RIETI Discussion Paper Series 25-E-047

Voting for Gender Balancing? The effect of a multiple-vote system on women's representation

ONO, Yoshikuni
RIETI

MIWA, Hirofumi
Gakushuin University

KASUYA, Yuko
Keio University

Research Institute of Economy, Trade & Industry, IAA

The Research Institute of Economy, Trade and Industry
https://www.rieti.go.jp/en/

RIETI Discussion Paper Series 25-E-047
May 2025

Voting for Gender Balancing?
The effect of a multiple-vote system on women's representation[1]

Yoshikuni ONO
Waseda University, Research Institute of Economy, Trade and Industry

Hirofumi MIWA
Gakushuin University

Yuko KASUYA
Keio University

Abstract

Does allowing voters to choose multiple candidates foster diversity in legislative bodies? Majoritarian systems typically restrict voters to casting a ballot for one candidate, yet research suggests that permitting voters to select multiple candidates could boost the election of women and racial minorities. Despite indications of greater diversity under multiple-vote systems, voter behavior evidence remains scarce. To address this, our survey experiment varied the number of selectable candidates on a ballot in local elections. Results revealed that respondents alternated candidate genders, particularly in their second and third choices, supporting the theory that multiple voting promotes diverse representation. However, men more frequently became their first choice when multiple candidates were selectable, giving male candidates an overall advantage on the aggregate level.

Keywords: Voting behavior, Electoral Systems, Gender, Survey experiment
JEL classification: J16, D72, D91

[1] This study was conducted as part of the project titled "Challenges to Achieving a Sustainable Society: Exploring solutions through a social science approach utilizing experiments and data" at RIETI. The draft of this paper was presented at the RIETI DP Seminar for their helpful comments.

1 Introduction

The effects of proportional representation (PR) and quota systems on women's representation in legislative bodies have been extensively studied. In contrast, little attention has been given to majoritarian electoral systems. Typically, in these systems, each voter is limited to choosing only a single candidate. Majoritarian systems do, however, come in a variety of forms, such as ranked-choice voting and alternative vote systems that allow individual voters to choose multiple candidates. Prior studies, such as those by Santucci and Scott (2021), John, Smith and Zack (2018), McGing (2013), suggest that these multiple-vote systems encourage a more diverse range of candidates to run for office. Yet, empirical evidence remains limited in demonstrating how voter behavior in these systems contributes to a more diverse elected body.

To address this gap in the literature, we compare voter behavior in two majoritarian plurality electoral systems that allocate seats based on the number of votes received: the Block Vote (BV) system, which allows each voter to cast votes for multiple candidates and elect more than one candidate per district, and the Single Non-Transferable Vote (SNTV) system, which permits each voter to cast a single vote for one candidate while still electing more than one candidate per district. A key commonality between the two systems is that winners are determined by the number of votes received, filling all seats in a multi-member district in descending order of votes. A notable difference, however, lies in the number of candidates for whom voters can cast their ballots. Our study aims to explore whether the option to cast multiple votes in majoritarian systems influences voters to seek greater representation diversity.

Our research was inspired by Japan's experience, which provides a compelling example of the potential influence of majoritarian electoral systems on the diversity of representation. Japan adopted the BV system for its national-level lower house election in 1946 but transitioned to the SNTV system in subsequent elections, continuing this system until 1993. In the 1946 election conducted under the BV system, women held 8.4% of the seats in the lower house of the Diet—the national parliament of Japan. This proportion was remarkably high, given that the average percentage of female legislators in the lower or single houses of 26 sovereign states at that time was only 3.0%.[1] However, under the SNTV system, the proportion of female representatives

[1] Inter-Parliamentary Union, Women in National Parliaments: 50 Years of History at a

significantly declined, hovering around 2–3% until the 1980s.

