

The Gender Gap in Political Careers Under Proportional Representation

Tobias Nowacki*

27th January 2022

Abstract

Does proportional representation feature gender gaps in incumbency advantages? I argue that closed-list PR (compared to open-list PR) can help to attenuate the gender difference by reducing voters' influence over individual candidates' electoral outcomes. Using difference-in-discontinuity designs, I contrast gender-specific incumbency effects in municipal elections across Norway (open-list) and Spain (closed-list). In Norway, women experience an incumbency advantage in winning future elections that is up to 60% smaller than men's; in Spain, female candidates enjoy incumbency effects equal to or greater than men's. Additional results suggest that these findings are unlikely to be due to unobserved country characteristics or differential selection. Further evidence also points towards smaller increases in preference votes for women in right-wing parties as a key mechanism for the gender gap in open-list PR. All told, this paper highlights the importance of electoral rules in ensuring gender equity in political representation by addressing leaky pipelines in office.

Keywords: incumbency advantage; gender gap; local elections; proportional representation; electoral systems

Word count: 11,971

*Ph.D. Candidate, Stanford University, Calif., USA. tnowacki@stanford.edu. I thank Kirk Bansak, Chris Berry, Gary Cox, Andy Eggers, Jon Fiva, Olle Folke, Ignacio Lago, Justin Grimmer, Andy Hall, Apoorva Lal, Sandy Handan-Nader, Jens Hainmueller, Jonathan Rodden, Brett Parker, Tine Paulsen, Soledad Prillaman, Johanna Rickne, Dan Smith, Dan Thompson, and participants in the Democracy and Polarization Lab and at the APSA 2021 panel for helpful comments.

Women are underrepresented in politics, which is concerning for intrinsic normative reasons, as well as for its consequences for downstream policy outcomes (Mansbridge, 1999, 2003; Chattopadhyay and Duflo, 2004; Dahlerup, 2013). A rich literature explains the empirical fact that countries using proportional representation (PR) are more likely to feature higher proportions of women elected to the legislature – and thus less ‘sticky’ floors (Norris, 1985; Matland and Studlar, 1996; Matland, 1998; Thames and Williams, 2010; Roberts, Seawright, and Cyr, 2013; Casas-Arce and Saiz, 2015; Thames, 2017; Golder et al., 2017; Verge and Wiesehomeier, 2019; Lucardi and Micozzi, 2020).¹ Yet, simply counting the number of women elected to the legislature is insufficient for achieving gender equity in politics: their career trajectory, time in office, and political experience, as well as their efficiency as policymakers representing women’s preferences all matter for this objective (Smrek, 2020). If female officeholders suffer from a ‘leaky pipeline’ (Cipullo, 2021) in their careers – that is, once elected, they are still less likely to win re-election and promotion to higher offices – concerns about unequal representation will persist.² Do PR systems struggle with a gender gap in political career trajectories? More specifically, is there a gender gap in the effect of incumbency on re-election and future career outcomes?

In this paper, I study how different forms of PR (open versus closed-list) shape the gender gap in the effect of winning office on downstream election outcomes – typically the next step in the career trajectory after winning office. I argue that elections using closed-list PR –

¹ Though reduced, sticky floors still represent significant hurdles for women entering the political arena (Folke and Rickne, 2016).

² Recent evidence from single-member, plurality systems suggests that women might be less persistent in pursuing political careers (Wasserman, 2020) and suffer from a lower probability of promotion to higher office (Brown et al., 2019; Cipullo, 2021), though see also Bernhard and de Benedictis-Kessner (2021) and Folke and Rickne (2020). In the context of proportional representation, Piscopo et al. (2021) show that female newcomers in Chile suffer from a disadvantage in access to campaign finance.

compared to those with open-list PR – are more likely to plug the ‘leaky pipeline’ by limiting voters’ influence over individual candidates’ ranking, which may otherwise disadvantage women disproportionately due to voter or systemic bias. In closed-list PR, incumbency advantages arise from party elites’ decisions to elevate candidates throughout list ranks. Though party elites may have initial negative biases against women, they have little reason to treat bare winners (compared to bare losers) differentially by gender (Luhiste, 2015; Hazan and Rahat, 2006).³ By contrast, in open-list systems, the effect of being elected on future election outcomes may be differentially greater for men due to a larger increase in personal votes for barely elected male candidates. This gap will be especially pronounced where voters are biased against women, or systemic disadvantages (e.g., smaller name recognition bonuses for women as an effect of incumbency) exist. Overall, this framework implies that women may face diminished incumbency effects in open-list PR systems, while we can expect an equal or even greater incumbency advantage for women in closed-list PR systems.

Empirical results from difference-in-discontinuity designs in municipal elections – typically the first rung in the political career ladder (Cirone, Cox, and Fiva, 2021) – in Norway and Spain are consistent with the theoretical argument. In Norway’s open-list elections, barely elected women enjoy a significantly smaller incumbency advantage compared to their male colleagues. Whereas men who barely won (compared to men who barely lost) have an approximately 10 percentage point higher probability of winning in the next election, the effect decreases to about 4 percentage points for women who barely won (compared to women who barely lost) – a 60% decline vis-a-vis men.⁴ Women in Spain’s closed-list

³ If anything, their initial bias towards women should diminish upon observing them perform in office (Kjaer and Krook, 2019)

⁴ Voters award preference votes to candidates, which determines their list ranking within the set of candidates fielded by the party. Voter preferences have a meaningful impact on deciding what candidates

elections, by contrast, likely enjoy a differentially greater incumbency advantage than their male colleagues. Importantly, I find no evidence of large gender gaps in career *persistence* in either setting: the effect of winning on the probability of running again is likely similar for men and women in either country.

I supplement the difference-in-discontinuity design with additional evidence to guard against concerns about case selection and differential selection into the threshold sample. While the difference-in-discontinuity design causally identifies the incumbency advantage within each gender, male and female candidate threshold samples may differ in individual-level (age, experience) as well as election-level and geographic attributes. I find that the magnitude of the gender gap in open-list persists when restricting the comparison to female candidates and the most similar male candidate with respect to age, experience and election year in the same party and county. These matching estimates suggest that gender, rather than a correlated covariate or differential selection, likely explains the effect heterogeneity between male and female candidates (Teele, Kalla, and Rosenbluth, 2018; Bansak, 2021). Second, if the magnitude of differential selection varies across cases, my contrast between open-list and closed-list PR may be an artefact of case selection. To address this concern, I leverage both within- and cross-country variance in electoral rules using data from Norway and Poland to show that the contrast between open- and closed-list PR, rather than idiosyncratic country effects, can explain the gender gap in incumbency advantages. Jointly, these results further point towards the role of electoral rules in affecting the gender gap in incumbency advantages.

get elected: Bergh, Bjørklund, and Hellevik (2010) estimate about 25% of candidates won office because voters' preference votes, rather than list ranking. While more complex classifications of PR systems may describe Norway's system as 'flexible-list', I follow the existing literature that recognizes Norway's municipal elections as 'open-list' (Fiva and Røhr, 2018; Fiva, Folke, and Sørensen, 2018).

Finally, I provide suggestive evidence for the theorized mechanism behind the negative gender gap in Norway – voters awarding a smaller increase in preference votes to barely elected women as a result of attaining incumbency. Specifically, although women are unlikely to face a differential penalty imposed by their *party* for being elected (Kjaer and Krook, 2019), the effect of winning in right-wing parties results in a far smaller increase, or potentially even decrease, in personal preference votes cast by voters for women when compared to men. Notably, no such discrepancy exists among candidates in left-wing parties, whose voters are more positively predisposed towards women in local politics. Voters’ influence over list ranking – the key difference between open and closed-list PR – thus likely contributes to the stark difference in incumbency advantage gender gaps across PR systems.

This paper connects and contributes to a number of existing literatures. It offers a new explanation for how electoral rules can affect gender equity in politics (Krook and Norris, 2014; O’Brien and Rickne, 2016; Krook, 2018; Luhiste, 2015). Although existing work suggests that voter biases that affect women’s *unconditional* probability of being elected in the first place may be context- and election-specific (Folke and Rickne, 2016; Anzia and Bernhard, 2019; Beaman et al., 2009; Broockman and Soltas, 2020; Eymoud and Vertier, 2017), this paper is among the first to study how candidates’ gender affects incumbency advantage mechanisms under PR (Schwindt-Bayer, 2005; Jankowski and Marcinkiewicz, 2019; Smrek, 2020), and whether voters reward incumbency differentially by gender. Beyond making the important distinction between open- and closed-list PR, the paper also offers a benchmark to evaluate the incumbency gender gap in PR systems against plurality (Wasserman, 2020; Cipullo, 2021; Brown et al., 2019). Finally, the paper speaks to the broader literature on political careers (Cirone, Cox, and Fiva, 2021; Kerevel, 2019; Folke, 2014) and offers an im-

portant theoretical distinction between the mechanisms driving incumbency advantages in open and closed-list PR systems more generally.

1 Electoral Systems and the Gender Gap in Political Careers

1.1 PR Boosts Number of Women Running (and Winning)

Proportional representation (PR) can boost the number of female candidates running (and potentially winning) for a number of reasons: Where party elites have control over nominations, they may follow gender-agnostic seniority norms (Fiva and Røhr, 2018; Cirone, Cox, and Fiva, 2021); incumbency advantages are less severe than in plurality, thus allowing ascending women to replace men more easily (Fiva, Folke, and Sørensen, 2018; Meserve, Pemstein, and Bernhard, 2020); in multi-member districts, nominating elites no longer have to cater to the ‘lowest common denominator’ and can instead attempt a more equitable distribution of viable spots within a district (Matland, 1998). Finally, PR systems can allow for a more fragmented party system in which left-leaning parties with a strong emphasis on gender equity push forward with nominating more women into viable spots, and other parties are forced to follow (Matland and Studlar, 1996).

Although increasing women’s numerical representation on the ballot and in office is essential for ensuring equity in downstream policy outcomes, it is not sufficient on its own. While existing work points towards PR attenuating the problem of a “sticky floor” for women, it remains an open question whether it can also close the “leaky pipeline” along women’s career

trajectories (Cipullo, 2021). This question is distinct from whether voters or parties treat female candidates differently unconditionally: once the conditional set of successful candidates is elected, male and female candidates may still interact differently with a given electoral system’s incumbency advantage and career promotion mechanisms.

Evaluating the gender gap in political careers under PR is thus important for improving representation and democratic legitimacy. If women, once elected, have shorter tenures in office or miss out on promotion to senior office, their representation might still carry less weight, as they wield less power and experience in legislative processes. Moreover, if women suffer from diminished effects of being elected on future career progression, this might also contribute to a greater gender gap in representation at higher levels of political office.

One key aspect that may drive the gender gap in political careers is whether there is a difference in the effect of being elected on running again and winning again as the most proximate downstream outcomes of winning office. Any gender differential in the effect of being re-elected will have knock-on effects on downstream career outcomes (such as running for or winning election to the national legislature). Studying this first outcome therefore sheds light on where the ‘leaky pipeline’ begins. While scholars have begun to examine this question in the context of plurality elections, (Brown et al., 2019; Bernhard and de Benedictis-Kessner, 2021; Wasserman, 2020; Cipullo, 2021), it remains understudied in the domain of PR (Smrek, 2020).

1.2 When Does PR Create Similar Incumbency Advantages for Men and Women?

How, if at all, should we expect incumbency advantages under PR to differ by candidates' gender? More candidate-centred (i.e., open-list) systems, in which voters can determine the ranking of individual candidates, may render women more vulnerable to voter and systemic biases, and lead to a smaller incumbency advantage compared to men.

Table 1 summarizes the argument. The primary incumbency advantage mechanism in closed-list systems – rank increases due to seniority – is unlikely to produce a gender gap, even if party elites are unconditionally biased against women. By contrast, a key mechanism in open-list PR systems – an increase in personal votes due to name recognition – may be reduced for women and produce a smaller incumbency advantage for female candidates if voters are biased or update less on female incumbents due to systemic bias. Altogether, closed- and open-list PR systems have different mechanisms through which the incumbency advantage operates. This, in turn, affects the gender gap in political careers – irrespective of any unconditional differences. Below, I elaborate on the incumbency mechanisms and their vulnerability towards gender biases in each setting.

Closed-list PR. In closed-list PR, the rank order of list positions decides individuals' electoral fortunes: higher list ranks translate into greater electoral security. Mechanically – as long as voters' decision what party to vote for does not depend on the position of individual candidates – any incumbency advantage for barely elected politicians must come

	Closed-list	Open-list
Incumbency Effect Mechanisms		
<i>Party-driven</i>	Seniority (rank advance)	Seniority (pre-vote rank advance)
<i>Voter-driven</i>		Name Recognition (pref vote increase)
Gender Gap Mechanisms	Social Learning (+)	Social Learning (+) Voter Bias (-) Systemic Bias (-)

Table 1: Incumbency Advantage Mechanisms under Proportional Representation

from an improved list rank.⁵ Put differently, candidates who are just barely elected move up to a safer list position with a higher chance of winning office in the next election. This mechanism is often institutionalized as a seniority norm in closed-list PR systems (cf. [Cirone, Cox, and Fiva, 2021](#)).

Under closed-list PR, we should expect party elites’ behavior to show no meaningful difference regarding the incumbency advantage between men and women. Party leadership may still exhibit unconditional negative biases towards women – for example, in the form of lower initial placements or a penalty to their perceived ability ([Luhiste, 2015](#); [Murray, Krook, and Opello, 2012](#); [Krook and Norris, 2014](#)). If a female candidate overcomes the initial bias (e.g. with greater ability) and proceeds to ascend through the list ranks towards the set of borderline elected candidates, it is hard to imagine why parties would further increase their negative bias of women as a result of being elected. Another reason why parties ought to be consistent in their treatment of borderline candidates is strategy: if the party were to deny women an incumbency advantage, they would be less likely to run in the first place.

⁵ The larger the district magnitude, the easier this assumption becomes to satisfy. With dozens of candidates on a party list, voters are unlikely to switch their ballot to another party because of a single candidate that they dislike.

One corollary of this argument is that female candidates in the sample of elected politicians may likely be, on average, of higher quality than men (Ashworth, Berry, and Bueno de Mesquita, 2021).⁶ If anything, then, parties should reduce their negative stereotypes upon observing a female candidate in office perform well: their perception should move closer to the true ability. If this is the case, and the initial perception was biased downward, women should actually experience a *greater* incumbency advantage than men. An important caveat to this argument is that women may suffer from lower persistence in political careers even after winning elections due to asymmetric outside factors (e.g., higher divorce rates, see Folke and Rickne (2020)), or more general differences in persistence or motivation to run for office (Lawless and Fox, 2005; Wasserman, 2020; Ashworth, Berry, and Bueno de Mesquita, 2021)).

Open-list PR. In open-list (and more candidate-centred) PR systems, part of the incumbency advantage comes from voters awarding elected candidates a boost in preference votes, therefore making it more likely that they will be re-elected (Dahlgaard, 2016; Dettman, Pepinsky, and Pierskalla, 2017; Jankowski and Müller, 2021).⁷ In lower-level, low-information elections, such as municipal elections, voters rarely observe the quality of candidates precisely. Instead, they might follow established heuristics in the form of incumbency status, (pre-vote) list rank, and name recognition. Voters typically only have a limited number of preference votes (e.g. in Norway, one-quarter of the council size) that they can allocate

⁶ This follows from women having to overcome the bias term in the first place, similar to Ashworth, Berry, and Bueno de Mesquita (2021).

⁷ Fiva and Røhr (2018), also studying Norwegian local elections, find the overall effect of incumbency on personal vote shares to be small and statistically indistinguishable from zero. This is likely the result of a particular operationalisation of personal vote share. I discuss this issue further in Appendix E.4 and show that my results – demonstrating that there is a meaningful effect for male candidates – hold up across different operationalisations.

across the party list.⁸

In this context, women may experience a diminished incumbency advantage if a subset of voters awards their preference votes to incumbent men, but not women. Consider a pool of voters in which some voters allocate preference votes based on candidates’ true ability (perhaps with a noisy, but unbiased signal), while others use name recognition and incumbency as a heuristic; a subset of these heuristic voters only awards preference votes to male incumbents. Women may still be able to win initial election, carried forward by unbiased voters. However, barely elected men will experience a greater increase in preference votes than barely elected women thanks to male-favoring heuristic voters. (Appendix A illustrates this argument further with a toy model simulation). With a smaller increase in preference votes, women will also observe a smaller increase in their probability of winning again as an effect of barely winning.⁹

There may be multiple, observationally equivalent, reasons why some voters do not award preference votes to female incumbents. One answer may be simple voter bias. A rich literature studying unconditional voter biases against women yields ambiguous results (Krook and Norris, 2014; Krook, 2018; Golder et al., 2017; Ragauskas, 2021). Survey experiments and quasi-experimental evidence conducted in plurality settings suggest that unconditional voter biases against women exist, but may be context- and party-specific (Anzia and Bernhard,

⁸ If parties can still decide the order in which candidates appear on the ballot, they may still be able to affect candidates’ success, which is strongly correlated with the initial list position (Appendix E.5). In that sense, parties in open-list determine which candidates enter the set of *viable* candidates, but do not have fully discrete power to decide who gets elected.

⁹ My theoretical argument does not make a claim about an unconditional difference in candidate type between male and female candidates. If women suffer from a ‘sticky floor’, the entire threshold sample of female candidates may be of higher quality compared to the respective male one. Yet, any alternative explanation focussed on differential selection needs to explain why there is an additional effect for men (or women) coming into force as a result of being borderline elected.

2019; Eymoud and Vertier, 2017; Horiuchi, Smith, and Yamamoto, 2020; Cipullo, 2021). It is less costly for voters to act on their biases in an open-list PR system: they can vote for co-partisans and thus need not trade off ideological proximity with a preference for candidates of a certain sex. Existing literature also focusses on voter biases preventing women from being elected in the first place; it remains an open question through what channels voter bias may affect typical incumbency mechanisms, and whether voters change their biases in response to women getting elected.¹⁰ Finally, systemic bias, rather than voter bias, may account for the gender differential in preference vote increases: if local media report less frequently or less positively about female incumbents, then they may attain smaller name recognition bonuses, and will accrue smaller preference vote boosts despite (heuristic) voters themselves being unbiased.

In sum, my argument leads to the following expectation for open-list PR: the differential increase in preference votes as an effect of winning is an incumbency advantage mechanism that is voter-driven, therefore only available in open-list PR, and will only kick in for incumbent men and women, but not for barely losing candidates. As a consequence, we should observe a smaller incumbency advantage for women compared to men.

2 Data and Design

How can we study the gender gap in political careers, and incumbency advantages more specifically, under proportional representation? In this section, I describe my case selection, data sources, and research design.

¹⁰ For example, male-preferring voters may increase their observable bias after the election of a women as they update on the probability of a women elected and/or societal preferences.

2.1 Cases and Data

In an ideal world, my case selection would exploit within-country (and within-election) variation in electoral rules to mitigate the concern of unobserved cross-country differences. Unfortunately, no such straightforward setting exists that would allow for estimation with sufficient statistical power.¹¹ Instead, I turn to a comparison of local elections across countries. By studying local elections – the typical entry point for political careers – we can understand whether electoral systems can have repercussions for women’s political careers that cascade into higher levels of office, and contribute to fewer and less experienced female candidates in more senior offices (Cirone, Cox, and Fiva, 2021).

