}\left(v_i - m(s_i)\right)^2\right)
\end{aligned}
$$

Then the maximum likelihood estimator is

$$\hat{\mu}_g = \frac{\bar{\sigma}_2^2\sum_{i\in N_1}\left(v_i - cs_i\right) + \bar{\sigma}_1^2\sum_{i\in N_2}\left(v_i - cs_i\right)}{\bar{\sigma}_2^2 n_1 + \bar{\sigma}_1^2 n_2} \tag{10}$$

The distribution of $\hat{\mu}_g$ is

$$\hat{\mu}_g \sim N\left(\mu_g, \frac{\bar{\sigma}_1^2\bar{\sigma}_2^2}{\bar{\sigma}_2^2 n_1 + \bar{\sigma}_1^2 n_2}\right) \tag{11}$$

The posterior distribution about the value of $\mu_g$ follows the same structure as before.

# Appendix E  Mechanisms

## E.1  Computing the running variable in the regression discontinuity design of section 7.3

The purpose of the regression discontinuity design of section 7.3 is to compare the strategies of parties that marginally won a PR councilor to those that marginally lost a PR councilor in the previous election. Thus, we are interested in the causal effect of having won a female PR councilor. We take marginal parties to be the *two* parties that either marginally won or lost the *last* PR seat. In order to differentiate marginal winners from losers, we measure how far off the vote share received by a party was, from the share it needed to win that seat. The running variable for party $p$ in municipal council $c$ at election cycle $t$ equals $v_{cpt} = voteshare_{cp,t-1} - \bar{v}_{cp,t-1}$, where $\bar{v}$ denotes the

verdict-determining vote share.

To compute $\bar{v}$, we first need to describe the rules by which PR seats get allocated:

---

**Rules for allocating PR seats**

1. Among parties running for prop rep in a municipality, only the parties getting >=5% of votes qualify.

2. Of the qualifying parties, first compute $X$ =(number of prop MP seats in the municipality)×(vote share of each qualifying party).

3. Allocate to each qualifying party the number of seats equal to the integer part of $X$.

4. Allocate the remaining seats by the ranking of the decimal part of $X$.

*E.g.* Municipal council A has 3 PR seats. There are 3 parties (1, 2, and 3) running for proportional representation. The vote shares of the parties are: party 1: 60%, 2: 38%, and 3: 2%. Party 3 got less than 5%, so it does not qualify. Among the qualifying parties, the vote shares are then party 1: $60/(60 + 38) \approx 61.22\%$, and 2: $38/(60 + 38) \approx 38.77\%$. The values of $X$'s are party 1: $3 \times 0.6122 \approx 1.83$, and 2: $3 \times 0.3877 \approx 1.16$. Parties 1 and 2 both have 1 in the integer part of X, so they first get one PR councilor each. The last PR seat goes to party 1, because $0.83 > 0.16$.

---

Below, we compute $\bar{v}$ for all possible contest scenarios.[29] While doing so, we distinguish whether the marginal candidate that won – or nearly won – the last PR seat corresponds to the first, second, or third PR candidate in a party's list. For notational convenience, we call $s$ the position in the party list of the marginal candidate of a party, and $V$ the sum of the vote shares (among qualifying parties) received by the two marginal parties.

$\boxed{1}$ When there is one PR seat in the municipality

  i) The two most popular parties contest over the only PR seat. Marginal parties: ranks 1 and 2

   − Rank 1: $\bar{v} = \frac{V}{2}$, $s = 1$
   − Rank 2: $\bar{v} = \frac{V}{2}$, $s = 1$

---

[29] An example is when there are three PR seats in a municipality, and the rank-1 and rank-2 parties contest over the last seat. Let $v_n$ denote the vote share (among qualifying parties) received by the rank-$n$ party. Rank 1 wins if $3v_1 - 2 > 3v_2 \iff v_1 > v_2 + \frac{2}{3}$. Therefore, $\bar{v}$ for the rank-1 party equals $v_2 + \frac{2}{3}$. On the other hand, $\bar{v}$ for the rank-2 party equals $v_1 - \frac{2}{3}$.