The initial surge in the number of female representatives observed in the 1946 election under the BV system, followed by a sharp decline in subsequent years under the SNTV system, suggests that the BV system may have encouraged voters to select a more diverse slate of candidates in a political environment where the party system was still evolving, thereby increasing the number of women elected. Although transitioning from a single-vote to a multiple-vote system in majoritarian multi-member districts (or vice versa) might seem a minor reform, it is crucial not to underestimate such changes, especially when they significantly influence descriptive representation in election outcomes.

To explore this hypothesis, we conducted a preregistered survey experiment involving 5,400 Japanese voters. This experiment included conjoint tasks where we varied the number of candidates that respondents could vote for on their ballots in local elections, a context in which political parties play a limited role. Our research offers a unique perspective to the literature by presenting empirical evidence regarding the impact of majoritarian electoral systems on representation diversity, particularly in terms of gender composition. The findings from our conjoint experiment reveal that respondents were more inclined to choose candidates of the opposite gender as their second and third choices under the BV condition, consistent with our hypothesis. However, these choices did not extend to other candidate attributes such as age and education. Despite this trend, female candidates were overall less likely to be elected under the BV condition compared to the SNTV condition. This outcome stems from male candidates being more frequently selected as the first choice under the BV system. Thus, the increased proportion of women representatives elected in Japan's 1946 election cannot be solely attributed to the BV system. Nevertheless, our findings confirm that, in a majoritarian election system, voters are more likely to support diverse representation when they are given the opportunity to cast multiple votes.

Glance, http://archive.ipu.org/wmn-e/history.htm, accessed May 1, 2023.

2 Theory and Hypotheses: Voting for Gender Balancing

2.1 Block Vote and the Japanese Experience

The design of electoral systems significantly influences women's representation, and various aspects of these systems have been extensively studied. A substantial body of research has focused on the effects of the electoral formula, specifically examining the differences between majoritarian systems and PR systems. There is a general consensus that PR systems more effectively facilitate the election of female candidates compared to majoritarian or plurality systems (Krook 2018; Profeta and Woodhouse 2022). Furthermore, within the scope of PR systems, closed-list systems may be more effective to the election of women than open-list systems (Audinga et al. 2019; Dhima et al. 2021; Golder et al. 2017; Gonzalez-Eiras and Sanz 2021). Another focal point is the gender quota system (Hughes et al. 2019). Among the various institutional designs related to quotas, reserved seat quotas have been found to be more effective than voluntary party quotas in increasing women's representation (Tripp and Kang 2008).

In some majoritarian systems, voters have the option to express preferences for multiple candidates on their ballots. Research has highlighted that ranked-choice voting (Santucci and Scott 2021) and the alternative vote (John, Smith and Zack 2018) can result in the election of more women compared to single-vote plurality systems. However, the specific voter behavior driving this outcome has not been fully explored. The BV is another method that has not received enough attention but has the potential to improve minority representation. BV systems, which allow voters to cast multiple votes and elect candidates by plurality (wherein candidates receiving the most votes are elected until all seats are filled), are currently employed in various countries, including in the upper house of Spain and the unicameral legislature of Gibraltar. Historically, BV systems have also been utilized in sub-national elections across several U.S. states, including Pennsylvania and New Jersey, as well as in some UK municipalities.

A variant of the BV system, known as the limited-vote, was implemented in Japan's lower house election in April 1946. Each electoral district had between four and fourteen seats to be allocated, and each voter was permitted to cast a ballot for up to two to three candidates. Beginning with the 1947 election, however, the SNTV system was adopted, and until the 1993