To estimate the gender gap in incumbency advantages, I use local elections between 2003 and 2019 in two countries – Norway (open-list) and Spain (closed-list), along with supplementary data from Poland’s open-list county elections. In the closed-list system of Spain, voters can only cast their ballot for parties, not candidates: here, the order of list rankings is *ex ante* determined by the local party leadership. By contrast, in the Norwegian open-list system, voters cast preferences for individual candidates *within* a party, which ultimately determines list rankings and, consequently, which candidates are successful in winning a seat. Party elites do come up with an *ex ante* ordering of candidates appearing on the ballot. They also award a ‘pre-advantage’ status to a limited number of candidates that allocates them additional preference votes compared to non-advantaged candidates.¹²

Nonetheless, it is ultimately voters’ allocation of preference votes that decides on the *ex post*

¹¹ I do leverage variation in Norway’s electoral rules across tiers of government later in the paper. However, the results from higher-order elections are noisy, as difference-in-discontinuity designs require a lot of data for sufficient power.

¹² Each candidate with pre-advantage is awarded a bonus number of preference votes equal to 25% of all votes cast for the party list.

rank ordering.¹³

Apart from the different electoral system, the two cases share many similarities. Both countries feature indirectly elected municipal executives at the local level and hold elections in a regular, fixed four-year cycle. Both countries also rank highly in terms of their overall share of women’s representation in the national legislature, as well as in lower-ranked levels of government.¹⁴ They also rank below average in terms of societal biases against women: The UN Gender Social Norms Index reports that about 20% of respondents in Norway showed a negative bias towards women in politics; the proportion was close to 30% in Spain. Despite great care to account for differences between the two cases, and additional analyses consistent with the argument, it is important to stress that the cross-country comparison nonetheless remains suggestive: we still require more work to credibly estimate the true effect of different shades of PR on the gender gap in political careers.

I collect my data for this analysis from two main sources. Data on elections and candidates in Norway come from [Fiva, Sorensen, and Vollo \(2020\)](#), who helpfully already code unique candidate identifiers across multiple election cycles. I drop all candidates that the original authors flagged as inconsistent, as well as those from minor or non-partisan lists.¹⁵ Data on elections and candidates in Spain come from the country’s Ministry of Interior (<http://www.infoelectoral.mir.es>). To link candidate records across time, I converted

¹³ Voters’ preference votes have proved decisive in electing about 25% of candidates – see footnote 3.

¹⁴ In the analysis sample of bare losers and bare winners, 38% of candidates in Norwegian local elections and 35% in Spanish local elections are women.

¹⁵ The data source flags around 30% of observations in the threshold sample as belonging to municipalities where non-threshold candidates’ preference vote records may be missing (usually from minor parties). However, all of the candidates in the threshold sample itself feature fully recorded preference votes, list rank and election outcomes. Including flagged municipalities is therefore unlikely to affect data quality in the threshold sample itself. Appendix [E.1](#) reports results from a more restricted sample and finds no meaningful difference in the results.

candidate names to lower case, substituted common abbreviations and linked records if the Jaro-Winkler distance between two full names in the same municipality was less than 0.1.¹⁶ Because the original data do not report candidates' gender before 2007, I classified their gender in 2003 based on the probability that the same first name was assigned as either male or female in subsequent elections.¹⁷

In both cases, I drop observations from cities with a population above 250,000 to ensure that locations with unusually high district magnitude and, in some cases, additional institutional powers, do not drive my results.¹⁸ Moreover, in order to maintain the as-if random assumption around the threshold, I follow Fiva and Røhr (2018) and restrict my sample to candidates whose list rank meant that they were the last one within their party to win a seat or the first ones to be defeated.

2.2 Empirical Strategy

2.2.1 Defining the Gender Gap in the Incumbency Advantage

I first discuss the estimand of interest before introducing the estimator and empirical design. Put simply, in each country, I am interested in the gender gap (the difference between men and women) in the effect of being elected (the 'incumbency effect') on downstream outcomes. This *difference* between two causal estimates, which I call the gender gap in the incumbency advantage, is formally defined as follows (using potential outcomes notation):

¹⁶ The Jaro-Winkler distance measures the edit distance between two strings: a lower score implies fewer edits are necessary to move from one string to another.

¹⁷ I fit a linear regression model predicting gender on post-2003 observations' first name, and use the fitted model to predict first names' gender in 2003. I drop observations whose first name only appears in 2003.

¹⁸ Municipalities such as Madrid or Barcelona, for example, feature exceptionally high-profile local politics, with larger budgets and greater policy-making powers.

$$\tau_F - \tau_M = (E[Y_{1i}|s_i = F] - E[Y_{0i}|s_i = F]) - (E[Y_{1i}|s_i = M] - E[Y_{0i}|s_i = M]) \quad (1)$$

where Y denotes the (potential) outcome of interest, and s_i is the candidate’s gender. Note that this estimand focuses on whether women who *do* get elected experience a ‘leaky pipeline’ in their careers, rather than the difficulty of all female candidates getting elected in the first place. This set-up is closely related to estimating the moderation effect of gender on the incumbency effect, although the difference in heterogeneous effects subsumes both the causal effect of gender and that of correlated attributes (Bansak, 2021).¹⁹

My key downstream outcomes are whether the candidate *runs again* in the next election, and whether they *win* in the next election. Following De Magalhaes (2015), I am interested in the unconditional effect of winning in t on election in $t + 1$ – that is, if a candidate does not run again in $t + 1$, they are coded as a ‘0’ on re-election.²⁰ I focus on the next election outcome as my primary interest, because it is the immediate next step in politicians’ careers. If a gender gap already exists at this stage, it will proliferate into downstream outcomes such as running (or winning) in $t + 2$ or running (or winning) in higher-order elections. Estimating the gender gap at the first stepping stone therefore provides the most direct measurement of the level of gender inequity.

¹⁹ Throughout this paper, I estimate the incumbency effect with respect to gender as a ‘bundle’ of characteristics (cf. Sen and Wasow, 2016). This is different from estimating the effect of a female candidate counterfactually running as a man (but with all other characteristics kept constant otherwise (Marshall, 2021)).

²⁰ Conditioning on running again in the future runs the risk of introducing post-treatment bias. See also Hyttinen et al. (2018); Cirone, Cox, and Fiva (2021).

2.2.2 Estimating the Gender Gap Using A Difference-in-Discontinuity Design

How can we estimate the theoretical quantity of interest? A typical regression discontinuity (RD) design, widely applied in the study of incumbency effects (e.g. [Lee, 2001](#); [De Magalhaes, 2015](#); [Folke, Persson, and Rickne, 2016](#)), compares candidates who have just missed out on being elected by a few votes to those who have just crossed the threshold and found themselves elected. The design recovers the local average treatment effect of being elected. It retains its causal interpretation when restricted to a comparison within a subgroup, in which case the RD design recovers the subgroup-specific local average treatment effect.

While the incumbency effect *within* each gender can be causally identified using a regression discontinuity between bare winners and bare losers, the *difference* between the two conditional effects is not causally identified ([Wasserman, 2020](#); [Brown et al., 2019](#)). Male and female candidates around the election threshold may differ from one another on a number of observable as well as unobservable characteristics. In that sense, the main objective of the paper, consistent with the previously defined estimand, is to estimate whether there is heterogeneity in the magnitude of the incumbency effect based on candidates' gender (and the bundle of differences associated with it).

To study the heterogeneity in the incumbency effect between male and female candidates, I estimate the following difference-in-discontinuity specification:

$$\begin{aligned} y_{it} = & f(MV_{it}) + \beta_1 D_{it} + \beta_2 F_{it} + f(MV_{it}, D_{it}) + f(MV_{it}, F_{it}) + \\ & \beta_3(D_{it} \times F_{it}) + f(D_{it}, F_{it}, MV_{it}) + \theta_i + \phi_{it} + \epsilon_{it} \end{aligned} \tag{2}$$

where MV denotes the margin of victory (the running variable), D is a dummy for whether the candidate is elected or not, and F is a dummy for whether the candidate is female. i is a subscript for the individual candidate running in an election at time t . In this setup, $\hat{\beta}_1$ recovers the incumbency effect for males, and $\hat{\beta}_3$ recovers the gender gap in the incumbency effect. I also include county or province (θ_i) and year-by-party fixed effects (ϕ_{it}). Unless noted otherwise, I report robust standard errors clustered by municipality.

Throughout the paper, I present estimates using local linear regression with first-order polynomials (cf. [Gelman and Imbens, 2019](#)) and a uniform kernel.²¹ Because the estimation of the treatment effect in either group depends on the functional form of the conditional expectation function modelling the outcome at the threshold, my estimates might be biased if my functional form $f(MV)$ is misspecified. I therefore also report results using a second-order polynomial specification in [Appendix B.1](#). Finally, the estimates from regression discontinuity may be sensitive to the chosen bandwidth around the threshold. To compute the optimal bandwidth on which I fit my specification, I use Calonico, Cattaneo and Titiunik’s ([2014](#)) approach when run on the full sample (including both men and women).²² Again, to show robustness across modelling choices, I also present estimates when the bandwidth is set to one-half and twice the optimal width in the main tables, and show robustness to further bandwidth choices in [Appendix B.2](#).

As with every RD-based design, my identification assumptions may be confounded if candidates have the ability to sort themselves around the threshold, or if parties can anticipate the election result with high accuracy, thus successfully predicting which list ranks will be

²¹ My results remain robust to dropping fixed effects or using other common (e.g., triangular) kernels (cf. [Appendices B.3 and B.4](#)).

²² I follow [Cipullo \(2021\)](#) in this approach.

elected. To allay that concern, I perform McCrary density tests on the observations around the threshold (McCrary, 2008). Appendix C.1 reports the results of these tests and demonstrates that the density of the running variable around the threshold is indeed smooth. I also check for the continuity of the share of female candidates around the threshold. While we may expect sorting of male and female candidates in the election at-large (e.g., women might run in less successful list positions overall), if treatment assignment for candidates in the neighborhood of the threshold is as good as random, we should also observe no jump in the share of female candidates at the threshold. Appendix C.2 confirms that this is, indeed, the case.

2.2.3 Identifying Bare Winners And Bare Losers In Proportional Representation

In the context of PR elections, constructing the running variable and identifying which candidates came close to barely losing or barely winning is not as straightforward as in the archetypal plurality case (Lee, 2001, 2008; Fiva and Halse, 2016; Fiva, Folke, and Sørensen, 2018). To do so, I use and extend methods developed and applied by Folke (2014) and Fiva and Røhr (2018) in the context of closed-list and open-list PR elections, respectively. I describe both of these methods in greater detail in Appendix D.

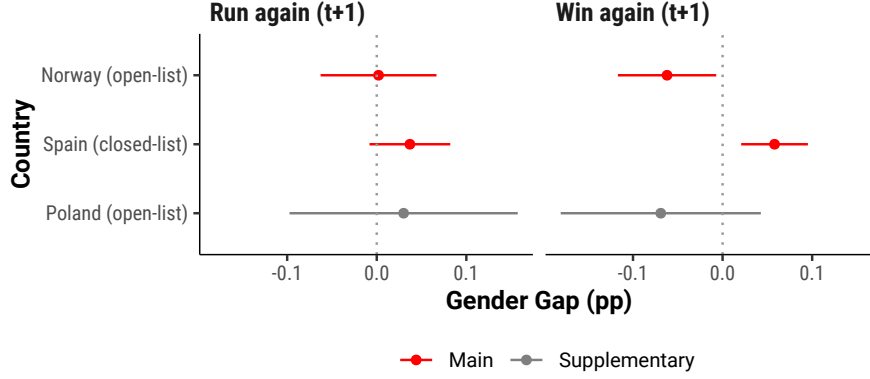


Figure 1: Estimated Gender Gap in Incumbency Advantages Across Cases. Point estimates from preferred difference-in-discontinuity specifications, with 95% Confidence Intervals.

3 Main Results: Negative Gender Gap In Open-List, But Not Closed-List PR

Figure 1 plots the difference in estimated incumbency effects between men and women for local elections in Norway (open-list) and Spain (closed-list), along with noisier, supplementary, estimates from Polish county-level elections (open-list). There is no statistically significant gender gap in the effect of winning on running again in any country. There is, however, a pronounced contrast in the effect on winning again: in Norway, women suffer from a far smaller incumbency advantage than their male colleagues (by about 7 pp), whereas female candidates in Spain experience a differentially larger incumbency advantage than men (by about 6 pp). Both estimates are statistically significant at conventional levels. The gender gap under open-list PR likely replicates to similar settings outside Norway, as the results for Poland suggest.

Figure 2 presents these findings in a graphical representation of the difference in discontinuities. Panel (a) shows the data from Norwegian open-list elections. Women, while

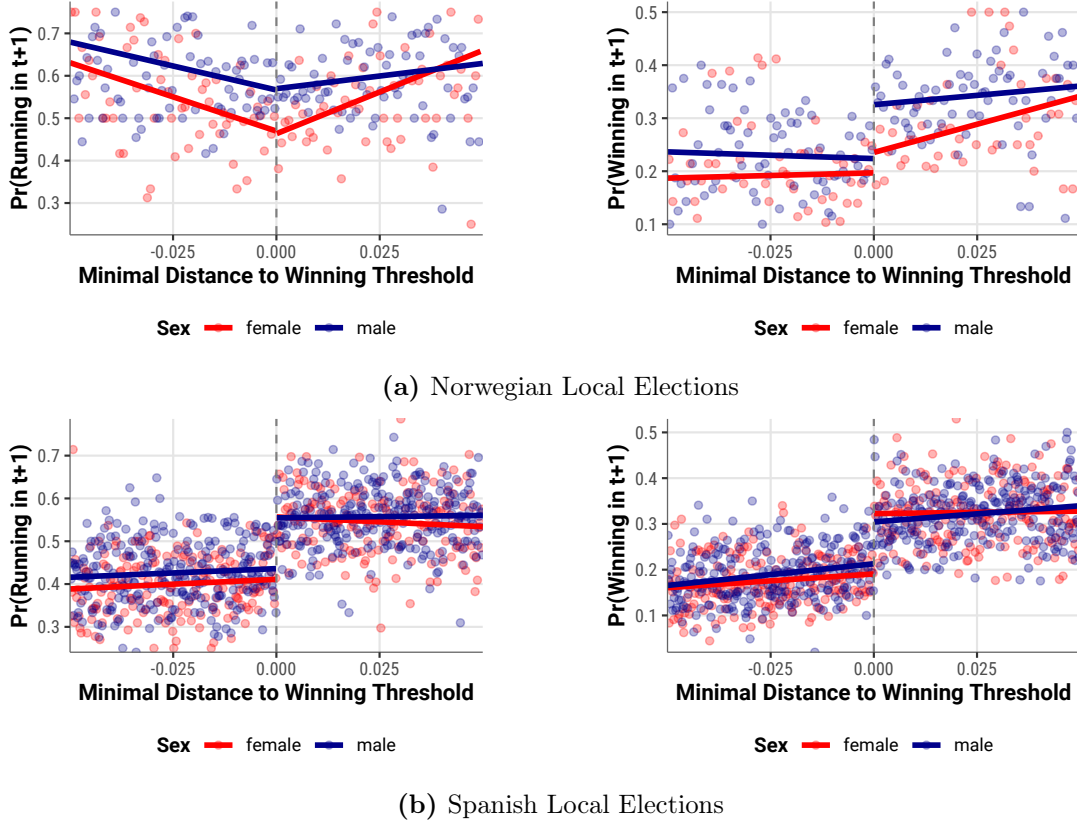


Figure 2: Effect of Winning Election on Running Again and Winning Again, By Gender And Case. The visual regression discontinuities compare bare winners and bare losers by candidates' gender.

unconditionally less likely to run again (left panel), do not suffer from any additional disadvantage in *running again* compared to men once elected: losing an election does not diminish women's persistence differentially. While barely winning increases the probability of *winning again* for candidates of either gender (right panel), the effect is clearly greater for male candidates (4 percentage points for women vs. 10 percentage points for men). This matches the results from Figure 1. Turning to Spanish closed-list elections in (b), we similarly observe no gender difference in the effect of winning on running again; candidates of either gender experience an increase of approximately 9 percentage points in the probability of winning again as a result of getting elected. Unlike in the Norwegian case, the Figure does not show

a visually obvious gender gap; if anything, the effect appears somewhat larger for women.

In the remainder of this section, I parse these headline results in greater detail and provide additional results to guard against likely alternative explanations.

3.1 Norway (Open-List PR)

I first present detailed results from the open-list setting in Norway’s municipal elections.

Table 2 backs up the graphical intuition with formal evidence in the form of difference-in-discontinuity estimates. Columns 1 to 3 report estimates of the effect of winning on running again in $t + 1$. Both the effect of being elected, as well as the gender gap coefficient, are close to zero across all bandwidth choices.

Columns 3 to 6 report the effect of winning on the probability of winning again in $t + 1$. For men, there is an increase of about 10 pp., which is statistically significant at all conventional levels. The interaction coefficient suggests that the incumbency advantage for women is approximately 6 pp. lower, which corresponds to a 60% decrease in the magnitude of the effect. Although the precise magnitude of the difference between men and women varies between 5 and 10 pp. depending on the bandwidth, all interaction estimates are statistically significant.²³

Together, these results constitute first evidence that female candidates in an open-list setting may suffer from a diminished incumbency advantage in future re-election attempts.

²³ For further robustness checks, Appendix B.2 reports the coefficients for a wide range of bandwidths. The results are also robust to dropping all municipalities in which some personal votes for candidates outside the threshold sample are recorded as missing (Appendix E.1), as well as to dropping outlying cities with very high populations (Appendix E.8.1).

Table 2: Difference-in-Discontinuity Estimates For Incumbency Advantage In Norwegian Municipalities. Women face a diminished incumbency effect on winning again.

	Run ($t + 1$)			Win ($t + 1$)		
	(1)	(2)	(3)	(4)	(5)	(6)
Elected	0.002 (0.019)	-0.010 (0.016)	0.016 (0.024)	0.107 (0.017)	0.102 (0.015)	0.109 (0.021)
Female	-0.097 (0.021)	-0.092 (0.018)	-0.085 (0.026)	-0.027 (0.018)	-0.024 (0.015)	-0.026 (0.021)
Elected x Female	0.002 (0.033)	0.024 (0.028)	-0.030 (0.039)	-0.062 (0.028)	-0.047 (0.024)	-0.096 (0.034)
Bandwidth	0.054	0.11	0.027	0.05	0.099	0.025
BW Type	Optimal	2x Opt	0.5x Opt	Optimal	2x Opt	0.5x Opt
Outcome Mean	0.564	0.575	0.551	0.261	0.263	0.257
N (left)	4617	5529	3667	4502	5428	3549
N (right)	4666	5590	3711	4551	5489	3592

All estimates are reported with robust standard errors clustered at the municipality level in parentheses. Each observation is a candidate's election attempt. 'Elected' is an indicator for observations where the candidate obtained a seat in the municipal council. 'Female' is an indicator for observations identified as female. 'Elected' times 'Female' is the interaction between the two variables. Other coefficients not reported. Regression run on all candidates in elections between 2003 and 2015.

3.2 Spain (Closed-List PR)

I now turn to the results from closed-list elections in Spanish municipalities.