$\boxed{2}$ When there are two PR seats in the municipality

   i) The contest is over whether the rank-2 party wins the second PR seat. Marginal parties: ranks 1 and 2

     – Rank 1: $\bar{v} = \frac{2V+1}{4}$, $s = 2$

     – Rank 2: $\bar{v} = \frac{2V-1}{4}$, $s = 1$

$\boxed{3}$ When there are three PR seats in the municipality

   i) The contest is over whether the third PR seat goes to the rank-1 party or the rank-2 party. Marginal parties: ranks 1 and 2

     – Rank 1: $\bar{v} = \frac{3V+2}{6}$, $s = 3$

     – Rank 2: $\bar{v} = \frac{3V-2}{6}$, $s = 1$

   ii) Where the rank-2 party wins a seat for sure, the contest is over whether the third PR seat goes to the rank-1 party or the rank-3 party. Marginal parties: ranks 1 and 3

     – Rank 1: $\bar{v} = \frac{3V+1}{6}$, $s = 2$

     – Rank 3: $\bar{v} = \frac{3V-1}{6}$, $s = 1$

$\boxed{4}$ When there are four PR seats in the municipality

   i) The contest is over whether the fourth PR seat goes to the rank-1 party or the rank-2 party. Marginal parties: ranks 1 and 2

     – Rank 1: $\bar{v} = \frac{4V+3}{8}$, $s = 4$

     – Rank 2: $\bar{v} = \frac{4V-3}{8}$, $s = 1$

   ii) Where the rank-1 party wins two seats for sure and the rank-2 party wins a seat for sure, the contest is over whether the fourth PR seat goes to the rank-1 party or the rank-2 party. Marginal parties: ranks 1 and 2

     – Rank 1: $\bar{v} = \frac{4V+1}{8}$, $s = 3$

     – Rank 3: $\bar{v} = \frac{4V-1}{8}$, $s = 2$

   iii) Where the rank-1 party wins two seats for sure and the rank-2 party wins a seat for sure, the contest is over whether the fourth PR seat goes to the rank-2 party or the rank-3 party. Marginal parties: ranks 2 and 3

     – Rank 2: $\bar{v} = \frac{4V+1}{8}$, $s = 2$

     – Rank 3: $\bar{v} = \frac{4V-1}{8}$, $s = 1$

iv) Where the rank-2 and rank-3 parties win a seat each for sure, the contest is over whether the fourth PR seat goes to the rank-1 party or the rank-4 party. Marginal parties: ranks 1 and 4

- Rank 1: $\bar{v} = \frac{4V+1}{8}$, $s = 2$
- Rank 4: $\bar{v} = \frac{4V-1}{8}$, $s = 1$

As an example, take the case of the rank-1 party in a municipality with two PR seats and two qualifying parties. The party's $\bar{v} = 0.75$, according to the computation given above. Indeed, Figure A4 shows that among such rank-1 parties, those receiving a vote share greater than 0.75 win two PR councilors whereas those receiving a vote share below 0.75 win one PR councilor.

**Figure A4:**
**MARGINAL WINNERS AND LOSERS OF THE LAST PR SEAT, AMONG RANK-1 PARTIES IN MUNICIPALITIES WITH TWO PR SEATS AND TWO QUALIFYING PARTIES**
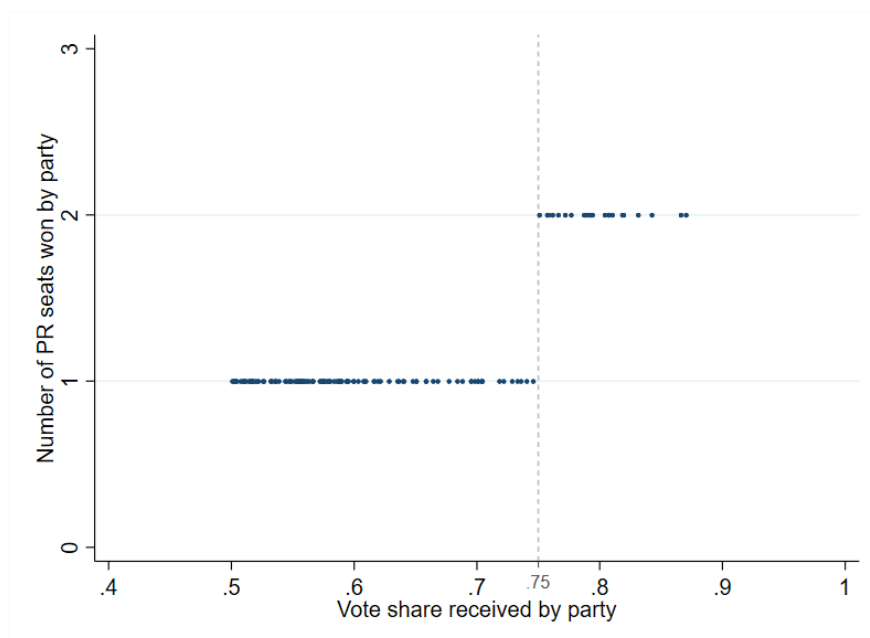