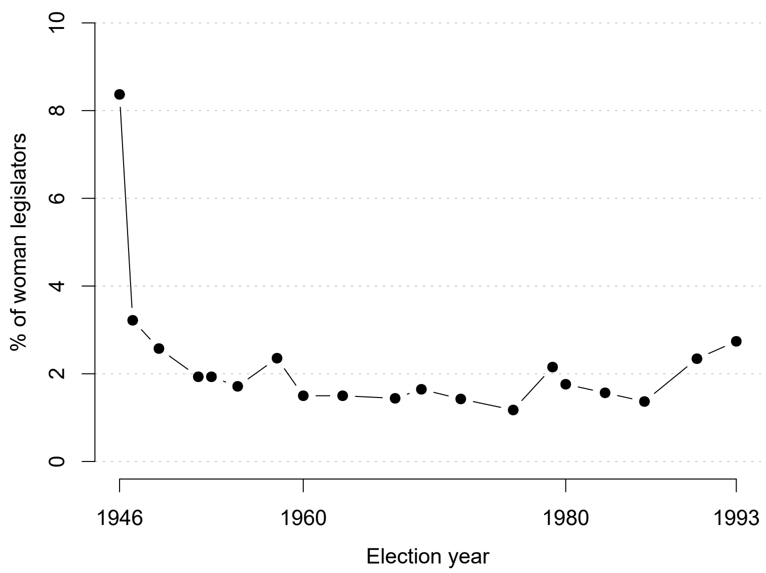

Figure 1: Percentage of Women Legislators in the House of Representatives under the BV and SNTV Systems (1946–1993)

Source: Gender Equality Bureau Cabinet Office (2023)

election, each voter was allowed to vote for no more than one candidate in a three- to five-member district. Figure 1 shows the percentage of female legislators elected to the lower house of the Diet from the 1946 election to the 1993 election. It illustrates that the 1946 election yielded an unprecedented percentage (8.4%) of female legislators compared to the elections held under the SNTV system, where the representation of women in the lower house remained below 3%. In the 1946 election, there were 79 female candidates, and 39 of them were elected. It appears that the BV system facilitated representation diversity, aligning with findings from prior research.

2.2 Gender Balancing Hypotheses

Voters may appear to choose political parties or candidates in elections, but their decisions often revolve around policies from which they stand to benefit or in which they have an interest. However, the final policy outcomes are not solely determined by individual voter preferences; they are the result of negotiations and agreements among political actors. In a presidential system, policies are typically determined through compromises between the president and Congress. In coalition governments, policies are often shaped through agreements among coalition parties. Previous studies indicate that voters who prioritize policy outcomes take into account the anticipated composition of the president and Congress, or the coalition government, and engage in split-ticket voting—choosing different parties for different institutions or tiers in mixed-electoral systems—to achieve a balance and secure desirable policy outcomes (Alesina, Rosenthal et al. 1995; Fiorina 1996; Kedar 2006; Lacy and Paolino 1998; Mebane 2000; Mebane and Sekhon 2002).

We apply this concept of split-ticket voting for policy balancing to explain potential voter behavior regarding representation diversity, particularly gender balancing, under the BV system. In multi-member district plurality systems like the BV and SNTV, the prominence of party names is diminished for individual candidate success. Candidates must compete not only with those from opposing parties but also with their own party members. Under these conditions, electoral campaigns tend to be candidate-centered, prompting voters to focus more on individual candidates rather than on political parties, as party labels do not significantly distinguish candidates from each other (Carey and Shugart 1995).[2] This tendency is particularly pronounced in local politics, where the role of political parties in decision-making is diminished and the attributes of individual candidates become increasingly significant. In such contexts, a candidate's gender can be a salient attribute, serving as a crucial cue for voters in nonpartisan elections (see Badas and Stauffer 2019; Ono and Burden 2019).

Given this context, we expect different patterns of voting behavior between the BV and SNTV systems. In the SNTV system, where voters are allowed to cast only one vote in muti-member districts, vote balancing by ticket-splitting is not feasible. If a significant number of voters exhibit a bias against women, they may predominantly choose male candidates as their preferred choice in the absence of party cues, resulting in the election of more men. In contrast, in the BV system, voters can cast more than one vote, enabling them to split the ticket.

[2] This situation is very different from a single-member district plurality system, a typical form of majoritarian electoral system, where political party names serve as important cues for voters. In such a system, candidates represent their political parties and promote their policy platforms during campaigns (see Catalinac 2016, 2018); thus, electoral campaigns tend to be party-centered.