The formal estimates in Table 3 corroborate the visual inspection and offer more precise results. Once again, columns 1 to 3 report the effect of winning on running again in the next election. There is a large and significant effect of winning on running again for men – an approximately 10 percentage point increase, in line with the result from Figure 2b. In this case, the gender gap is close to zero and statistically insignificant: the effect of winning an election on running again appears to be of a very similar (or slightly larger in the case of some bandwidths) magnitude for women.

Columns 4 to 6 provide clear evidence that the incumbency effect on winning again is at least as great, if not bigger, for women. Men experience an incumbency effect of an

Table 3: Difference-in-Discontinuity Estimates For Incumbency Advantage In Spanish Municipalities. Women likely enjoy a larger effect of winning on their probability to win again.

	Run ($t + 1$)			Win ($t + 1$)		
	(1)	(2)	(3)	(4)	(5)	(6)
Elected	0.109 (0.014)	0.119 (0.010)	0.100 (0.020)	0.072 (0.012)	0.095 (0.008)	0.076 (0.017)
Female	-0.024 (0.016)	-0.032 (0.012)	-0.012 (0.022)	-0.032 (0.013)	-0.026 (0.009)	-0.033 (0.018)
Elected x Female	0.037 (0.023)	0.033 (0.017)	0.018 (0.032)	0.058 (0.019)	0.045 (0.013)	0.057 (0.027)
Bandwidth	0.032	0.064	0.016	0.035	0.071	0.018
BW Type	Optimal	2x Opt	0.5x Opt	Optimal	2x Opt	0.5x Opt
Outcome Mean	0.486	0.483	0.485	0.256	0.251	0.254
N (left)	14840	26294	7984	16223	28426	8726
N (right)	14729	26886	7825	16159	29072	8530

All estimates are reported with robust standard errors clustered at the municipality level in parentheses. See Table 2 for additional details.

approximately 7 percentage point increase in their probability of winning the next election.²⁴

The gender gap coefficient is positive, and statistically significant at the optimal bandwidth (though not at other values), suggesting that the incumbency advantage is an additional 5 to 6 percentage points larger for women. This set of results suggests that women in Spanish local elections enjoy a somewhat *greater* incumbency advantage than their male colleagues – much in contrast to the findings in Norway’s open-list PR setting.

I conduct a number of additional analyses to confirm that my Spanish results from estimating the gender gap on the whole sample are unlikely to mask any underlying heterogeneity between parties or municipalities. In Appendix F.2, I check for any meaningful differences if I estimate the difference-in-discontinuity separately for either of Spain’s two major parties. Second, in Appendix F.3, I evaluate whether the results are meaningfully different for muni-

²⁴ See Appendix B.1 for robustness to polynomial order, and Appendix B.2 for robustness to bandwidth choice. I also leverage the large number of observations in order to estimate a diff-in-diff sufficiently close to the threshold (within 1 percentage point) in Appendix F.1. The estimates are consistent with the results from Table 3, and further point towards no large gender gap in the effect of winning on running again.

cipalities above that implemented a legally binding gender quota – mandating that at least 40% of all candidates be women – versus those that did not. I do not find evidence of a negative gender gap for women across either type of municipality. In Appendix F.4, I also find no meaningful evidence that estimates differ by cities’ population size, although the positive gender gap may be attenuated towards zero in cities in the highest population tercile.

All told, the direct comparison between the results from Norway’s open-list and Spain’s closed-list elections is consistent with the theoretical argument that women suffer from disadvantages in incumbency effects where voters can determine individual candidates’ list rankings. Still, on its own, this comparison cannot rule out that fundamental differences between Norway and Spain drive the set of results. Next, I leverage additional within-country and cross-country analyses to address this concern.

3.3 Countries’ Idiosyncracies Unlikely To Drive Gender Gap Difference

Are the divergent estimates of the gender gap the result of countries’ idiosyncratic characteristics, rather than the product of electoral rules? To address this concern, I first exploit within-country variation in Norway’s electoral systems, and next extend my analysis of open-list systems to another, though noisier, case: open-list elections in Poland. Jointly, these supplementary analyses address concerns about the cross-country comparison and strengthen the argument for the role of the electoral system vis-a-vis country-specific differences in explaining the gender gap.

3.3.1 Same Country, Different Electoral System: Likely No Gender Gap

Is the gender gap in incumbency advantages restricted to open-list PR, or is it a more general feature of Norwegian elections? I examine whether elections at the *county* and *national* level, where voters have far less control over individual list rankings, exhibit a similar gender gap as in municipal ones.²⁵ As the higher-level electoral setting (with fewer candidates and fewer marginal seats) leaves me underpowered for a difference-in-discontinuity design, I estimate a difference-in-difference specification among borderline elected and borderline losing candidates for all national and county elections between 2000 and 2012 using data from Fiva and Smith (2018).²⁶

Table 4 reports the estimates from the difference-in-differences specification fitted on national elections. Columns (1) and (2) leverage all borderline elected or losing candidates. The gender gap coefficient is positive and significant. Columns (3) and (4) use a bandwidth of 10 percentage points (so borderline candidates who win/lose by a large margin, and likely feature different characteristics are no longer included) with virtually unchanged point estimates. Finally, columns (5) and (6) use a bandwidth of 5 percentage points. The gender gap shrinks close to zero. Although it does not turn negative, it comes with large standard errors, rendering this bandwidth estimate difficult to interpret.

In addition, Appendix E.2 reports additional results for county-level elections. I observe the same pattern as in national elections, even though the estimates come with even greater uncertainty.

²⁵ It is possible that the gender gap may be attenuated at higher levels of office if only the highest-ability female candidates can overcome repeated gender gaps at lower, open-list rungs of the office pipeline, and party elites select with less bias in higher-level closed-list PR elections.

²⁶ These estimates may be biased, as candidates away from the threshold may no longer be quasi-randomly assigned to winning or losing their election bid. Still, Fiva and Halse (2016) employ the same strategy.

Table 4: Probabilities of Winning Re-Election in Norway’s National Elections (Difference-in-Differences).

	Win (in t+1)					
	(1)	(2)	(3)	(4)	(5)	(6)
Elected	0.290 (0.051)	0.282 (0.046)	0.261 (0.053)	0.257 (0.045)	0.236 (0.073)	0.252 (0.068)
Female	-0.064 (0.035)	-0.052 (0.031)	-0.047 (0.043)	-0.030 (0.040)	-0.069 (0.061)	-0.013 (0.068)
Elected x Female	0.101 (0.058)	0.111 (0.049)	0.098 (0.069)	0.102 (0.062)	0.028 (0.090)	0.004 (0.087)
R2	0.143	0.193	0.130	0.199	0.116	0.232
Bandwidth	All	All	0.10	0.10	0.05	0.05
N	648	648	566	566	355	355
Outcome Mean	0.312	0.312	0.309	0.309	0.349	0.349
Region FE	Y	Y	Y	Y	Y	Y
Year FE	Y	N	Y	N	Y	N
Year x Party FE	N	Y	N	Y	N	Y

All estimates are reported with robust standard errors clustered at the election district level in parentheses. Each observation is a candidate’s election attempt. Regression run on all candidates in national elections between 2003 and 2009.

While these results are not well identified, and are estimated noisily, they remain consistent with the argument that the features of open-list PR, rather than a unique characteristic of Norwegian politics, drive the negative gender gap found at the local level.

3.3.2 Different Country, Same Electoral System: Gender Gap Replicates

Alternatively, is the gender gap in Norway’s open-list elections idiosyncratic to only this set of elections? If the claim about the effect of electoral systems holds, the negative gender gap ought to replicate in other countries using a similar electoral system. I implement this test by fitting the main difference-in-discontinuity specification on county-level elections from Poland, which also use open-list PR.²⁷

²⁷ I restrict my sample to candidates from the two main parties – Civic Platform (PO), and Law and Order (PiS) – running in the 2010 and 2014 elections at the county level. I leave out the 2018 elections for two reasons, both of which are likely to affect candidates’ career trajectories. First, an electoral reform extended office tenure length from four to five years starting in 2018. Second, in that year, Civic Platform merged with several smaller parties to form a broad opposition alliance.

The estimates of the gender gap in Polish county-level, open-list elections, albeit less precise, match those from Norwegian municipalities. I report the full set of estimates in Appendix G. The results suggest that male candidates enjoy an approximately 10 percentage point higher probability of winning in the next election as a result of getting elected. Women, however, experience an incumbency advantage that is about 6 percentage points smaller. Although, with a much smaller sample, the point estimates are not statistically significant at conventional levels, they are consistent across multiple bandwidth specifications and of a very similar magnitude as those in Norway.

Overall, the finding of a gender gap in incumbency advantages replicates for the same electoral system across different countries, but not for different electoral systems within the same country. This weakens (but cannot fully rule out) concerns about case selection driving the results and is consistent with the claim that differences in the type of proportional representation may affect the gender gap in political careers.

3.4 Differential Selection Of Candidates Into Sample Unlikely To Drive Gender Gap

In addition to concerns about the cross-country comparison discussed just above, we may also worry about threats to inference within any one case: if candidates' gender is also correlated with other characteristics, Norway's open-list results may simply reflect parties' or voters' preferences for the associated attribute, rather than gender itself (Bansak, 2021; Teele, Kalla, and Rosenbluth, 2018). Alternatively, female candidates in the threshold sample may run in different electoral environments (e.g. parties or geographies) that lead to smaller incumbency

advantages [Folke and Rickne \(2016\)](#). In both of these cases, differential selection into the sample could explain my results and threaten the inference drawn from the observed gender gap.

I present two checks to caution against this explanation. First, I substitute gender with age and experience as conditioning variables in the difference-in-discontinuities setup in the Norwegian case.²⁸ My results in [Appendix E.6](#) suggest that there is no difference in the magnitude of the incumbency advantage between older and younger, or more and less experienced candidates.

Second, using a matching approach discussed in [Bansak \(2021\)](#) and [Bansak and Nowacki \(2022\)](#), I compare female candidates to male candidates who are most similar on a number of individual- and election-level attributes. In [Appendix E.7](#), I present results from various matched comparisons that remain consistent with my main results. Matching on candidate attributes alone, the gender gap remains large (up to 7 pp.) and statistically highly significant. In the most demanding matching procedure – finding the most similar candidate by age, experience and election year in exactly the same party *and* county – women’s incumbency advantage is still 4 to 5 percentage points smaller (a 50% decrease compared to men), although this estimate is noisier due to a smaller sample size ($p = 0.13$). Overall, the finding of a negative gender gap in Norwegian local politics persists even when a stricter comparison group lessens concerns about differential selection.

Jointly, the results in this subsection offer suggestive evidence that gender itself, rather than a correlated attribute, is a major factor driving the observed gender gap in incumbency

²⁸ Unfortunately, this is not possible to replicate in the Spanish case due to a lack of available candidate covariates.

advantages – even though caveats about the finite nature of available covariates remain.

4 Mechanisms Behind the Gender Gap in Open-List PR

What mechanism accounts for the stark difference in gender-specific incumbency advantages under open-list PR in Norway? In this section, I show that the gender gap is most pronounced among candidates in right-wing parties. I also find that voters in right-wing parties award a smaller increase in preference votes to female candidates as a result of being barely elected. This channel, mechanically unavailable in closed-list PR, is likely to contribute to women’s smaller incumbency advantage in open-list settings.

4.1 Women’s Disadvantage Concentrated in Right-Wing Parties

Table 5: Difference-in-Discontinuity Estimates For Incumbency Advantage In Norwegian Municipalities, By Political Party Group.

	Run (t+1)		Win (t+1)	
	(1)	(2)	(3)	(4)
Elected	0.022 (0.031)	0.005 (0.026)	0.086 (0.027)	0.126 (0.024)
Female	−0.079 (0.034)	−0.044 (0.030)	−0.015 (0.027)	−0.001 (0.025)
Elected x Female	−0.040 (0.049)	−0.014 (0.045)	−0.053 (0.039)	−0.089 (0.039)
Parties	Left	Right	Left	Right
Bandwidth	0.05	0.068	0.052	0.069
Outcome Mean	0.546	0.59	0.253	0.26
N (left)	1568	2382	1580	2388
N (right)	1589	2413	1601	2419

All estimates are reported with robust standard errors clustered at the municipality level in parentheses. Each observation is a candidate’s election attempt.

Is the gender gap in the incumbency advantage persistent across parties from the entire

spectrum of political ideologies? Both parties and voters following a more socially conservative ideology might be more inclined to exhibit a negative bias towards women, thus diminishing their incumbency advantage.²⁹ Although merely suggestive, a stark contrast between party families may also point towards voter bias, rather than universal, systemic bias, as a key mechanism.

I estimate the difference-in-discontinuity separately on subsamples of candidates from left-leaning and right-leaning parties and report the results in Table 5.³⁰ Columns 1 and 2 report the effect of winning on running again by subsample. Although noisier than the main results, we see the same finding replicated as in Table 2. There is no meaningful effect of being elected on running again in the next election for male candidates; nor is there a statistically significant difference in the effect between genders.

Columns 3 and 4 report the effect of winning on whether the candidate is elected in the next election. In left-wing parties, both the absolute incumbency effect for men, as well as the magnitude of the gender gap are smaller. Though the point estimate still suggests a 5 pp. lower incumbency advantage for women, the gender gap coefficient is no longer statistically significant in this subsample. By contrast, the gender gap is more pronounced in right-wing parties. Here, male candidates who are bare winners enjoy an almost 13 percentage points higher chance of being re-elected (compared to male bare losers). For female candidates, however, this incumbency advantage is estimated to be almost 9 percentage points smaller,

²⁹ Appendix F.2 repeats the same procedure for the Spanish case. Though noisy, the results suggest that women likely enjoy similar or even greater incumbency advantages as men across the ideological spectrum.

³⁰ Parties classified as 'left' are Labour and the Socialist Left. Parties classified as 'right' are the Conservatives, Progress, the Christian Democratic Party, and the Liberal Party, corresponding to the two 'coalition blocs' in Norwegian politics. I exclude the Centre Party, which has joined both left-wing and right-wing coalitions in recent decades.

a statistically significant difference.

4.2 Voters, Rather Than Party, May Drive Gender Gap

Do party elites or voters drive the diminished incumbency advantage for women? Exploiting the specific rules of Norway’s open-list elections, I fit the difference-in-discontinuity design on additional outcomes pertaining to the next election and provide suggestive evidence that right-wing voters’ behavior contributes to the diminished incumbency advantage for women.

Recall that, party elites draw up an *ex ante* ranking of candidates on the list, and award a *pre-advantage* to select candidates (who receive a boost in preference vote shares as a result). However, if party elites are less likely to award these improvements to incumbent female candidates, this could explain the diminished incumbency advantage even though voters are no less likely to vote for women. Alternatively, if party elites exhibit no bias in the *ex ante* placement of women, but there is a gender gap in the increase in preference votes awarded by voters upon winning the first time, then voters’ male-favoring voting behavior may contribute to the diminished incumbency advantage for women.

Table 6 reports the results. Because previous results suggest that the gender gap is most pronounced in right-wing parties, I fit the specification separately on left-wing and right-wing candidates.³¹

In columns 1 and 2, I investigate whether winning an election renders candidates more likely to advance in *ex ante* list rank in the next election. Across both left- and right-wing parties, male winners enjoy a significant increase in their probability of improving their *ex ante* rank as a result of winning. I find no statistically significant gender gap, although the

³¹ The results are substantively similar when fitting second-order polynomials (Appendix E.3).

Table 6: Mechanisms leading to lower incumbency advantage: Difference-in-Discontinuity Estimates On Additional Outcomes, Norway.

	Orig. Rank Advance		Actual Rank Advance		Pre-Ad. (t+1)		Pers.V. Share (t+1)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Elected	0.088 (0.026)	0.066 (0.020)	0.026 (0.028)	0.044 (0.021)	0.037 (0.019)	0.058 (0.019)	0.003 (0.002)	0.005 (0.001)
Female	0.003 (0.027)	-0.003 (0.023)	-0.031 (0.029)	0.011 (0.023)	0.017 (0.022)	0.049 (0.021)	-0.003 (0.003)	0.000 (0.002)
Elected x Female	-0.063 (0.038)	-0.034 (0.034)	-0.013 (0.039)	-0.062 (0.032)	0.001 (0.031)	-0.041 (0.033)	0.001 (0.004)	-0.005 (0.003)
Parties	Left	Right	Left	Right	Left	Right	Left	Right
Bandwidth	0.068	0.092	0.061	0.097	0.033	0.055	0.025	0.052
Outcome Mean	0.257	0.253	0.229	0.211	0.119	0.163	0.0204	0.0175
N (left)	1655	2620	1624	2662	1421	2203	446	840
N (right)	1676	2655	1645	2698	1441	2229	458	856

All estimates are reported with robust standard errors clustered at the municipality level in parentheses. Each observation is a candidate's election attempt.

point estimates are still signed negative. In columns 3 and 4, examine whether candidates advanced in *ex post* ranks – after voters cast their preferences. In left-leaning parties, neither men nor women enjoy a statistically significant effect. By contrast, in right-wing parties, male winners are 4 percentage points more likely to improve their ex-post rank in the next election. The gender gap, close to significant at conventional levels, suggests that this effect is 6 percentage points *lower* for women. Contrasting this with the previous columns, it suggests that most of the diminished incumbency advantage that women face in right-wing parties comes from voters.

Another way of examining whether party elites are responsible for holding back female winners is to look at awarding pre-advantage status. Columns 5 and 6 suggest that men enjoy an increased probability (4 to 6 pp.) of receiving pre-advantage status when elected. In neither case, however, do we find a statistically significant gender gap in the effect of being elected on attaining pre-advantage.

Finally, I turn to the effect of winning on the share of preference votes received in the next election.³² For these specifications, I drop municipalities flagged as having incomplete personal vote records and condition the sample on running again in the first place.³³ On average, male candidates, upon being elected, can expect a 0.3 percentage point (0.5 in right-wing parties) higher preference vote share in the next election. While the effect is small in absolute size, it is meaningful in relation to candidates' average preference vote share of 2 (1.75) percentage points, representing a 15% (28%) increase for the average candidate in the sample. We see a stark difference in the estimated gender gap between the two party groups. In left-wing parties, the coefficient is close to zero, whereas in right-wing parties, it is negative, statistically significant, and drives the incumbency advantage in preference votes for women back to zero.

Overall, these results, though with the caveat of large uncertainty, point towards right-wing voters rewarding female candidates for being elected with a smaller increase in personal votes as an important factor in explaining the gender gap. An increase in personal votes as a meaningful incumbency advantage mechanism may thus not be available to all candidates to the same extent (Fiva and Røhr, 2018). Although I cannot precisely estimate whether women also face diminished effects for outcomes solely determined by party leadership, my results suggest that, at worst, both voters and party elites – perhaps responding strategically – play a role in decreasing women's incumbency advantage.

³² These estimates report the effect on the share of preference votes *relative to all votes cast in the municipality*. Measuring personal votes is somewhat more challenging than previously discussed binary outcomes. Appendix E.4 shows robustness to different operationalisations of personal vote measures.

³³ Unlike previous binary outcomes, it is less clear that imputing "0" personal votes for those not running again is justifiable. Following De Magalhaes (2015), conditioning is only problematic if barely elected and barely defeated candidates exhibit different probabilities of running again, which I do not observe in this context. Appendix E.4.2 discusses the sample restriction further and shows that results are consistent (though noisier) when including all observations.