*Notes*: This figure shows that in municipalities with two PR seats and two qualifying parties, the rank-1 parties must receive a vote share greater or equal to 0.75 in order to win both PR seats. The reason the vote share received is always greater than 0.5 is because these parties are the rank-1 parties. Note that the vote share is the share of votes among qualifying parties only.

## E.2 Robustness to the choice of $f(x)$ in the regression discontinuity design of section 7.3

**Table A8:**

**ROBUSTNESS: THE EFFECT OF MARGINALLY WINNING A PR WOMAN IN THE PREVIOUS ELECTION**

| | Main political parties | | | |
|---|---|---|---|---|
| | $\mathbb{1}$(Number-1 PR candidate in $t-1$ is a ward candidate in $t$) | | | |
| Bandwidth ($|v_{cpt}|$) | 0.20 | 0.15 | 0.10 | 0.05 |
| | (1) | (2) | (3) | (4) |
| **Panel A: All parties** | | | | |
| $f(x) = x$ | 0.44*** | 0.43*** | 0.35*** | 0.42*** |
| | (5.63) | (5.39) | (3.47) | (3.17) |
| $f(x) = x + x^2$ | 0.40*** | 0.41*** | 0.36*** | 0.43*** |
| | (5.59) | (5.35) | (3.88) | (3.41) |
| $f(x) = (x + x^2) \times Treat$ | 0.39*** | 0.39*** | 0.55*** | 0.75*** |
| | (4.10) | (3.53) | (4.24) | (4.63) |
| $N$ | 414 | 313 | 216 | 114 |
| **Panel B: 2nd PR candidate = Man** | | | | |
| $f(x) = x$ | 0.47*** | 0.52*** | 0.40*** | 0.47*** |
| | (5.64) | (5.84) | (3.67) | (3.00) |
| $f(x) = x + x^2$ | 0.44*** | 0.49*** | 0.41*** | 0.47*** |
| | (5.76) | (5.73) | (4.06) | (3.30) |
| $f(x) = (x + x^2) \times Treat$ | 0.47*** | 0.46*** | 0.58*** | 0.72*** |
| | (4.49) | (3.72) | (3.90) | (3.63) |
| $N$ | 308 | 241 | 165 | 90 |
| **Panel C: 2nd PR candidate = Woman** | | | | |
| $f(x) = x$ | 0.37* | 0.20 | 0.11 | 0.03 |
| | (1.88) | (0.99) | (0.45) | (0.09) |
| $f(x) = x + x^2$ | 0.23 | 0.22 | 0.12 | -0.07 |
| | (1.19) | (1.18) | (0.54) | (-0.18) |
| $f(x) = (x + x^2) \times Treat$ | 0.02 | 0.12 | -0.25 | -0.46 |
| | (0.09) | (0.39) | (-0.65) | (-0.76) |
| $N$ | 106 | 72 | 51 | 24 |

*Notes*: The Table reports the result of specification (6) on the share of women among ward and PR candidates in cycle $t$ (Panel A), and the share of women among ward and PR candidates in cycle t who were the number-1 PR candidate in cycle $t-1$ (Panel B). Two different specifications are considered for the functional form of $f(v_{cp,t-1})$. The first row reports the coefficient of the $Winner_{t-1}$ dummy assuming $f(x) = x$. The second row assumes $f(x) = (x + x^2) \times Treat$. The results are provided for all parties, parties for which the second PR candidate in cycle $t-1$ was a man, and parties where the second PR candidate was a woman. Each column shows the coefficient of the $Winner_{t-1}$ dummy considering different bandwidths. The standard errors are clustered at the municipality $\times$ party level. t statistics from standard errors clustered by municipality$\times$party in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.