However, when achieving policy balancing through voting for different parties is challenging, voters can diversify their choices by voting for different candidates across gender—a prominent factor distinguishing candidates.[3] The gender of a candidate can be a crucial indicator for voters, as male and female candidates frequently advocate for different policies, not only during campaigns but also after election (Atkinson and Windett 2019; Bauer and Santia 2022; Evans and Clark 2016; Milita, Ryan and Simas 2014; Ono and Miwa 2023; Schaffner 2005).[4] Thus, we hypothesize that in multi-member districts where multiple votes are allowed, voters are more likely to split their ticket by including candidates of the opposite gender for gender balancing. Specifically, our expectations of split-ticket voting for gender balancing in the BV system lead to the following hypothesis.

H1: *Under the BV system, voters are more likely to alternate between male and female candidates as they make their choices than consistently choosing candidates of the same gender.*

Of course, not all voters necessarily engage in balance-voting as hypothesized in H1. Yet, if voters, collectively, exhibit a preference for candidates of a particular gender as their first choice, we can also expect that the gap in support between male and female candidates in the second and subsequent choices will be smaller as a result of balance-voting. For example, consider a scenario where 60% of voters favor male candidates, and the remaining 40% prefer female candidates as their first choice. If 30% of these voters intend to balance their votes by selecting a candidate of the opposite gender, and this probability is independent of the gender of

[3] Besides gender, there are various signals that differentiate candidates, such as age, education, or occupational backgrounds; Carey and Shugart (1995, 435, fn.14) report journalist anecdotes that, in the 1993 Jordanian election which used the BV with a limited-vote system, voters tended to cast their first vote based on clan, the second vote on constituency service, and the third vote on ideological alignment. In our study, however, we focus specifically on candidate gender.

[4] Beyond policy considerations, voters might also have additional incentives for gender balancing, such as concerns about fairness and equity.

the first-choice candidate, then male and female candidates would receive 54% (= 60%$\times$0.7 + 40%$\times$0.3) and 46% (= 40%$\times$0.7 + 60%$\times$0.3) of the votes for the second-rank choice, resulting in a reduced gender disparity in vote share.[5] This leads to the following hypothesis.

H2: *Under the BV system, the difference in voter support between male and female candidates is smaller for the second and subsequent choices compared to the first choice.*

Our third hypothesis addresses the influence of the number of votes each voter holds in a multi-member district setting. In the SNTV system, voters are limited to casting a single vote, whereas in the BV system, voters have the opportunity to cast multiple votes. If voters cast their votes sincerely, their first choice in the BV system should correspond to their choice in SNTV. If H2 holds true, we can expect different election outcomes between the BV and SNTV systems, particularly in terms of the gender of elected candidates. This expected difference arises from voters' tendency to engage in balance-voting under the BV system. More specifically, we hypothesize that the gap in voter support between male and female candidates will be narrower in the BV system, where gender balancing is anticipated, compared to the SNTV system. This leads to the following hypothesis:

H3: *The difference in voter support between male and female candidates is smaller under the BV system than under the SNTV system.*

[5] It is important to note that this argument does not hold if *all* voters engage in balance-voting and switch the gender of the candidates they support. However, it is reasonable to assume that only a portion of voters will do so.

3 Research Design

3.1 Survey Design

We conducted a pre-registered survey that incorporated a conjoint experiment to test our hypotheses.[6] The survey was carried out from February 19 to 23, 2022, targeting Japanese residents aged 18 to 69 who had voluntarily enrolled in survey panels affiliated with Lucid Marketplace. We employed quotas based on gender, age (five categories), education (three levels), and region of residence (six regions) to match the census population, and successfully recruited 5,400 respondents who passed the attention check questions.[7]

In the survey, we collected demographic information from respondents and assessed their levels of sexism (as detailed below). This was followed by a candidate-choice conjoint experiment, where respondents were presented with six profiles of hypothetical candidates for their municipal assembly.[8] We asked them to select the candidates they would vote for. Each respondent completed the conjoint tasks five times.