4.3 Further Evidence for the Mechanism

Finally, I briefly report two additional analyses that contribute further suggestive evidence in support of the role of the theorized explanation for the gender gap under open-list PR.

Gender Gap By Population Size. If voter-driven incumbency advantages predominantly work through name recognition heuristics, the gender gap should attain its largest magnitude in big cities with a higher number of councillors. In small towns, tracking information about a handful of candidates may be less costly; in larger cities, however, voters are less likely to have information about an increasing number of candidates and may rely more strongly on heuristics such as name recognition and negative biases towards women. Appendix [E.8](#) reports the gender gap estimates for subsamples of the data binned into population terciles and provides evidence consistent with this logic.

Survey Results. If right-wing voters drive the gender gap in incumbency advantages, we should expect this group of voters to display more negative views towards women in politics in general. To test this claim, I use data from the 2015 Norwegian Local Election Survey. Results suggest that voters in right-wing parties were, indeed, less supportive of women’s advancement in local politics. I report and discuss the results in greater detail in Appendix [E.9](#).

5 Conclusion

Do different forms of proportional representation feature gender gaps in political careers? This paper is among the first to evaluate gender gaps in incumbency advantages under PR,

and the first to theorize about and empirically document a stark contrast in gender gaps between elections held under open-list and closed-list variants of PR. Under closed-list PR, voters are unable to determine individual candidates' list positions: consequently, female candidates are likely shielded from voters' biases and experience incumbency effects that are equal to or greater than men's. Meanwhile, under open-list PR, female candidates may suffer from diminished incumbency advantages as incumbency effect mechanisms operate through voter-based name recognition channels, which may be more vulnerable to gender stereotyping.

Jointly, the results in this paper point towards the role that electoral rules play not only for the *overall* share of women elected but also whether female candidates, once elected, enjoy equitable political careers. This finding may help explain why, despite an increase in the nominal share of women, downstream outcomes such as policy might not change (Ferreira and Gyourko, 2014; Bagues and Campa, 2021).

The paper's findings come with caveats that also point towards promising avenues for future research. First, although the paper provides evidence for the existing contrast between open-list and closed-list environments, and suggests that voters' ability to selectively cast preference votes in open-list systems likely contribute to the gender gap in this setting, more work is needed to provide more direct and more precise evidence for theorized mechanisms. Subsequent work should further evaluate which of the channels leading to smaller preference vote boosts for incumbent women is mainly responsible for the observed result. Second, additional work is also needed to test whether my findings generalize beyond the selected cases, in particular with respect to other variants of open-list (and flexible-list) PR. Lastly, we should also research the gender gap in women's promotion to higher-ranked offices, thus

studying whether the ‘leaky pipeline’ is indeed responsible for women’s underrepresentation in higher levels of politics, and compare the results with similar studies under plurality (Wasserman, 2020; Brown et al., 2019; Cipullo, 2021).

It is also important to emphasize that, even beyond concerns about limiting voter’s influence over who gets elected, closed-list PR is far from a panacea for equal representation: although female candidates enjoy equal or possibly greater incumbency advantages in Spanish municipalities, women are still a far way off from achieving equal representation. More generally, my results do not speak to the problem of ‘sticky floors’ (Cipullo, 2021): women may struggle to get elected in the first place, and parties might still disadvantage women *unconditionally* by placing them in lower or unwinnable ranks (Murray, Krook, and Opello, 2012; Luhiste, 2015; Casas-Arce and Saiz, 2015). Still, this paper demonstrates that parties, policymakers and electoral reformers should also consider the impacts of electoral systems on gender differences throughout candidates’ entire career trajectory.

References

- Anzia, Sarah F., and Rachel Bernhard. 2019. “How Does Gender Stereotyping Affect Women at the Ballot Box? Evidence from Local Elections in California, 1995-2016.”
- Ashworth, Scott, Christopher R Berry, and Ethan Bueno de Mesquita. 2021. “Sources of Women’s Underrepresentation in US Politics: A Model of Election Aversion and Voter Discrimination.”
- Bagues, Manuel, and Pamela Campa. 2021. “Can Gender Quotas in Candidate Lists Empower Women? Evidence from a Regression Discontinuity Design.” *Journal of Public Economics* p. 104315.
- Bansak, Kirk. 2021. “Estimating Causal Moderation Effects with Randomized Treatments and Non-Randomized Moderators.” *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 184(1): 65–86.
- Bansak, Kirk, and Tobias Nowacki. 2022. “Effect Heterogeneity and Causal Attribution in Regression Discontinuity Designs: Introducing the Moderation-in-Discontinuities Framework.”
- Beaman, Lori, Raghabendra Chattopadhyay, Esther Duflo, Rohini Pande, and Petia Topalova. 2009. “Powerful Women: Does Exposure Reduce Bias?” *Quarterly Journal of Economics* 124(4): 1497–1540.
- Bergh, Johannes, Tor Bjørklund, and Ottar Hellevik. 2010. “Personutvelgingen i Norske Valg.” *Norsk statsvitenskapelig tidsskrift* 26(2): 105–131.
- Bernhard, Rachel, and Justin de Benedictis-Kessner. 2021. “Men and Women Candidates Are Similarly Persistent after Losing Elections.” *Proceedings of the National Academy of Sciences* 118(26).
- Blais, André, and Ignacio Lago. 2009. “A General Measure of District Competitiveness.” *Electoral Studies* 28(1): 94–100.
- Broockman, David E., and Evan J. Soltas. 2020. “A Natural Experiment on Discrimination in Elections.” *Journal of Public Economics* 188: 104201.
- Brown, Ryan, Hani Mansour, Stephen D O’Connell, and James Reeves. 2019. “Gender Differences in Political Career Progression: Evidence from U.S. Elections.”
- Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. 2014. “Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs.” *Econometrica* 82(6): 2295–2326.
- Casas-Arce, Pablo, and Albert Saiz. 2015. “Women and Power: Unpopular, Unwilling, or Held Back?” *Journal of Political Economy* 123(3): 641–669.
- Chattopadhyay, Raghabendra, and Esther Duflo. 2004. “Women as Policy Makers: Evidence from a Randomized Policy Experiment in India.” *Econometrica* 72(5): 1409–1443.

- Cipullo, Davide. 2021. "Gender Gaps in Political Careers: Evidence from Competitive Elections."
- Cirone, Alexandra, Gary W. Cox, and Jon H. Fiva. 2021. "Seniority-Based Nominations and Political Careers." *American Political Science Review* 115(1): 234–251.
- Cox, Gary W., Jon H. Fiva, and Daniel M. Smith. 2020. "Measuring the Competitiveness of Elections." *Political Analysis* 28(2): 168–185.
- Dahlerup, Drude. 2013. *Women, Quotas and Politics*. Routledge.
- Dahlgaard, Jens Olav. 2016. "You Just Made It: Individual Incumbency Advantage under Proportional Representation." *Electoral Studies* 44: 319–328.
- De Magalhaes, Leandro. 2015. "Incumbency Effects in a Comparative Perspective: Evidence from Brazilian Mayoral Elections." *Political Analysis* 23(1): 113–126.
- Dettman, Sebastian, Thomas B. Pepinsky, and Jan H. Pierskalla. 2017. "Incumbency Advantage and Candidate Characteristics in Open-List Proportional Representation Systems: Evidence from Indonesia." *Electoral Studies* 48: 111–120.
- Eggers, Andrew C., Anthony Fowler, Jens Hainmueller, Andrew B. Hall, and James M. Snyder Jr. 2015. "On the Validity of the Regression Discontinuity Design for Estimating Electoral Effects: New Evidence from over 40,000 Close Races." *American Journal of Political Science* 59(1): 259–274.
- Eymmoud, Jean-Benoit, and Paul Vertier. 2017. "Gender Biases: Evidence from a Natural Experiment in French Local Elections."
- Ferreira, Fernando, and Joseph Gyourko. 2014. "Does Gender Matter for Political Leadership? The Case of US Mayors." *Journal of Public Economics* 112: 24–39.
- Fiva, Jon H., and Askill H. Halse. 2016. "Local Favoritism in At-Large Proportional Representation Systems." *Journal of Public Economics* 143: 15–26.
- Fiva, Jon H., and Daniel M. Smith. 2018. "Political Dynasties and the Incumbency Advantage in Party-Centered Environments." *American Political Science Review* 112(3): 706–712.
- Fiva, Jon H., and Helene Lie Røhr. 2018. "Climbing the Ranks: Incumbency Effects in Party-List Systems." *European Economic Review* 101: 142–156.
- Fiva, Jon H., Olle Folke, and Rune J. Sørensen. 2018. "The Power of Parties: Evidence from Close Municipal Elections in Norway." *The Scandinavian Journal of Economics* 120(1): 3–30.
- Fiva, Jon H., Rune J. Sørensen, and Reidar Vollo. 2020. "Local Candidate Dataset."
- Folke, Olle. 2014. "Shades of Brown and Green: Party Effects in Proportional Election Systems." *Journal of the European Economic Association* 12(5): 1361–1395.

- Folke, Olle, and Johanna Rickne. 2016. "Electoral Competition and Gender Differences in Political Careers." *Quarterly Journal of Political Science* 11(1): 59–102.
- Folke, Olle, and Johanna Rickne. 2020. "All the Single Ladies: Job Promotions and the Durability of Marriage." *American Economic Journal: Applied Economics* 12(1): 260–87.
- Folke, Olle, Torsten Persson, and Johanna Rickne. 2016. "The Primary Effect: Preference Votes and Political Promotions." *The American Political Science Review* 110(3): 559.
- Gelman, Andrew, and Guido Imbens. 2019. "Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs." *Journal of Business & Economic Statistics* 37(3): 447–456.
- Golder, Sona N., Laura B. Stephenson, Karine Van der Straeten, André Blais, Damien Bol, Philipp Harfst, and Jean-François Laslier. 2017. "Votes for Women: Electoral Systems and Support for Female Candidates." *Politics & Gender* 13(1): 107–131.
- Hazan, Reuven Y., and Gideon Rahat. 2006. "The Influence of Candidate Selection Methods on Legislatures and Legislators: Theoretical Propositions, Methodological Suggestions and Empirical Evidence." *The Journal of Legislative Studies* 12(3-4): 366–385.
- Horiuchi, Yusaku, Daniel M. Smith, and Teppei Yamamoto. 2020. "Identifying Voter Preferences for Politicians' Personal Attributes: A Conjoint Experiment in Japan." *Political Science Research and Methods* 8(1): 75–91.
- Hyttinen, Ari, Jaakko Meriläinen, Tuukka Saarimaa, Otto Toivanen, and Janne Tukiainen. 2018. "When Does Regression Discontinuity Design Work? Evidence from Random Election Outcomes." *Quantitative Economics* 9(2): 1019–1051.
- Jankowski, Michael, and Kamil Marcinkiewicz. 2019. "Ineffective and Counterproductive? The Impact of Gender Quotas in Open-List Proportional Representation Systems." *Politics & Gender* 15(1): 1–33.
- Jankowski, Michael, and Stefan Müller. 2021. "The Incumbency Advantage in Second-Order PR Elections: Evidence from the Irish Context, 1942–2019." *Electoral Studies* 71: 102331.
- Kerevel, Yann. 2019. "Empowering Women? Gender Quotas and Women's Political Careers." *The Journal of Politics* 81(4): 1167–1180.
- Kjaer, Ulrik, and Mona Lena Krook. 2019. "The Blame Game: Analyzing Gender Bias in Danish Local Elections." *Politics, Groups, and Identities* 7(2): 444–455.
- Krook, Mona Lena. 2018. "Electoral Systems and Women's Representation." In *The Oxford Handbook of Electoral Systems*. Oxford University Press pp. 175–192.
- Krook, Mona Lena, and Pippa Norris. 2014. "Beyond Quotas: Strategies to Promote Gender Equality in Elected Office." *Political Studies* 62(1): 2–20.
- Lawless, Jennifer L., and Richard L. Fox. 2005. *It Takes a Candidate: Why Women Don't Run for Office*. Cambridge University Press.

- Lee, David S. 2001. The Electoral Advantage to Incumbency and Voters' Valuation of Politicians' Experience: A Regression Discontinuity Analysis of Elections to the US. Working Paper National Bureau of Economic Research.
- Lee, David S. 2008. "Randomized Experiments from Non-Random Selection in US House Elections." *Journal of Econometrics* 142(2): 675–697.
- Lucardi, Adrián, and Juan Pablo Micozzi. 2020. "District Magnitude and Female Representation: Evidence from Argentina and Latin America." *American Journal of Political Science* (First View).
- Luhiste, Maarja. 2015. "Party Gatekeepers' Support for Viable Female Candidacy in PR-list Systems." *Politics & Gender* 11(1): 89–116.
- Mansbridge, Jane. 1999. "Should Blacks Represent Blacks and Women Represent Women? A Contingent" Yes"." *Journal of Politics* 61(3): 628–657.
- Mansbridge, Jane. 2003. "Rethinking Representation." *American Political Science Review* 97(4): 515–528.
- Marshall, John. 2021. "Can Close Election Regression Discontinuity Designs Identify Effects of Winning Politician Characteristics?"
- Matland, Richard E. 1998. "Women's Representation in National Legislatures: Developed and Developing Countries." *Legislative Studies Quarterly* 23(1): 109–125.
- Matland, Richard E., and Donley T. Studlar. 1996. "The Contagion of Women Candidates in Single-Member District and Proportional Representation Electoral Systems: Canada and Norway." *The Journal of Politics* 58(3): 707–733.
- McCrary, Justin. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics* 142(2): 698–714.
- Meserve, Stephen A., Daniel Pemstein, and William T. Bernhard. 2020. "Gender, Incumbency and Party List Nominations." *British Journal of Political Science* 50(1): 1–15.
- Murray, Rainbow, Mona Lena Krook, and Katherine AR Opello. 2012. "Why Are Gender Quotas Adopted? Party Pragmatism and Parity in France." *Political Research Quarterly* 65(3): 529–543.
- Norris, Pippa. 1985. "Women's Legislative Participation in Western Europe." In *Routledge Studies in Political Psychology*. Vol. 4 Routledge/Taylor pp. 90–101.
- O'Brien, Diana Z., and Johanna Rickne. 2016. "Gender Quotas and Women's Political Leadership." *American Political Science Review* 110(1): 112–126.
- Piscopo, Jennifer M., Magda Hinojosa, Gwynn Thomas, and Peter M. Siavelis. 2021. "Follow the Money: Gender, Incumbency, and Campaign Funding in Chile." *Comparative Political Studies* (Online).

- Ragauskas, Rimvydas. 2021. "Party-Determined Viability and Gender Bias in Open-List Proportional Representation Systems." *Politics & Gender* 17(2): 250–276.
- Roberts, Andrew, Jason Seawright, and Jennifer Cyr. 2013. "Do Electoral Laws Affect Women's Representation?" *Comparative Political Studies* 46(12): 1555–1581.
- Schwindt-Bayer, Leslie A. 2005. "The Incumbency Disadvantage and Women's Election to Legislative Office." *Electoral Studies* 24(2): 227–244.
- Sen, Maya, and Omar Wasow. 2016. "Race as a Bundle of Sticks: Designs That Estimate Effects of Seemingly Immutable Characteristics." *Annual Review of Political Science* 19: 499–522.
- Smrek, Michal. 2020. "Do Female Legislators Benefit from Incumbency Advantage? Incumbent Renomination in a Flexible-List PR System." *Electoral Studies* 66.
- Teele, Dawn Langan, Joshua Kalla, and Frances Rosenbluth. 2018. "The Ties That Double Bind: Social Roles and Women's Underrepresentation in Politics." *American Political Science Review* 112(3): 525–541.
- Thames, Frank C. 2017. "Understanding the Impact of Electoral Systems on Women's Representation." *Politics & Gender* 13(3): 379–404.
- Thames, Frank C., and Margaret S. Williams. 2010. "Incentives for Personal Votes and Women's Representation in Legislatures." *Comparative Political Studies* 43(12): 1575–1600.
- Verge, Tània, and Nina Wiesehomeier. 2019. "Parties, Candidates, and Gendered Political Recruitment in Closed-List Proportional Representation Systems: The Case of Spain." *Political Research Quarterly* 72(4): 805–820.
- Wasserman, Melanie. 2020. "Gender Differences in Politician Persistence." *Review of Economics and Statistics* (Forthcoming).

Online Appendix

Intended for online publication only.

Table of Contents

A Illustrative Simulation of Open-List Theory	3
A.1 Setup	3
A.2 Simulation	4
A.3 Results	6
B RD Robustness Checks: Main Results	7
B.1 Functional Form	7
B.2 Bandwidth Sensitivity	9
B.3 Kernel Choice	11
B.4 Estimates without Fixed Effects	13
C RD Robustness Checks: Identification Assumptions	15
C.1 Continuity Around Threshold	15
C.2 Continuity in Share Of Female Candidates	16
D Identifying Close Winners and Losers under Proportional Representation	17
E Additional Statistical Results: Norway	19
E.1 Gender Gap Persists In Restricted Sample	19
E.2 No Evidence of Negative Gender Gap in County-Level Elections	20
E.3 Additional Outcomes Robust To Functional Form Choices	21
E.4 Estimating the Effect on Personal Votes	22
E.5 Rank Position and Personal Votes	27
E.6 Differential Selection and Characteristics Correlated With Gender	28
E.7 Matching Estimates	30
E.8 Heterogeneity By Population Size	32
E.9 Survey Evidence On Voters' Attitudes Towards Female Candidates	34
F Additional Statistical Results: Spain	35
F.1 Estimates from Difference-in-Differences Specification	35
F.2 Heterogeneity Across Parties	36

F.3	Heterogeneity Across Quota Rules	37
F.4	Heterogeneity By Population Size	42
G	Additional Statistical Results: Poland	43
H	Descriptive Statistics	45

A Illustrative Simulation of Open-List Theory

My argument in the paper rests on the idea that in open-list PR, female incumbents may receive a smaller boost in preference votes compared to male incumbents, thereby inducing a smaller effect of winning on the probability of winning again (incumbency advantage). In this appendix, I sketch out a toy model explaining what conditions may induce such a result.

A.1 Setup

In this toy model, I focus on candidates running in one party across two elections, in $t = 0$ and $t = 1$. To keep exposition simple, I assume that the party always fields 8 candidates and is slated to win 4 seats. Under the open-list system resembling Norway's (without pre-advantage votes), the party first draws up a list of candidates; voters then assign (up to) 4 preference votes. The four candidates with the highest preference vote share then win election.

Timing of the Game.

Election at $t = 0$

1. Draw 2 initial incumbents with quality $\theta_i \sim \mathcal{U}[0.5, 1]$ and 6 additional candidates with quality $\theta_i \sim \mathcal{U}[0, 1]$. Here and thereafter, assume that every candidate drawn has an equal chance of being male and female.
2. The party elites observe a noisy (and possibly biased) quality signal, $\hat{\theta}_i$ and rank the initial list in descending order.
3. A unit mass of voters, comprising a share p of (unbiased) quality-oriented voters, a share r of (unbiased) heuristic voters, and a share q of biased heuristic voters, casts their preference votes.
4. The election is resolved; 4 successful candidates are designated as new incumbents.

Election at $t = 1$

1. Candidates from the previous election decide whether to run again; vacant candidacies are filled up with new draws, as before, with $\theta_i \sim \mathcal{U}[0, 1]$.
2. Incumbents who run again increase their quality θ_i by 0.05.
3. Once more, the party elites observe $\hat{\theta}_i$ and rank the initial list in descending order.
4. The same unit mass of voters casts their preference votes; the election is then resolved.

Actors and Steps. *Party Elites.* Party elites observe a noisy and possibly biased quality signal:

$$\hat{\theta}_i = \theta_i + b_i \cdot 1(Female) + v_i \cdot 1(Incumbent) + u_i \quad (3)$$

where $b_i \leq 0$ is a bias term for female candidates, $v_i \geq 0$ is a bias term for incumbents, and $u_i \sim \mathcal{N}(0, 0.05)$ is the noise term.

Voters. There are three types of voters.

1. Quality-oriented voters observe a noisy signal $\tilde{\theta} = \theta + e_i$ where $e_i \sim \mathcal{N}(0, 0.1)$.
2. Heuristic voters award preference votes to incumbents first. If there are fewer incumbents than available preference votes, they will award the remaining preference votes to non-incumbent candidates in descending list rank order, irrespective of gender.
3. Biased heuristic voters award preference votes in a similar manner to unbiased heuristic voters, except that they skip any and all female candidates.

Of the unit mass of voters, there are p quality-oriented voters, r heuristic voters, and q biased heuristic voters.

Candidates. At the beginning of $t = 1$, candidates who ran in $t = 0$ have to decide whether to run again. Incumbents will do so with probability $\pi_1 = 0.8$, while non-incumbents will do so with $\pi_0 = 0.4$.

A.2 Simulation

I simulate the model described above for different parameter values (scenarios). For every scenario, I run 100 iterations of 200 simulations each.

I present the distribution of the probability of winning for three different subgroups of candidates at time $t = 1$:

1. $Pr(Win_{t_1} | Won_{t_0})$, the win rate among incumbents;
2. $Pr(Win_{t_1} | Lost_{t_0})$, the win rate among previous losers;
3. $Pr(Win_{t_1} | Non - incumbent_{t_1})$, the win rate among all non-incumbents running in $t = 1$. This statistic gives us a sense of whether women face a ‘sticky floor’ in getting elected in the first place.

I introduce four different scenarios; for each scenario I run two sets of simulations: one with no party bias ($b_i = 0$) and one with some party bias ($b_i = -0.1$) against women, respectively. I also keep the party bias towards incumbents, $v_i = 0.1$ throughout.

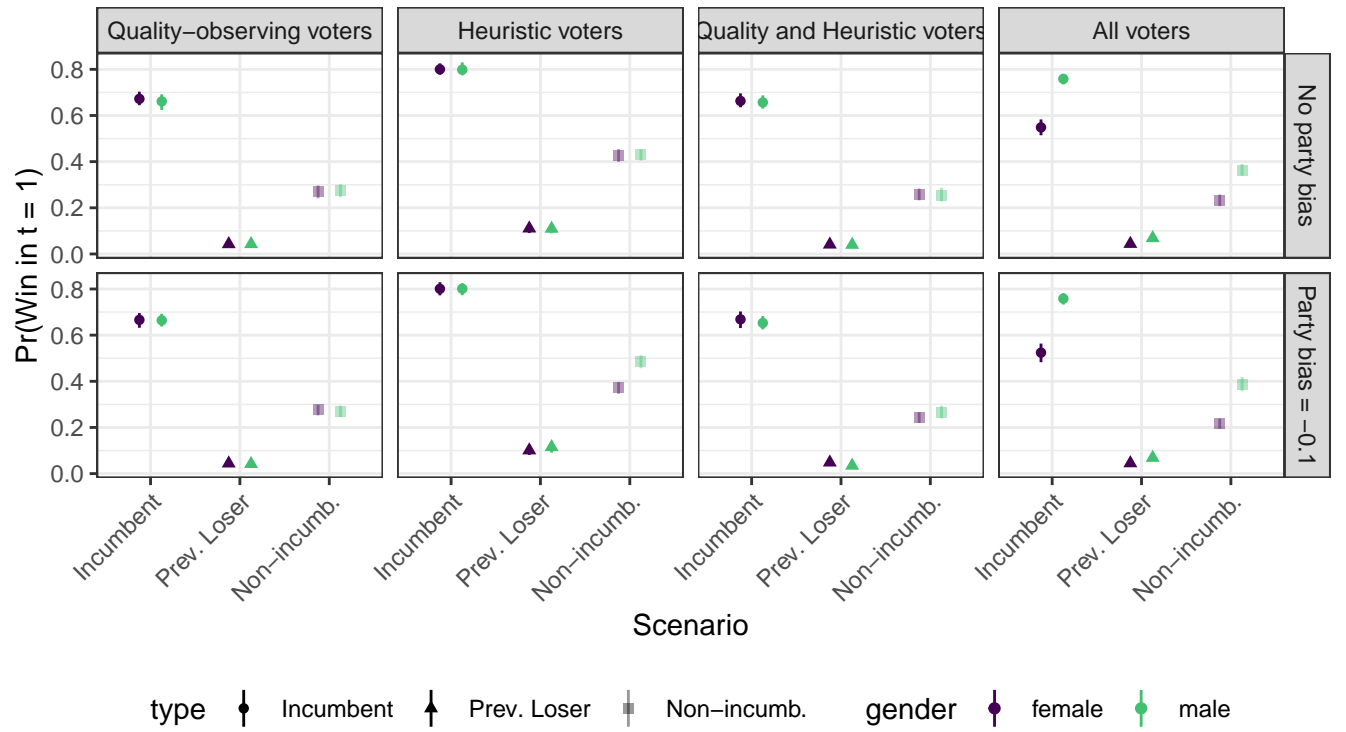


Figure A1: Simulation results showing probability of election for candidate subgroups across scenarios

A.3 Results

Figure A1 shows the shares of candidates (previous winners, previous losers, non-incumbents) who win in $t = 1$. Within each panel, we observe the probability of winning for candidates who won in $t = 0$, for candidates who lost in $t = 0$, and, in a lighter shade, for non-incumbents who ran in $t = 1$ (i.e., the unconditional probability of winning for non-incumbents).

Crucially, we see no gender gaps between previous winners and previous losers across the first three scenarios. It is only when we add the biased heuristic voters into the election that we see that female candidates observe a smaller increase in their win probability as a result of incumbency when compared to men.

B RD Robustness Checks: Main Results

In this section, I offer a number of checks to confirm that the main results in the paper are not sensitive to different specifications of the RD design.

B.1 Functional Form

In this section, I assess the robustness of the main RD estimates with respect to functional form. Tables B1 and B2 below report the results from the main specification fitted with second-order polynomials. The results are consistent with the estimates from the linear specification.

B.1.1 Norway

Table B1: Difference-in-Discontinuity Estimates from Norway with Second-Order Polynomial.

	Run ($t + 1$)			Win ($t + 1$)		
	(1)	(2)	(3)	(4)	(5)	(6)
Elected	0.000 (0.019)	-0.011 (0.016)	0.022 (0.023)	0.106 (0.017)	0.100 (0.015)	0.112 (0.022)
Female	-0.083 (0.020)	-0.080 (0.018)	-0.097 (0.023)	-0.025 (0.017)	-0.020 (0.014)	-0.032 (0.019)
Elected x Female	-0.015 (0.032)	0.003 (0.029)	-0.030 (0.038)	-0.074 (0.029)	-0.058 (0.025)	-0.091 (0.035)
Bandwidth	0.11	0.22	0.055	0.091	0.18	0.045
BW Type	Optimal	2x Opt	0.5x Opt	Optimal	2x Opt	0.5x Opt
Outcome Mean	0.576	0.587	0.565	0.263	0.262	0.261
N (left)	5553	6537	4660	5314	6216	4384
N (right)	5614	6609	4709	5373	6283	4431

All estimates are reported with robust standard errors clustered at the municipality level in parentheses. Each observation is a candidate's election attempt. 'Elected' is an indicator for observations where the candidate obtained a seat in the municipal council. 'Female' is an indicator for observations identified as female. 'Elected' times 'Female' is the interaction between the two variables. Other coefficients not reported. Regression run on all candidates in elections between 2003 and 2015.

B.1.2 Spain

Table B2: Difference-in-Discontinuity Estimates from Spain with Second-Order Polynomial..

	Run (in t+1)			Win (in t+1)		
	(1)	(2)	(3)	(4)	(5)	(6)
Elected	0.095 (0.020)	0.116 (0.014)	0.075 (0.029)	0.076 (0.014)	0.087 (0.010)	0.074 (0.020)
Female	-0.014 (0.023)	-0.014 (0.017)	0.008 (0.032)	-0.027 (0.015)	-0.030 (0.010)	-0.037 (0.021)
Elected x Female	0.043 (0.032)	0.021 (0.023)	0.011 (0.046)	0.045 (0.022)	0.053 (0.016)	0.056 (0.031)
Bandwidth	0.036	0.072	0.018	0.059	0.12	0.029
BW Type	Optimal	2x Opt	0.5x Opt	Optimal	2x Opt	0.5x Opt
Outcome Mean	0.487	0.481	0.486	0.253	0.244	0.255
N (left)	16564	28924	8917	24559	40700	13705
N (right)	16568	29555	8708	25132	40620	13587

All estimates are reported with robust standard errors clustered at the municipality level in parentheses. Each observation is a candidate's election attempt. 'Elected' is an indicator for observations where the candidate obtained a seat in the municipal council. 'Female' is an indicator for observations identified as female. 'Elected' times 'Female' is the interaction between the two variables. Other coefficients not reported. Regression run on all candidates in elections between 2003 and 2015.

B.2 Bandwidth Sensitivity

Tables 2 and 3 report results of the difference-in-discontinuity specification with optimal bandwidth as selected by Calonico, Cattaneo, and Titiunik (2014) on the full sample, along with one-half and double that bandwidth. Below, I plot the coefficient estimates of interest with additional bandwidth parameters to ascertain that the results are robust to the choice of bandwidth.

B.2.1 Norway

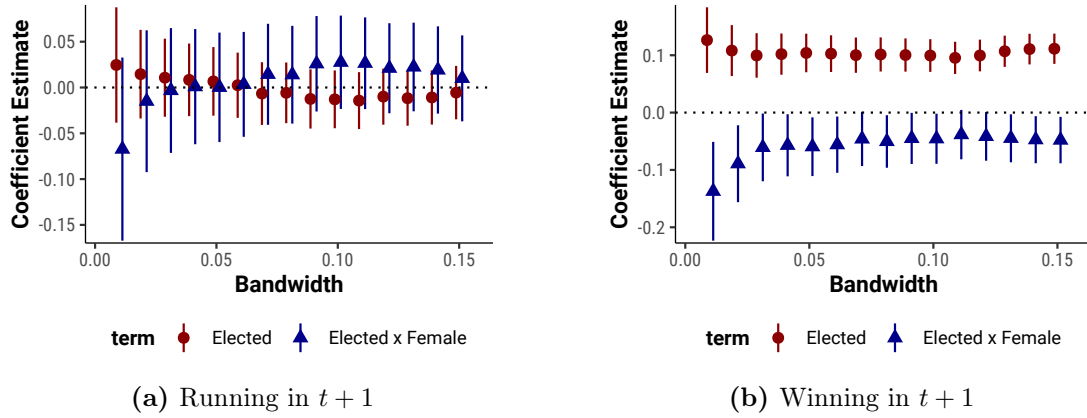
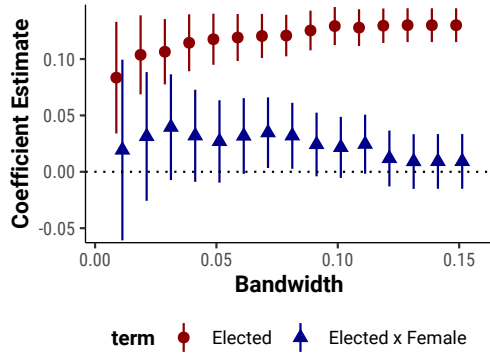
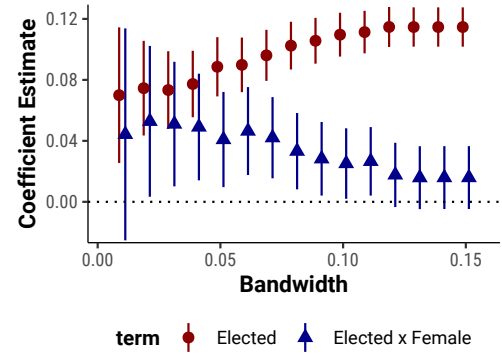


Figure B1: Bandwidth Sensitivity Check (Norway). The plot shows the RD coefficients for ‘Elected’ and ‘Elected x Female’ in the case of Norway for different bandwidth choices.

B.2.2 Spain



(a) Running in $t + 1$



(b) Winning in $t + 1$

Figure B2: Bandwidth Sensitivity Check (Spain). The plot shows the RD coefficients for ‘Elected’ and ‘Elected x Female’ in the case of Spain for different bandwidth choices.

B.3 Kernel Choice

Below, I present estimates from the main specification with triangular, rather than uniform kernels. The results are very similar to the preferred estimates in the main body of the paper in both countries.

Table B3: Difference-in-Discontinuity Estimates For Incumbency Advantage In Norwegian Municipalities, with triangular kernels. Women Face Diminished Incumbency Effect On Winning Again.

	Run ($t + 1$)			Win ($t + 1$)		
	(1)	(2)	(3)	(4)	(5)	(6)
Elected	0.011 (0.020)	-0.003 (0.017)	0.018 (0.025)	0.106 (0.018)	0.102 (0.015)	0.112 (0.022)
Female	-0.089 (0.023)	-0.089 (0.019)	-0.081 (0.027)	-0.031 (0.019)	-0.030 (0.016)	-0.022 (0.024)
Elected x Female	-0.019 (0.035)	0.007 (0.030)	-0.040 (0.041)	-0.071 (0.030)	-0.053 (0.025)	-0.102 (0.037)
Bandwidth	0.054	0.11	0.027	0.05	0.099	0.025
BW Type	Optimal	2x Opt	0.5x Opt	Optimal	2x Opt	0.5x Opt
Outcome Mean	0.564	0.575	0.551	0.261	0.263	0.257
N (left)	4617	5529	3667	4502	5428	3549
N (right)	4666	5590	3711	4551	5489	3592

All estimates are reported with robust standard errors clustered at the municipality level in parentheses. Each observation is a candidate's election attempt. 'Elected' is an indicator for observations where the candidate obtained a seat in the municipal council. 'Female' is an indicator for observations identified as female. 'Elected' times 'Female' is the interaction between the two variables. Other coefficients not reported. Regression run on all candidates in elections between 2003 and 2015.

Table B4: Difference-in-Discontinuity Estimates For Incumbency Advantage In Spanish Municipalities, with triangular kernels. Women enjoy a larger effect of winning on their probability to win again.

	Run ($t + 1$)			Win ($t + 1$)		
	(1)	(2)	(3)	(4)	(5)	(6)
Elected	0.103 (0.016)	0.116 (0.011)	0.093 (0.022)	0.075 (0.013)	0.088 (0.009)	0.074 (0.019)
Female	-0.018 (0.018)	-0.027 (0.013)	-0.005 (0.025)	-0.030 (0.014)	-0.026 (0.010)	-0.030 (0.019)
Elected x Female	0.038 (0.025)	0.031 (0.018)	0.020 (0.035)	0.055 (0.021)	0.046 (0.015)	0.049 (0.029)
Bandwidth	0.032	0.064	0.016	0.035	0.071	0.018
BW Type	Optimal	2x Opt	0.5x Opt	Optimal	2x Opt	0.5x Opt
Outcome Mean	0.486	0.483	0.485	0.256	0.251	0.254
N (left)	14840	26294	7984	16223	28426	8726
N (right)	14729	26886	7825	16159	29072	8530

All estimates are reported with robust standard errors clustered at the municipality level in parentheses.

B.4 Estimates without Fixed Effects

Below, I show results from difference-in-discontinuity specifications mirroring the main specification, but estimated without fixed effects. This specification matches the heterogeneity-in-discontinuities estimator discussed in [Bansak and Nowacki \(2022\)](#). The results are very similar.

Table B5: Difference-in-Discontinuity Estimates For Incumbency Advantage In Norwegian Municipalities, without fixed effects. Women Face Diminished Incumbency Effect On Winning Again.

	Run (t + 1)			Win (t + 1)		
	(1)	(2)	(3)	(4)	(5)	(6)
Elected	-0.002 (0.019)	-0.014 (0.016)	0.014 (0.023)	0.110 (0.017)	0.104 (0.015)	0.110 (0.021)
Female	-0.106 (0.021)	-0.099 (0.018)	-0.092 (0.026)	-0.025 (0.017)	-0.023 (0.015)	-0.022 (0.021)
Elected x Female	0.008 (0.033)	0.032 (0.028)	-0.031 (0.039)	-0.069 (0.028)	-0.050 (0.024)	-0.099 (0.035)
Bandwidth	0.054	0.11	0.027	0.05	0.099	0.025
BW Type	Optimal	2x Opt	0.5x Opt	Optimal	2x Opt	0.5x Opt
Outcome Mean	0.564	0.575	0.551	0.261	0.263	0.257
N (left)	4617	5529	3667	4502	5428	3549
N (right)	4666	5590	3711	4551	5489	3592

All estimates are reported with robust standard errors clustered at the municipality level in parentheses. Each observation is a candidate's election attempt. 'Elected' is an indicator for observations where the candidate obtained a seat in the municipal council. 'Female' is an indicator for observations identified as female. 'Elected' times 'Female' is the interaction between the two variables. Other coefficients not reported. Regression run on all candidates in elections between 2003 and 2015.

Table B6: Difference-in-Discontinuity Estimates For Incumbency Advantage In Spanish Municipalities, Without Fixed Effects. Women enjoy a larger effect of winning on their probability to win again.

	Run ($t + 1$)			Win ($t + 1$)		
	(1)	(2)	(3)	(4)	(5)	(6)
Elected	0.109 (0.014)	0.120 (0.010)	0.097 (0.020)	0.074 (0.012)	0.097 (0.008)	0.080 (0.017)
Female	-0.020 (0.016)	-0.030 (0.012)	-0.010 (0.022)	-0.029 (0.013)	-0.023 (0.009)	-0.026 (0.018)
Elected x Female	0.034 (0.023)	0.031 (0.017)	0.019 (0.032)	0.054 (0.019)	0.043 (0.014)	0.054 (0.027)
Bandwidth	0.032	0.064	0.016	0.035	0.071	0.018
BW Type	Optimal	2x Opt	0.5x Opt	Optimal	2x Opt	0.5x Opt
Outcome Mean	0.486	0.483	0.485	0.256	0.251	0.254
N (left)	14840	26294	7984	16223	28426	8726
N (right)	14729	26886	7825	16159	29072	8530

All estimates are reported with robust standard errors clustered at the municipality level in parentheses.

C RD Robustness Checks: Identification Assumptions

In this section, I discuss additional checks verifying that there is no sorting around the RD threshold.

C.1 Continuity Around Threshold

In order for the regression discontinuity design to identify the causal effect of being elected within each gender, the key assumption is continuity around the threshold – that is, observations on either side of the threshold are comparable in all respects except for the treatment assignment and the outcome.