The attributes and their respective levels for the conjoint experiment, presented randomly to each respondent, included: gender (man and woman), age (30s, 40s, 50s, 60s, and 70s), educational attainment (high school, vocational college, private university, national university, and graduate school), prior occupation (business employee, business executive, government employee, self-employed, part-time worker, and secretary of a legislator), hometown (outside and inside of the municipality), and experience as an assembly member (no experience, four years, eight years, and twelve years). To reflect the predominantly non-partisan nature of local elections in Japan, party affiliation was not included as an attribute.[9]

---

[6] Our pre-analysis plan is available on AsPredicted (https://aspredicted.org/W8C_J4J).

[7] Unfortunately, we cannot perform robustness analyses including survey participants who failed attention checks, as their participation was immediately terminated before they could proceed to our survey questions. However, since these attention checks were conducted prior to the experiment, excluding such respondents does not induce post-treatment bias.

[8] The exact wording of the survey is provided in Appendix A.

[9] According to the latest statistics, 68% of municipality assembly candidates were indepen-

To ensure that respondents could accurately identify each candidate's gender, we included a pictogram next to each candidate's profile information, depicting either a man's or a woman's face (refer to Figure 2).[10] The profiles used in the survey were designed to approximate the marginal distribution of local assembly members in Japan, enhancing the external validity of the study (de la Cuesta, Egami and Imai 2022). The exception to this was gender, which was distributed uniformly.[11]

The key focus of this experiment lies in the random assignment of respondents into one of the two groups. Respondents in the first group were instructed to select a single candidate, whereas those in the second group were asked to choose three candidates in order of preference. We made it explicit to all respondents that the district magnitude of this hypothetical election was three. In other words, for the first group, we implemented the SNTV system, and for the second group, the BV system was used.[12] Hereafter, we will refer to the first group as the "SNTV group" and the second as the "BV group." Respondents were assigned to these groups in a 1:3

---

dent, and over half of the party-affiliated candidates belonged to two non-major parties in the 2019 local elections (Somusho Jichi Gyosei Kyoku Senkyo Bu 2021).

[10] These pictograms were used under license from Adobe Stock. Some might argue that attaching pictograms could excessively emphasize the candidate's gender, potentially leading to an experimenter demand effect. However, we believe that ensuring respondents' awareness of the candidate's gender is not unreasonable, considering that one limitation of the conjoint design is that respondents might overlook some attributes (Jenke et al. 2021). It is also crucial to replicate real-world conditions, where, unlike other candidate characteristics, gender is almost always apparent to voters through names and campaign materials such as posters and manifestos. Moreover, recent studies indicate that the experimenter demand effect is generally negligible, even when researchers' expectations are disclosed (de Quidt, Haushofer and Roth 2018; Mummolo and Peterson 2019).

[11] Detailed settings of the conjoint experiment, including the exact wording of attributes and levels and the sources of the real-world profile distribution, are available in Appendix B.

[12] Although block voting (and limited voting) in practice does not require voters to rank candidates, we opted for this approach in our experiment to assess H1 and H2.

Figure 2: Screenshot of the conjoint experiment (BV group)

ratio, resulting in 1,456 and 3,944 respondents in the SNTV and BV groups, respectively.[13]

We focused our conjoint experiment on municipal assemblies, which continue to elect their members using the SNTV system.[14] This choice enhances the realism of our hypothetical

[13] We assigned more respondents to the BV group than to the SNTV group because the former was used to test all hypotheses, whereas the latter was specifically for H3.

[14] The SNTV system is also employed in certain prefectural districts during the upper house elections in Japan.

election. Although most municipal assemblies operate with at-large districts, typically having a district magnitude greater than three, there are exceptions like Osaka, Japan's second-largest metropolitan area after Tokyo.