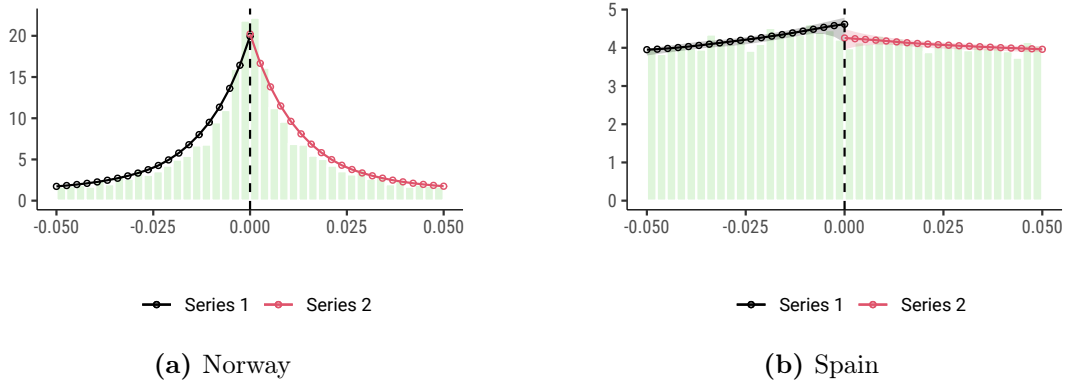


Figure C1: No Discontinuity in Running Variable Density Around Threshold. The plots show a local linear regression fitted on the density of the running variable.

C.2 Continuity in Share Of Female Candidates

Similarly, if there is no sorting around the threshold, we should observe no discontinuous jump in the share of female candidates at the threshold. Figure C2 confirms that this is the case.

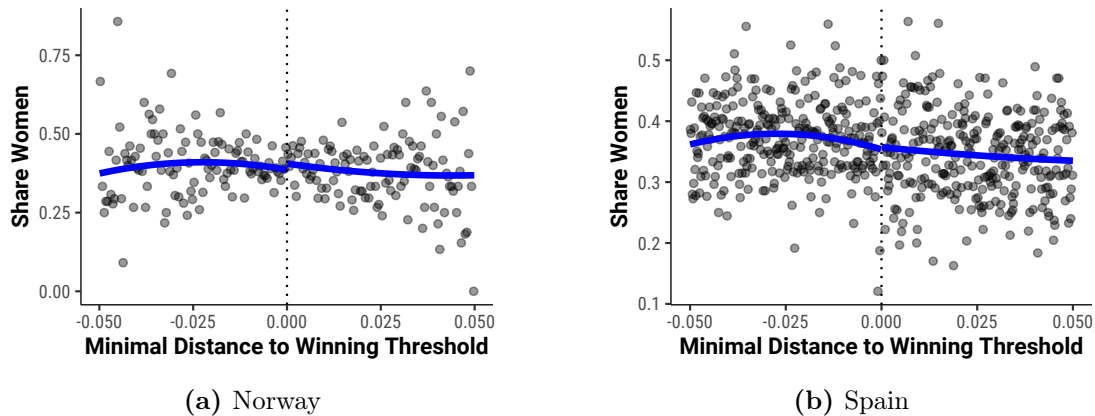


Figure C2: No Discontinuity in Share Of Female Candidates Around Threshold. The plots show a local quadratic regression predicting the share of female candidates around the threshold.

D Identifying Close Winners and Losers under Proportional Representation

In the context of PR elections, constructing the running variable and identifying which candidates came close to barely losing or barely winning is not as straightforward as in the archetypal plurality case (Lee, 2001, 2008; Fiva and Halse, 2016; Fiva, Folke, and Sørensen, 2018). These individual-level measures are different from Blais and Lago (2009) and Cox, Fiva, and Smith (2020), who measure competitiveness at the district level.

Norway (open-list). I am constructing the margin of victory (running variable) following Fiva and Røhr (2018). Effectively, the strategy identifies bare winners and bare winners within each party as those candidates who have just about gained enough preference votes (or just about missed out on enough preference votes) to be elected (to be defeated).³⁴ The as-if random element comes from the distribution of preference votes across candidates of the same list. Following Fiva and Rohr’s notation, the margin of victory for candidate i on party list l is defined as:

$$MV_{il} = \begin{cases} \frac{Poll_{il} - Poll_l^{S_{l+1}}}{PartyVotes_l} & \text{if elected} \\ \frac{Poll_{il} - Poll_l^{S_l}}{PartyVotes_l} & \text{if not elected} \end{cases} \quad (4)$$

where $Poll_{il}$ denotes candidates’ personal vote share, $Poll_l^{S_{l+1}}$ denotes the vote share of party l ’s first loser, and $Poll_l^{S_l}$ denotes the vote share of party l ’s last winner. $PartyVotes_l$ is the total number of votes cast for party l . In other words, I construct the margin of victory using the difference in personal votes between candidate i and the vote threshold to winning a seat, normalized by the total number of votes the party obtained.

I construct the running variable in the same way for my supplementary analysis of county-level elections in Poland.

Spain (closed-list). I construct the running variable following Folke (2014) and calculate a ‘minimum distance’ to winning (or losing) an extra seat for the legislator’s party.³⁵ The as-if random element comes from the discontinuous translation from parties’ vote shares to a discrete number of seats, where a small perturbation in the vote share may lead to a

³⁴ One issue with this choice of running variable is that the observations closest to either side of the threshold are likely to have greater population and district magnitude. This follows from the distribution of personal votes across the number of candidates within a list – very tight margins (e.g. less than a percentage point) are more likely when there are many candidates with similar, small shares of personal votes. Although this issue does not confound the LATE estimate at the threshold (since units do not sort according to treatment status), it does pose a problem of adequate functional form for outcomes such as personal vote shares in $t + 1$. See Appendix E.4 for more details.

³⁵ Note, in the original application of this method, the algorithm does not account for electoral thresholds below which parties are excluded from the seat allocation altogether. In the Spanish case, there is a 4% threshold. However, because of low district magnitudes, the *effective* threshold necessary for parties to clear to win seats, is typically higher than this. Still, I exclude candidates from lists that scored less than 10% in an election from my sample to ensure that the minimum distance calculated is not affected by the legal threshold.

different candidate winning the last available seat. I compute the smallest change in the vote share distribution that would cause a party to either lose or win an additional seat. In other words, in each municipal election, I identify the closest winner (the last seat assigned) and the closest loser (the first candidate to miss out on a seat), and calculate the necessary change vote share to flip the last seat from the winner to the loser.

Note that the setup of the running variable in the two countries identifies bare winners and bare losers in different ways. However, since my quantity of interest is the LATE subgroup estimate at the threshold, both running variables identify bare winners and bare losers in their respective electoral context. It is typical for different electoral systems to require different operationalisations of the margin of victory. See [Eggers et al. \(2015\)](#) for plurality, [Folke \(2014\)](#); [Fiva, Folke, and Sørensen \(2018\)](#); [Fiva and Halse \(2016\)](#) for closed-list PR, [Folke, Persson, and Rickne \(2016\)](#); [Fiva and Røhr \(2018\)](#) for open-list PR. Moreover, since I am using party-by-year fixed effects in my specifications, my results in Norway are also not driven by selection of candidates into specific parties.

E Additional Statistical Results: Norway

In this section, I report supplementary analyses and additional results from the Norwegian case.

E.1 Gender Gap Persists In Restricted Sample

Below, I report the results from Table 2 estimated on the sample restricted to municipality-years in which no personal votes are recorded as missing. There are two reasons why these adjustments should have little effect. First, most records with missing personal votes are outside the borderline sample – candidates from minor parties or in very small municipalities. Second, even if personal vote information were missing, other election details such as list rank and whether the candidate is running, and winning (again) persist. This feature of the data should therefore not affect the main conclusions. Indeed, Table E1 demonstrates that this is the case, although the uncertainty associated with the estimates increases due to a smaller sample size.

Table E1: Difference-in-Discontinuity Estimates For Incumbency Advantage In Norwegian Municipalities. Run On Restricted Sample.

	Run (t + 1)			Win (t + 1)		
	(1)	(2)	(3)	(4)	(5)	(6)
Elected	0.001 (0.024)	−0.004 (0.020)	0.016 (0.029)	0.100 (0.021)	0.095 (0.018)	0.099 (0.025)
Female	−0.098 (0.024)	−0.090 (0.021)	−0.063 (0.029)	−0.032 (0.020)	−0.027 (0.017)	−0.017 (0.025)
Elected x Female	0.003 (0.038)	0.015 (0.031)	−0.045 (0.045)	−0.061 (0.033)	−0.048 (0.028)	−0.093 (0.041)
Bandwidth	0.053	0.11	0.027	0.049	0.098	0.025
BW Type	Optimal	2x Opt	0.5x Opt	Optimal	2x Opt	0.5x Opt
Outcome Mean	0.561	0.573	0.545	0.259	0.26	0.253
N (left)	3227	3853	2559	3156	3783	2486
N (right)	3272	3909	2599	3201	3839	2525

All estimates are reported with robust standard errors clustered at the municipality level in parentheses. Each observation is a candidate’s election attempt. ‘Elected’ is an indicator for observations where the candidate obtained a seat in the municipal council. ‘Female’ is an indicator for observations identified as female. ‘Elected’ times ‘Female’ is the interaction between the two variables. Other coefficients not reported. Regression run on all candidates in elections between 2003 and 2015, excluding municipality-years in which at least some candidates are recorded as missing personal votes.

E.2 No Evidence of Negative Gender Gap in County-Level Elections

Table E2: Probabilities of Winning Re-Election in Norway’s Regional Elections.. Female incumbents are unlikely to experience a smaller incumbency advantage than men in this context.

	Win (in t+1)					
	(1)	(2)	(3)	(4)	(5)	(6)
Elected	0.183 (0.036)	0.164 (0.039)	0.106 (0.052)	0.121 (0.055)	0.153 (0.059)	0.156 (0.059)
Female	-0.051 (0.032)	-0.056 (0.036)	-0.056 (0.068)	-0.032 (0.072)	-0.091 (0.050)	-0.092 (0.050)
Elected x Female	0.015 (0.051)	0.015 (0.054)	0.019 (0.076)	-0.007 (0.079)	0.064 (0.075)	0.061 (0.076)
R2	0.058	0.098	0.069	0.117	0.057	0.073
Parties	All	All	Left	Left	Right	Right
N	1700	1700	412	412	825	825
Outcome Mean	0.248	0.248	0.257	0.257	0.273	0.273
Region FE	Y	Y	Y	Y	Y	Y
Year FE	Y	N	Y	N	Y	N
Year x Party FE	N	Y	N	Y	N	Y

All estimates are reported with robust standard errors clustered at the region level in parentheses. Each observation is a candidate’s election attempt. ‘Elected’ is an indicator for observations where the candidate obtained a seat in the municipal council. ‘Female’ is an indicator for observations identified as female. ‘Elected’ times ‘Female’ is the interaction between the two variables. Regression run on all candidates in regional elections between 2003 and 2015.

E.3 Additional Outcomes Robust To Functional Form Choices

Below, I replicate Table 6 with second-order polynomials. The conclusions are unchanged, except for the gender gap in personal votes. As I discuss in Appendix E.4 below, estimation of the personal vote share can be challenging due to obstacles in operationalisation. When using alternative measures (for example raw personal vote counts) as an outcome, the gender gap in personal votes persists even with second-order polynomials (see Table E6).

Table E3: Difference-in-Discontinuity Estimates On Additional Outcomes, Norway, With Second-Order Polynomials. The results remain broadly consistent.

	Orig. Rank Advance		Actual Rank Advance		Pre-Ad. (t+1)		Pers.V. Share (t+1)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Elected	0.092 (0.031)	0.084 (0.024)	0.037 (0.032)	0.065 (0.024)	0.041 (0.025)	0.047 (0.023)	-0.000 (0.003)	0.001 (0.003)
Female	0.011 (0.032)	0.001 (0.023)	-0.023 (0.033)	0.008 (0.023)	0.034 (0.028)	0.053 (0.021)	-0.001 (0.003)	-0.000 (0.002)
Elected x Female	-0.079 (0.048)	-0.064 (0.043)	-0.023 (0.046)	-0.069 (0.039)	-0.005 (0.042)	-0.061 (0.040)	0.001 (0.004)	0.000 (0.004)
Parties	Left	Right	Left	Right	Left	Right	Left	Right
Bandwidth	0.068	0.092	0.061	0.097	0.033	0.055	0.025	0.052
Outcome Mean	0.257	0.253	0.229	0.211	0.119	0.163	0.0204	0.0175
N (left)	1655	2620	1624	2662	1421	2203	446	840
N (right)	1676	2655	1645	2698	1441	2229	458	856

All estimates are reported with robust standard errors clustered at the municipality level in parentheses. Each observation is a candidate's election attempt. 'Elected' is an indicator for observations where the candidate obtained a seat in the municipal council. 'Female' is an indicator for observations identified as female. 'Elected' times 'Female' is the interaction between the two variables. Other coefficients not reported. Regression run on all candidates in elections between 2003 and 2015.

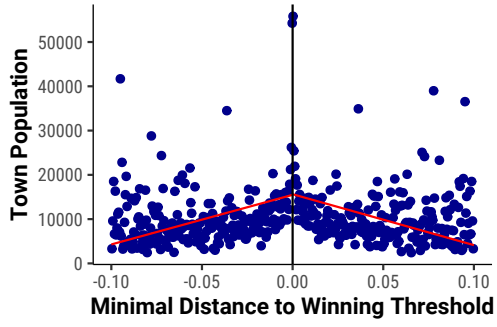
E.4 Estimating the Effect on Personal Votes

One additional challenge in the Norwegian case is that larger municipalities are more likely to be close to the threshold, because differences in candidates’ personal vote scores are likely to be smaller in districts with large district magnitude and a large number of candidates (e.g., two candidates at the threshold are more likely to be separated by just a small vote share if there are 50 other candidates on the list).

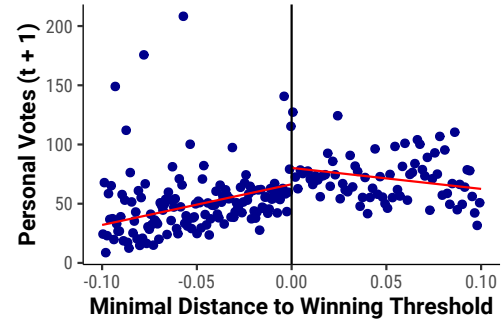
Fortunately, this does not raise concerns with the main results, where outcomes are binary, as we have seen that rates of running again and winning again are relatively constant (or linear) along the running variable (cf. Figure 2). It does, however, bring about a problem when using outcomes related to personal votes in $t + 1$. Candidates closest to the threshold in t are not only more likely to run in larger cities, they are also more likely to have smaller personal vote shares (compared to the total number of votes cast for either the party or in the municipality). While this does not violate the no-sorting condition *across* the threshold, it introduces the threat of non-linear conditional expectation functions on either side of the threshold, which makes estimation of the true LATE much more difficult.

Figures E1 and E2 illustrate this problem. Figure E1 plots the average town population for bins near the threshold. We see that candidates closest to the threshold are more likely to run in larger towns. Fortunately, in the case of raw personal votes (right panel), no significant non-linearities appear: towns near the threshold are larger, but with higher district magnitudes, threshold candidates are also placed further down the list, so their raw vote count does not change all that much. Note that any estimate of the gender gap in the effect of being elected on raw personal votes may be distorted if men and women, on average, receive different vote counts to start with. I find no evidence for this concern in the threshold sample, with the average male candidate receiving 52 personal votes in t , and the average female candidate receiving 51.

The story is different for measures of personal vote *shares*. First, there are multiple ways to express the share: either as a share of all votes cast in the town (‘Total Votes’), or as a share of all votes cast in the town for the candidate’s particular party (‘Party Votes’, following Fiva, Folke, and Sørensen (2018)). In both cases, we observe that the outcome decreases severely close to the threshold, likely making estimation of the effect at the threshold more difficult. While an appealing operationalisation of candidates’ preference votes (and hence reported in the main text), some of the additional estimates using share-based measures end up being overly noisy due to this issue.

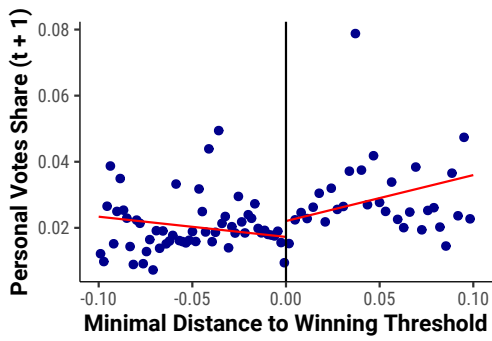


(a) Municipality Population

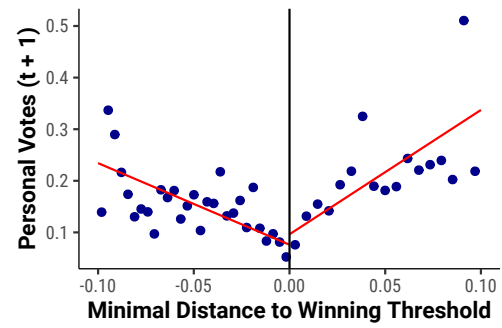


(b) Raw Personal Votes ($t + 1$)

Figure E1: Distribution of Municipality Population size and Raw Personal Votes around Threshold



(a) Personal Vote / Total Votes ($t + 1$)



(b) Personal Vote / Party Votes ($t + 1$)

Figure E2: Distribution of Personal Vote Shares Around Threshold

E.4.1 Robustness to Personal Vote Measurement

Because of the aforementioned challenges, I estimate the difference-in-discontinuity specification with all the different outcome operationalisations discussed above. Table E.4 reports the results. The gender gap in the effect of incumbency on personal vote increases across right-wing parties persists across all specifications, although somewhat imprecisely estimated.

Table E4: Sensitivity of Results to Different Personal Vote Measures. The gender gap estimate is consistent across preferred measures.

	Raw PV		log(Raw PV)		PV / Total Votes		PV / Party Votes	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Elected	13.842 (6.882)	23.573 (5.948)	0.243 (0.085)	0.340 (0.067)	0.003 (0.002)	0.005 (0.001)	0.012 (0.009)	0.031 (0.010)
Female	-9.542 (8.383)	20.199 (27.847)	-0.153 (0.100)	-0.019 (0.101)	-0.003 (0.003)	0.000 (0.002)	0.006 (0.008)	0.004 (0.011)
Elected x Female	12.549 (14.673)	-34.571 (29.305)	0.095 (0.156)	-0.229 (0.139)	0.001 (0.004)	-0.005 (0.003)	-0.009 (0.013)	-0.024 (0.015)
Parties	Left	Right	Left	Right	Left	Right	Left	Right
Bandwidth	0.067	0.093	0.05	0.073	0.025	0.052	0.013	0.046
Outcome Mean	65.6	69.1	3.77	3.76	0.0204	0.0175	0.069	0.124
N (left)	586	1019	553	941	446	840	352	791
N (right)	604	1048	564	962	458	856	358	803

All estimates are reported with robust standard errors clustered at the municipality level in parentheses. Each observation is a candidate's election attempt. 'Elected' is an indicator for observations where the candidate obtained a seat in the municipal council. 'Female' is an indicator for observations identified as female. 'Elected' times 'Female' is the interaction between the two variables. Other coefficients not reported. Regression run on all candidates in elections between 2003 and 2015.

E.4.2 Robustness to Data Sample

All analyses on personal vote outcomes so far were restricted to candidates in municipalities where all personal votes are recorded. This is important because we want our denominators for the shares to be right. Still, for robustness’s sake, I also report the estimates for the full sample in Table E5. Note that I can still extend the analyses to these additional observations because the personal vote measure(s) are not missing for the borderline candidates – they are merely missing for other (usually minor) candidates running in the municipality.

The estimates for raw personal vote and vote share (relative to total votes) remain consistent, and even gain some precision. The estimates for PV / Party Votes do change, however and are very noisy. This is, as mentioned before, likely due to the estimation challenge with this outcome in particular, where the results are sensitive to even a small degree of noise or measurement error.

Table E5: Sensitivity of Results to Different Personal Vote Measures. Including Municipalities With Incomplete Personal Vote Data.

	Raw PV		log(Raw PV)		PV / Total Votes		PV / Party Votes	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Elected	18.040 (6.289)	25.384 (4.625)	0.228 (0.068)	0.381 (0.056)	0.003 (0.002)	0.004 (0.001)	0.009 (0.007)	0.016 (0.009)
Female	-3.911 (6.929)	16.047 (20.862)	-0.098 (0.092)	-0.018 (0.089)	-0.002 (0.003)	0.001 (0.001)	0.004 (0.006)	0.012 (0.011)
Elected x Female	2.739 (11.299)	-36.455 (22.658)	0.014 (0.133)	-0.282 (0.127)	0.000 (0.003)	-0.004 (0.002)	-0.010 (0.011)	-0.006 (0.014)
Parties	Left	Right	Left	Right	Left	Right	Left	Right
Bandwidth	0.059	0.087	0.038	0.068	0.022	0.048	0.012	0.039
Outcome Mean	60	64.8	3.69	3.66	0.0188	0.0168	0.0648	0.118
N (left)	856	1457	769	1341	657	1166	519	1064
N (right)	861	1463	773	1343	663	1160	530	1068

All estimates are reported with robust standard errors clustered at the municipality level in parentheses. Each observation is a candidate’s election attempt. ‘Elected’ is an indicator for observations where the candidate obtained a seat in the municipal council. ‘Female’ is an indicator for observations identified as female. ‘Elected’ times ‘Female’ is the interaction between the two variables. Other coefficients not reported. Regression run on all candidates in elections between 2003 and 2015.

E.4.3 Robustness to Functional Form

Finally, I also report estimates from specifications with second-order polynomials in Table E6. The results become a lot noisier; the general pattern of results, along with the estimates' sign and magnitude, still holds up for raw and logged personal vote counts. As discussed before, the specifications using share-based measures as an outcome are much more susceptible to noise – especially with more demanding specifications such as second-order polynomials – and should therefore be treated with caution.

Table E6: Sensitivity of Results to Different Personal Vote Measures, With Second-Order Polynomials. The results become noisier, but remain consistent.

	Raw PV		log(Raw PV)		PV / Total Votes		PV / Party Votes	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Elected	6.901 (9.262)	36.748 (7.169)	0.102 (0.107)	0.424 (0.090)	-0.000 (0.003)	0.001 (0.003)	0.011 (0.009)	0.008 (0.012)
Female	-11.687 (10.981)	33.684 (36.674)	-0.126 (0.126)	-0.002 (0.122)	-0.001 (0.003)	-0.000 (0.002)	0.005 (0.009)	-0.006 (0.011)
Elected x Female	31.003 (20.053)	-59.560 (37.342)	0.175 (0.204)	-0.225 (0.182)	0.001 (0.004)	0.000 (0.004)	-0.008 (0.014)	-0.014 (0.018)
Parties	Left	Right	Left	Right	Left	Right	Left	Right
Bandwidth	0.067	0.093	0.05	0.073	0.025	0.052	0.013	0.046
Outcome Mean	65.6	69.1	3.77	3.76	0.0204	0.0175	0.069	0.124
N (left)	586	1019	553	941	446	840	352	791
N (right)	604	1048	564	962	458	856	358	803

All estimates are reported with robust standard errors clustered at the municipality level in parentheses. Each observation is a candidate's election attempt. 'Elected' is an indicator for observations where the candidate obtained a seat in the municipal council. 'Female' is an indicator for observations identified as female. 'Elected' times 'Female' is the interaction between the two variables. Other coefficients not reported. Regression run on all candidates in elections between 2003 and 2015.

E.5 Rank Position and Personal Votes

Below, I plot the average number of personal votes by list rank, grouped by size of the council. We see that pre-vote rank ordering is a strong (though imperfect) predictor of the number of personal votes.

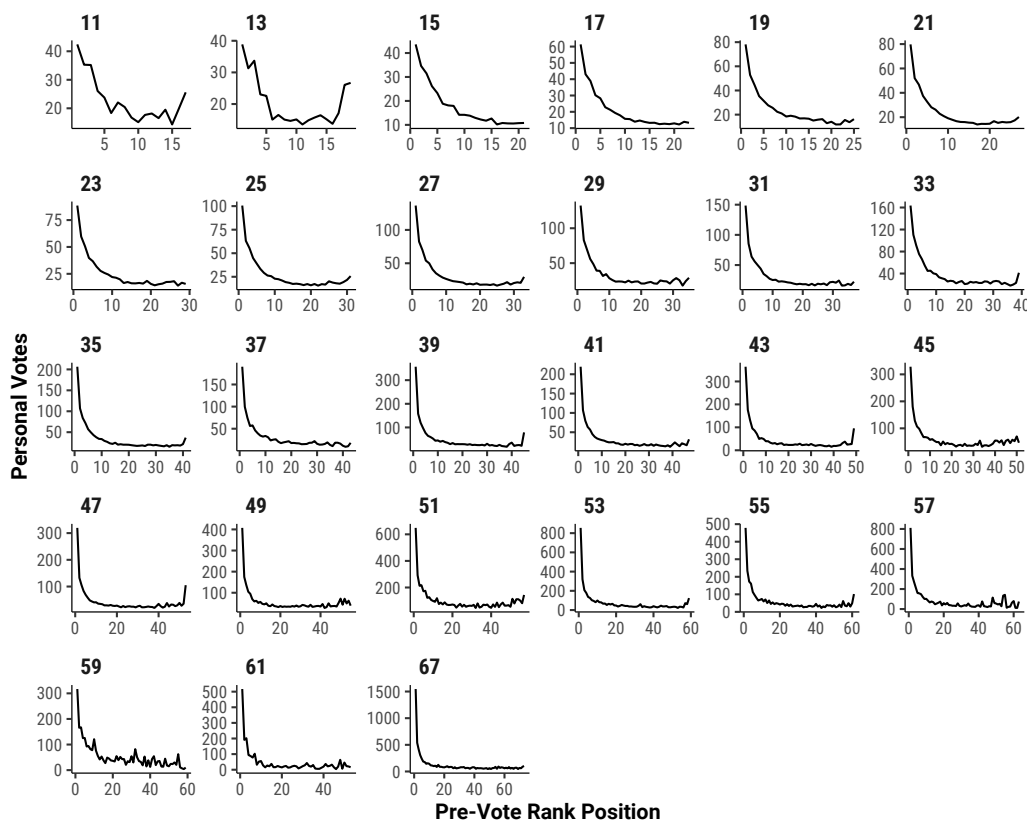


Figure E3: Average number of personal votes, by list rank and council size.

E.6 Differential Selection and Characteristics Correlated With Gender

Is the 'gender gap' specific to this characteristic, or do other candidate characteristics also moderate the incumbency advantage? Using the same difference-in-discontinuities specification as in the main results, I check whether there is a meaningful difference in incumbency effects between candidates below and above the median age. Unfortunately, due to data availability, I can only provide these estimates for Norway.

Table E7: Difference-in-Discontinuity Estimates For Incumbency Advantage In Norwegian Municipalities, By Age Group. No Meaningful Difference Between Younger And Older Candidates.

	Run ($t + 1$)			Win ($t + 1$)		
	(1)	(2)	(3)	(4)	(5)	(6)
Elected	0.007 (0.019)	-0.002 (0.017)	0.013 (0.024)	0.076 (0.018)	0.075 (0.016)	0.064 (0.022)
Age 48+	0.098 (0.020)	0.088 (0.018)	0.108 (0.024)	-0.043 (0.015)	-0.037 (0.014)	-0.032 (0.019)
Elected x Age 48+	-0.012 (0.029)	-0.002 (0.026)	-0.023 (0.036)	0.013 (0.025)	0.018 (0.022)	0.012 (0.030)
Bandwidth	0.054	0.11	0.027	0.05	0.099	0.025
BW Type	Optimal	2x Opt	0.5x Opt	Optimal	2x Opt	0.5x Opt
Outcome Mean	0.564	0.575	0.551	0.261	0.263	0.257
N (left)	4617	5529	3667	4502	5428	3549
N (right)	4666	5590	3711	4551	5489	3592

All estimates are reported with robust standard errors clustered at the municipality level in parentheses. Each observation is a candidate's election attempt. 'Elected' is an indicator for observations where the candidate obtained a seat in the municipal council. 'Age 48+' is an indicator for candidates older than 47 at the time of the election. 'Elected' times 'Female' is the interaction between the two variables. Other coefficients not reported. Regression run on all candidates in elections between 2003 and 2015.

The results offer no evidence of any age-based gap in incumbency advantages.

Next, I also code candidates on whether they have run as candidate before or not. I use this binary 'experience' variable as an additional moderator for my difference-in-discontinuity design. As with age, I find no significant differential effect in the incumbency effect based on experience.

Table E8: Difference-in-Discontinuity Estimates For Incumbency Advantage In Norwegian Municipalities, By Political Experience. No Difference Between First-Time And Multiple-Time Candidates.

	Run (t + 1)			Win (t + 1)		
	(1)	(2)	(3)	(4)	(5)	(6)
Elected	−0.010 (0.019)	−0.001 (0.016)	−0.009 (0.023)	0.083 (0.017)	0.093 (0.015)	0.069 (0.021)
Experience	0.050 (0.020)	0.074 (0.017)	0.044 (0.026)	−0.026 (0.017)	−0.013 (0.015)	−0.025 (0.021)
Elected x Experience	0.027 (0.032)	−0.003 (0.026)	0.032 (0.040)	−0.005 (0.026)	−0.025 (0.022)	0.002 (0.032)
Bandwidth	0.054	0.11	0.027	0.05	0.099	0.025
BW Type	Optimal	2x Opt	0.5x Opt	Optimal	2x Opt	0.5x Opt
Outcome Mean	0.564	0.575	0.551	0.261	0.263	0.257
N (left)	4617	5529	3667	4502	5428	3549
N (right)	4666	5590	3711	4551	5489	3592

All estimates are reported with robust standard errors clustered at the municipality level in parentheses. Each observation is a candidate's election attempt. 'Elected' is an indicator for observations where the candidate obtained a seat in the municipal council. 'Female' is an indicator for observations identified as female. 'Elected' times 'Female' is the interaction between the two variables. Other coefficients not reported. Regression run on all candidates in elections between 2003 and 2015.

E.7 Matching Estimates

In order to further reduce concerns about differential selection, I want to find the closest possible comparison group to female candidates in the threshold sample. To do so, I match observations on a number of dimensions: candidates' age, experience, the year of the election, their party and the province / county in which they run. I describe the matching procedure in greater detail and report the results below in Appendix E.7. Importantly, even in the strictest setting – matching the closest male threshold candidate in terms of age, experience and election year in exactly the same party and county – my estimates report a meaningful gender gap in the incumbency advantage (albeit with somewhat larger uncertainty).

Matching Procedure. For all matching estimates, I retain the set of female threshold candidates within the bandwidth of 0.05 and attempt to find the closest possible male candidate within the same bandwidth that matches this observation. Unlike in the original setting, I cannot rely on automatic bandwidth selection procedures, since the choice of bandwidth also affects the choice of matching pairs (that is, selecting h on the full sample may no longer be MSE-optimal after only retaining matches within h).

Throughout, I match without replacement. First, I use simple distance-based (Mahalanobis) matching to find the closest male candidate in terms of running variable, age, experience, and election year (columns 1 and 2). I then restrict possible matches to those within exactly the same party (column 3) and exactly the same province (column 4). Finally, I pick the male candidate with the closest election margin, age, experience and election year who runs in both the same party as well as the same province (column 5).

Results. The results suggest that even when comparing the incumbency advantage among men and women with similar profiles, and from the same party and geography, there is a notable gender gap, decreasing the incumbency advantage for women by more than 50% on average. That said, it is worth emphasising that the matching estimates in columns (4) and (5) are somewhat reduced in magnitude and no longer statistically significant at the conventional level of .05. While these estimates reduce the concern that differential selection or candidate attributes other than, yet correlated with gender can explain the contrast in incumbency advantages, a remaining caveat is that my list of candidate covariates is finite and non-exhaustive.

Table E9: Matched Estimates.

	Win ($t + 1$)				
	(1)	(2)	(3)	(4)	(5)
Elected	0.106 (0.020)	0.108 (0.021)	0.113 (0.020)	0.081 (0.021)	0.087 (0.022)
Female	-0.032 (0.019)	-0.032 (0.020)	-0.040 (0.019)	-0.045 (0.019)	-0.043 (0.020)
Elected x Female	-0.064 (0.029)	-0.067 (0.030)	-0.073 (0.029)	-0.039 (0.029)	-0.047 (0.031)
Bandwidth	0.05	0.05	0.05	0.05	0.05
N	7128	7128	7068	7128	6804
Individual Match	Y	Y	Y	Y	Y
Year Match	Y		Y	Y	Y
Year (Factor) Match		Y			
Party Match			Y		Y
Province Match				Y	Y

All estimates are reported with robust standard errors clustered at the municipality level in parentheses. 'Individual Match' indicates 1:1 distance-based matching on age and experience; 'Year' match indicates 1:1 distance-based matching on election year; 'Year (Factor)' treats election year as a dummy variable; 'Party Match' and 'Province Match' indicates 1:1 exact matching on party and province, respectively.

E.8 Heterogeneity By Population Size

In this section, I examine whether cities that are outliers in terms of population drive my results. Although the gender gap is concentrated among cities in the upper tercile (consistent with the name recognition mechanism), I find no evidence that the overall result is sensitive to excluding the largest cities.

E.8.1 Sensitivity to Outliers

My estimates hold up even when restricting the sample to towns below a population of 20,000. When restricting the sample to towns smaller than 10,000 population, the effect of winning on winning again moves towards zero, but remains negative. In line with Appendix E.8.2, I interpret this as evidence for name recognition mechanisms in larger municipalities driving the effect.

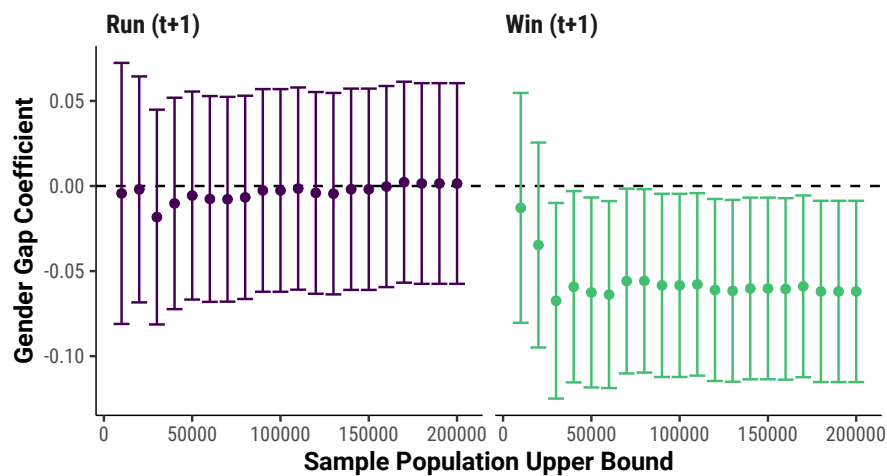


Figure E4: Gender Gap in Incumbency Effect, Norway, By Maximum Municipality Size Included.

E.8.2 Effect Magnitude by Population Tercile

Next I also estimate the gender gap on subsamples binned by population tercile. The results suggest that the negative effect in the Norwegian case is primarily driven by the upper tercile, i.e. the largest municipalities (but, as the previous appendix section suggests, not just its extreme outliers). This finding is consistent with the argument that name recognition is a key driver of personal vote incumbency advantages in Norway: in smaller municipalities (with shorter lists and fewer candidates), voters may know the handful of elected female candidates better.

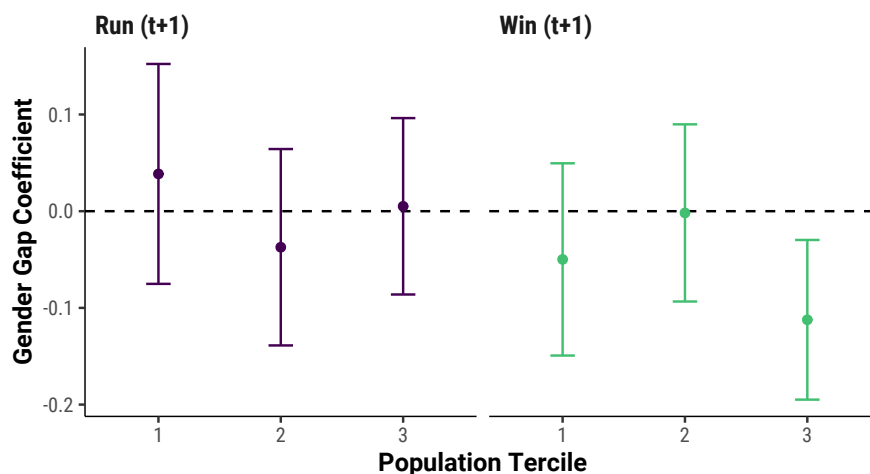


Figure E5: Gender Gap in Incumbency Effect, Norway, By Population Tercile.

E.9 Survey Evidence On Voters' Attitudes Towards Female Candidates

Next, I examine data from the 2015 Local Election Survey in Norway (conducted by the University of Oslo) to test whether right-wing voters hold more negative views about women in politics. I plot the share of respondents from each party agreeing that “Women should play a *greater* role in local [municipal] politics.” The sample analysed included 882 respondents. A majority of respondents in left-leaning parties agrees with this statement, whereas the share of respondents agreeing in right-wing parties hovers around 40%.