As in other studies based on candidate-choice conjoint experiments, the unit of analysis is each candidate within each task.[15] To test our hypotheses, we employed two types of dependent variables. The first is a binary variable indicating whether each candidate was chosen (referred to as the "choice dummy"). The second, which is exclusively applicable to respondents in the BV group, consists of binary variables indicating whether each candidate was ranked first, second, or third by respondents (referred to as the "1st dummy," "2nd dummy," and "3rd dummy," respectively). Since each respondent completed five tasks, and each task included six candidates, the total number of observations amounts to 162,000. However, for the model estimating the 2nd and 3rd dummies, we excluded data related to higher-ranked candidates.[16] Standard errors were clustered by respondent, and the significance level was set at 5% for all analyses in this study.

To test our hypotheses, following our pre-registered experiment design, we conducted analyses on both the entire sample and two specific subgroups—male respondents and individuals with high scores on a hostile sexism scale. Hostile sexism was measured using a five-battery scale proposed by Schaffner (2022). We calculated respondents' scores by summing their responses, categorizing those with scores above the median as having high levels of hostile sexism. The purpose of analyzing these subgroups is to ascertain whether they also engage in gender-balance-

[15] We confirmed that our conclusions hold even when we analyze the data by choice level rather than candidate level to address dependency between candidates within each task. We discuss the results of the choice-level analyses in Appendix C.6.

[16] Excluding higher-ranked candidates introduced dependencies into the marginal profile distributions. Although applying weighting is necessary for accurate estimation, we chose to estimate the models without weighting due to impracticality of specifying weights for each respondent. To address concerns arising from dependencies in the marginal profile distributions (i.e., candidates' gender and other quality-related attributes), we conducted robustness checks in the process of testing our hypotheses.

voting under the BV system. We expect these subgroups to exhibit a lower preference for female candidates compared to the overall sample. For instance, male respondents may favor male candidates due to gender affinity effects (Badas and Stauffer 2019).

First, we test H1 by using only data from the BV group and estimating the following model for $r = 2, 3$:

$$y_{ijk}^{(r)} = \alpha^{(r)} + \beta^{(r)} x_{ijk} + \gamma^{(r)} v_{ij}^{(r-1)} + \delta^{(r)} x_{ijk} v_{ij}^{(r-1)} + \boldsymbol{\lambda}^{(r)\prime} \boldsymbol{w}_{ijk} + \varepsilon_{ijk}^{(r)}. \quad (1)$$

The superscript (r) indicates that the dependent variable is the r-th dummy. For the k-th candidate for respondent i's j-th task, y_{ijk} is the r-th dummy, x_{ijk} is a dummy variable indicating that the candidate is a woman, and $\boldsymbol{w}_{ijk}$ is a vector of dummy variables for the attribute-levels of the candidate other than his or her gender. The dummy variable $v_{ij}^{(r-1)}$ indicates that respondent i chose a male candidate as the $(r-1)$-th rank in j-th task. α is an intercept, β, γ, δ, and $\boldsymbol{\lambda}$ are coefficients, and ε_{ijk} is an error term. H1 hypothesized that when respondents voted for a male candidate as the $(r-1)$-th rank, female candidates would be more likely to be selected as the r-th rank, and vice versa for a female candidate as the $(r-1)$-th rank. Therefore, we expected $\delta^{(r)}$ to be significantly negative.[17]

Second, we test H2 by using only the data of the BV group and estimating the following model for $r = 1, 2, 3$:

$$y_{ijk}^{(r)} = \alpha^{(r)} + \beta^{(r)} x_{ijk} + \boldsymbol{\lambda}^{(r)\prime} \boldsymbol{w}_{ijk} + \varepsilon_{ijk}^{(r)}. \quad (2)$$

This is a standard linear model used to estimate the average marginal component effects (AMCEs) (Hainmueller, Hopkins and Yamamoto 2014, Proposition 3-2), with $\beta^{(r)}$ representing the AMCE of being a woman in the r-th rank choice. Our focus is on the difference between ranks in the difference in voter support between male and female candidates, which is represented by

[17] Although we pre-registered that we would conduct the F-test to compare Model (1) and the model excluding $\delta^{(r)} x_{ijk} v_{ij}^{(r-1)}$, we decided to use the t-test of δ instead. This is because it is easier to consider clustering by respondent in the t-test than in the F-test. We confirmed that the conclusion remained consistent when conducting the F-test; however, the t-test is a more conservative method.