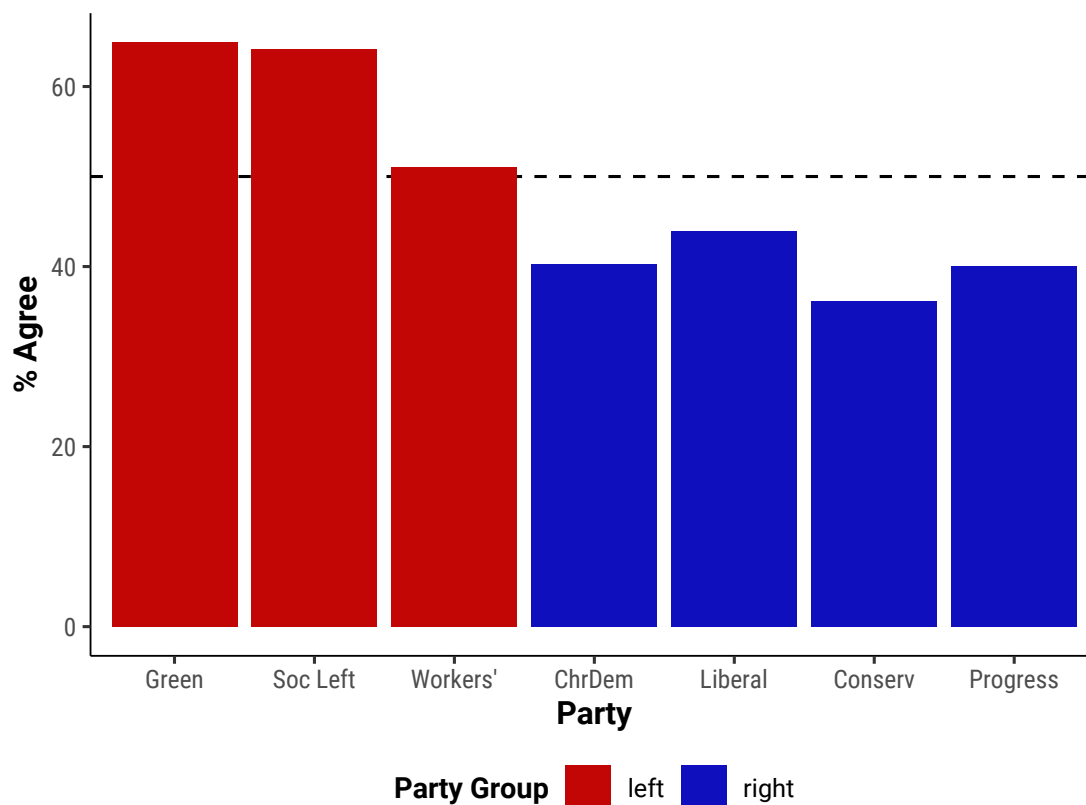


Figure E6: Share Respondents Agreeing That Women Should Play Greater Role In Local Politics, By Party.

F Additional Statistical Results: Spain

F.1 Estimates from Difference-in-Differences Specification

The large number of observations in the Spanish case allows me to compare the outcome mean of candidates extremely close to the threshold without making additional functional form assumptions. To do so, I estimate regressions of the form

$$y_{it} = D_{it} + F_i + D_{it} \times F_i + \phi_{it}(\cdot) + \varepsilon_{it} \quad (5)$$

on all observations within 1 (0.5, 0.25) percentage point(s) of the election margin. D_{it} is a binary indicator whether the candidate i in election t was elected, and F_i is a binary indicator for whether the candidate is recorded as female. As before, I include province and party-by-year fixed effects. The results, in Table F1, yield estimates that are consistent with those obtained using the difference-in-discontinuity approach.

Table F1: . Estimates from Difference-in-Differences in Spanish Close Elections.

	Run Again			Win Again		
	(1)	(2)	(3)	(4)	(5)	(6)
Elected	0.131 (0.014)	0.115 (0.019)	0.104 (0.028)	0.099 (0.012)	0.071 (0.017)	0.094 (0.025)
Female	-0.022 (0.016)	0.002 (0.023)	0.021 (0.033)	-0.016 (0.013)	-0.029 (0.019)	-0.015 (0.026)
Elected x Female	0.031 (0.023)	0.004 (0.032)	-0.003 (0.047)	0.041 (0.020)	0.048 (0.029)	0.027 (0.041)
Bandwidth	0.01	0.005	0.0025	0.01	0.005	0.0025
Mean Out	0.49	0.49	0.49	0.25	0.25	0.25
Avg # votes	26.17	13.4	6.82	26.17	13.4	6.82

Note:

All estimates are reported with robust standard errors clustered at the municipality level in parentheses. Each observation is a candidate's election attempt. 'Elected' is an indicator for observations where the candidate obtained a seat in the municipal council. 'Female' is an indicator for observations identified as female. Bandwidth indicates the range of running variable margins included in the estimation. Avg. # Votes is the average number of votes between the extreme end of the selected bandwidth and the election threshold.

F.2 Heterogeneity Across Parties

In this section, I report the estimates from the main specification on different subgroups of the Spanish data. If the estimates fitted on left- and right-wing parties differ, the overall gender gap estimate in the main part of the manuscript might mask some important heterogeneities.

First, I fit the main specification on samples from left- and right-wing parties (PSOE and PP) separately. Table F.2 reports the results with running and winning in the next election as outcomes.

There is a notable difference in the gender gap in the effect on running again: female candidates in left-wing parties experience a much smaller increase (relative to men) in their probability of running again in left-wing parties, compared to the same difference in right-wing parties. However, the estimate is also somewhat imprecise.

More importantly, we see little difference in the gender gap in the incumbency effect on winning again between the two party samples. This suggests that there is no party-specific mechanism that drives results (and is consistent with an ideology-agnostic seniority norm), and that would be masked by pooling the sample.

Table F2: Difference-in-Discontinuity Estimates For Incumbency Advantage In Spanish Municipalities, By Political Party Group. The estimates for the gender gap are similar in both left- and right-wing parties.

	Run ($t + 1$)		Win ($t + 1$)	
	(1)	(2)	(3)	(4)
Elected	0.091 (0.024)	0.119 (0.023)	0.069 (0.020)	0.087 (0.021)
Female	-0.026 (0.026)	-0.006 (0.027)	-0.010 (0.021)	-0.042 (0.022)
Elected x Female	0.069 (0.036)	0.003 (0.039)	0.042 (0.032)	0.058 (0.034)
Parties	Left	Right	Left	Right
Bandwidth	0.03	0.032	0.032	0.03
Outcome Mean	0.478	0.484	0.259	0.26
N (left)	5695	5422	6017	5162
N (right)	5544	5402	5855	5139

All estimates are reported with robust standard errors clustered at the municipality level in parentheses. Each observation is a candidate's election attempt. 'Elected' is an indicator for observations where the candidate obtained a seat in the municipal council. 'Female' is an indicator for observations identified as female. 'Elected' times 'Female' is the interaction between the two variables. Other coefficients not reported. Regression run on all candidates in elections between 2003 and 2015.

F.3 Heterogeneity Across Quota Rules

Another worry is that estimating the average pooled over all municipalities masks some important heterogeneity across towns. In 2007, a gender quota started to apply to places with a population of above 5,000: lists in municipalities above this threshold must now feature at least 40% of candidates from either gender. The requirement was extended to all municipalities above 3,000 pop. in 2011.

In this section, I exploit the discontinuity in the application of this quota to compare the estimates of incumbency effects in municipalities near the population threshold. This allows me to check whether the lack of a meaningful gender gap in incumbency advantages is driven by either municipalities with or without a quota.

Ideally, to check whether the mandated quota drives these results, I would run a two-dimensional regression discontinuity with margin of victory and population size as independent assignment variables. This would allow me to check whether the estimates in Equation 2 meaningfully change at the threshold. Unfortunately, even with the amount of data from Spanish municipalities, such a design would be underpowered and consequently suffer from high uncertainty and noisy estimates.

Instead, I report results from two types of specifications. First, I fit the typical difference-in-discontinuity specification. The results offer no clear interpretation because of large uncertainty that comes with these estimates. In addition, I also offer estimates from a difference-in-difference strategy. While less clearly identified, these estimates come with lower variance. Together, I find no strong evidence that the legal gender quota in small municipalities is responsible for the lack of a gender gap in political careers in Spain.

F.3.1 Difference-in-Discontinuity Estimates

Following the introduction of the 3,000 pop. threshold in 2011, I compare candidates in municipalities between 250 and 3,000 population to candidates in municipalities between 3,001 and 6,000 population.³⁶ The approach assumes that municipalities within the bandwidth are comparable to one another, and any difference between the samples is driven by the quota alone. That said, the results are also noisier because I have to restrict my sample to post-2007 (i.e., post-quota introduction) observations.

Unsurprisingly, the estimates obtained from the difference-in-discontinuity specification come with a high degree of uncertainty: there simply are not that many observations close to the municipality threshold *and* the electoral threshold. For each of the two outcomes, the estimates of the gender gap change signs across the two municipality samples (although in different directions). At the same time, the confidence intervals for these estimates also firmly includes null and the respective other point estimate. These results therefore offer little evidence that would suggest that the quota plays an important role in attenuating the gender gap.

³⁶ I exclude candidates from municipalities that cross the population threshold in between elections in this sample. The choice of bandwidth in this exercise is such that I have enough power to say something meaningful. Estimates that restrict the bandwidth to a narrower slice to municipalities around the 3,000 threshold end up being extremely noisy.

Table F3: Difference-in-Discontinuity Estimates For Incumbency Advantage In Spanish Municipalities, By Quota Law. The estimates for the gender gap are similar in municipalities with and without the gender quota on lists.

	Run (t+1)		Win (t+1)	
	(1)	(2)	(3)	(4)
Elected	0.127 (0.032)	0.113 (0.068)	0.039 (0.029)	0.160 (0.063)
Female	0.055 (0.039)	-0.060 (0.070)	-0.018 (0.033)	0.003 (0.059)
Elected x Female	-0.062 (0.056)	0.037 (0.102)	0.035 (0.050)	-0.029 (0.095)
Pop	0-3k	3-6k	0-3k	3-6k
Bandwidth	0.03	0.031	0.029	0.029
Outcome Mean	0.475	0.522	0.247	0.272
N (left)	2839	856	2812	801
N (right)	2947	780	2912	739

All estimates are reported with robust standard errors clustered at the municipality level in parentheses. Each observation is a candidate's election attempt. 'Elected' is an indicator for observations where the candidate obtained a seat in the municipal council. 'Female' is an indicator for observations identified as female. 'Elected' times 'Female' is the interaction between the two variables. Other coefficients not reported. 'Pop' indicates the population of a candidate's municipality at time t . Regression run on all candidates in elections between 2011 and 2015. Observations in municipalities that crossed the 3,000 population threshold in between elections are excluded.

F.3.2 Difference-in-Differences Estimates

An alternative approach is to estimate a regression akin to a triple difference-in-differences design using barely losing / barely winning candidates within a selected vote share bandwidth around the threshold of victory. This specification no longer estimates the LATE at the threshold of being elected, but may reduce variance by making fewer functional form assumptions.

In Table F.3, I report the estimates with different bandwidth choices for candidates' margin of victory, keeping the window of municipalities at 1,000 to 5,000. They are estimated somewhat more precisely than the results from the difference-in-discontinuity specification. For both the simple interaction (capturing the gender gap in cities below the threshold), as well as the double interaction (capturing the *additional* gender gap in cities above the threshold), the estimates are close to and statistically indistinguishable from zero. While these results need to be treated with the appropriate caution, I find no convincing evidence that municipalities with a quota have fundamentally different gender gaps from those without a quota.

While these results do not represent clear-cut evidence, they are nonetheless consistent with the argument that elections using closed-list PR exhibit fewer disadvantages for women's political careers, regardless of whether a gender quota is enforced on lists.³⁷ They are also consistent with Bagues and Campa (2021), who find that the introduction of the gender quota in Spanish municipalities has little effect on downstream policy outcomes.

³⁷ One potential explanation for the lack of noticeable differences could be that the Spanish gender quota law is not a particularly strict one: the gender gap might be affected by more stringent requirements such as zipped lists.

Table F4: Difference-in-Difference Estimates For Incumbency Advantage In Spanish Municipalities, By Quota Law. The estimates for the gender gap are similar in municipalities with and without the gender quota on lists.

	Win (t+1)			
	(1)	(2)	(3)	(4)
Elected	0.122 (0.023)	0.148 (0.016)	0.120 (0.023)	0.148 (0.016)
Female	0.012 (0.025)	0.008 (0.016)	0.012 (0.025)	0.008 (0.016)
Quota	-0.015 (0.030)	0.032 (0.023)	-0.105 (0.130)	-0.023 (0.087)
Elected x Female	-0.030 (0.038)	-0.010 (0.026)	-0.028 (0.038)	-0.010 (0.026)
Elected x Quota	0.086 (0.047)	-0.000 (0.035)	0.087 (0.047)	-0.001 (0.035)
Female x Quota	-0.012 (0.046)	-0.042 (0.033)	-0.012 (0.046)	-0.043 (0.033)
Elected x Female x Quota	-0.013 (0.070)	0.029 (0.050)	-0.014 (0.070)	0.029 (0.050)
Pop Window	1-5k	1-5k	1-5k	1-5k
Margin Window	0.025	0.05	0.025	0.05
Linear Pop Trend	N	N	Y	Y
N	3285	6374	3285	6374

All estimates are reported with robust standard errors clustered at the municipality level in parentheses. Each observation is a candidate's election attempt. 'Elected' is an indicator for observations where the candidate obtained a seat in the municipal council. 'Female' is an indicator for observations identified as female. 'Quota' is an indicator for candidates in cities with a population greater than 3,000. Regression run on all candidates in elections between 2011 and 2015. Observations in municipalities that crossed the 3,000 population threshold in between elections are excluded.

F.4 Heterogeneity By Population Size

In the case of Spain, the results remain consistent throughout the range of upper bounds.

F.4.1 Sensitivity to Outliers

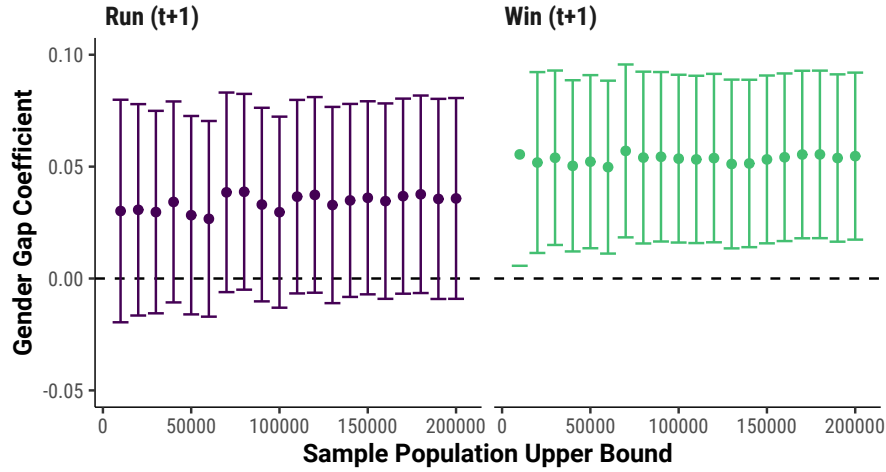


Figure F1: Gender Gap in Incumbency Effect, Spain, By Maximum Municipality Size Included.

F.4.2 Effect Magnitude By Population Tercile

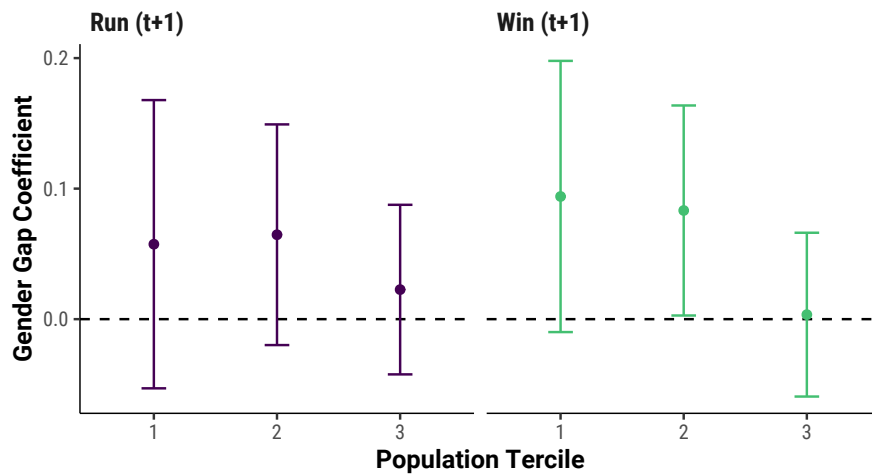


Figure F2: Gender Gap in Incumbency Effect, Spain, By Population Tercile.

G Additional Statistical Results: Poland

In this section, I report additional results from the difference-in-discontinuity specification fitted on open-list PR, county-level elections in Poland. The sample includes all threshold candidates from Poland's major two parties (Civic Platform, Law and Order) at the time of the 2010 election running for election in counties or cities with county powers. These elections are one level up from local elections (where council members are elected using plurality or plurality at-large).

The estimates yield gender gap magnitudes that are similar to those of Norway's open-list PR case, albeit less precise. For running again, the results suggest that men enjoy a 10-12 percentage point increase in their probability of running again upon being elected the first time. Here, the gender gap coefficient is positive but highly insignificant and exhibits meaningful variance across the bandwidth choices.

For winning again, the incumbency advantage for men is around 13 percentage points. The gender gap is negative and suggests that the incumbency advantage diminishes by about 7 percentage points for women.

Table G1: Difference-in-Discontinuity Estimates For Incumbency Advantage In Polish Counties and County-Like Cities. The gender gap is similar in magnitude to that of Norwegian municipalities.

	Run (t + 1)			Win (t + 1)		
	(1)	(2)	(3)	(4)	(5)	(6)
Elected	0.109 (0.032)	0.110 (0.026)	0.122 (0.043)	0.134 (0.030)	0.121 (0.023)	0.157 (0.042)
Female	-0.026 (0.047)	-0.053 (0.037)	-0.045 (0.056)	-0.001 (0.035)	-0.027 (0.027)	0.054 (0.049)
Elected x Female	0.030 (0.065)	0.048 (0.053)	0.001 (0.081)	-0.069 (0.057)	-0.057 (0.044)	-0.116 (0.078)
Bandwidth	0.065	0.13	0.033	0.059	0.12	0.029
BW Type	Optimal	2x Opt	0.5x Opt	Optimal	2x Opt	0.5x Opt
Outcome Mean	0.369	0.374	0.365	0.155	0.163	0.153
N (left)	1227	1681	827	1160	1619	772
N (right)	1221	1673	822	1154	1611	767

All estimates are reported with robust standard errors clustered at the municipality level in parentheses. Each observation is a candidate's election attempt. 'Elected' is an indicator for observations where the candidate obtained a seat in the municipal council. 'Female' is an indicator for observations identified as female. 'Elected' times 'Female' is the interaction between the two variables. Other coefficients not reported. Regression run on all candidates from two main parties (PO, PiS) in county-level elections in 2010.

H Descriptive Statistics

In this section, I report key summary statistics for both cases.

Table H1: Summary Statistics For Borderline Sample, Norway.

	Mean	SD	Min	Max
Share Female	0.39	0.49	0.00	1.00
Elected	0.50	0.50	0.00	1.00
Population	11950.16	20058.80	214.00	244620.00
Total Seats	27.28	10.11	11.00	85.00
Running (t+1)	0.60	0.49	0.00	1.00
Winning (t+1)	0.26	0.44	0.00	1.00

Table H2: Summary Statistics For Borderline Sample, Spain.

	Mean	SD	Min	Max
Share Female	0.38	0.48	0.00	1.00
Elected	0.50	0.50	0.00	1.00
Population	12070.60	26552.11	251.00	248150.00
Total Seats	12.56	5.27	7.00	27.00
Running (t+1)	0.49	0.50	0.00	1.00
Winning (t+1)	0.25	0.44	0.00	1.00