the absolute value of $\beta^{(r)}$. Although we had no specific expectation regarding whether female or male candidates would be more preferred (and thus, we had no expectation for the sign of β), H2 posited that the gender difference would be smaller in the second and third choices than in the first choice. Thus, we expected $|\beta^{(1)}| > |\beta^{(2)}|$ and $|\beta^{(1)}| > |\beta^{(3)}|$. Since the standard errors of these differences cannot be analytically computed, we conducted 1,000 bootstrapping iterations clustered by respondent and obtained the bootstrapped sampling distribution of the differences to test H2.

Third, we test H3 by estimating the following linear model, using the pooled data of the SNTV and BV groups:

$$y_{ijk} = \alpha + \beta x_{ijk} + \gamma z_i + \delta x_{ijk} z_i + \boldsymbol{\lambda}' \boldsymbol{w}_{ijk} + \varepsilon_{ijk}. \tag{3}$$

y_{ijk} is the choice dummy, and z_i is a dummy variable that takes one if respondent i was assigned to the BV group. Note that, though the meaning of y_{ijk} is not strictly the same between the two groups (one in six candidates was selected in the SNTV group, while three in six were selected in the BV group), it is taken into account by including the term γz_i. In addition to Model (3), we estimate the standard linear model to calculate the AMCEs using the SNTV and BV groups data separately to visualize the estimation results. As with the analysis of H2, we had no expectation about the sign of β. However, H3 stated that the gender difference in respondents' preferences would be smaller in the BV group (represented by $\beta + \delta$) than in the SNTV group (represented by β). Therefore, we expected that $\beta + \delta$ would be closer to zero than β and that δ would be statistically significant.[18]

Note that the differences in the effect of candidate gender on voting preferences between the SNTV and the BV systems, as well as across different ranks within the BV, may stem from the selectability of female and male candidates. For instance, in the SNTV condition, if a respondent strongly favors male candidates, they can select a male candidate unless all six

[18] For the same reason explained in footnote 17, we conducted the t-test of the coefficient rather than the F-test for model comparison, though we confirmed that the F-test gave as the same conclusion.

displayed candidates are women (with a probability of 0.016). However, in the BV condition, respondents are compelled to choose at least one woman if fewer than three of the six displayed candidates are men (a scenario occurring with a probability of 0.344). Furthermore, while respondents can unconditionally choose a male candidate as their first choice in the BV, they might need to opt for a female candidate as their second or subsequent choice due to the limited number of male candidates available. To ensure the robustness of our findings and to confirm that they are not skewed by these selectability issues, we will present the results of robustness checks for each hypothesis after presenting our main findings.

4 Results

4.1 Main Results

We begin by presenting the results for H1. The top panel of Figure 3 shows the AMCEs of being a woman as the second-rank choice, conditioned on the gender of the candidate selected as the first-rank choice by respondents in the BV group. The bottom panel shows the AMCEs for the third-rank choice, conditional on the second-rank choice. In both panels, the dots represent the point estimates of the AMCE for candidate gender, with segments indicating the associated 95% confidence intervals. A positive value suggests a preference for female candidates over male candidates.

The figure illustrates that the AMCE of candidate gender for the second-rank choice does not vary depending on the previous rank choice. However, this pattern shifts for the third-rank choice. Respondents who selected a man as their second-rank choice were 1.6 percentage points more likely to choose a woman as their third-rank choice. In contrast, those who opted for a woman as their second-rank choice were 1.3 percentage points less likely to select another female candidate as their third-rank choice. This difference in the AMCEs, denoted as $\delta^{(3)}$ in Model (1), was estimated to be -0.029 and was statistically significant. On the other hand, the point estimate for $\delta^{(2)}$ was -0.003 and did not reach statistical significance. Therefore, the results regarding the third-rank choice corroborate H1, while those regarding the second-rank choice do not.

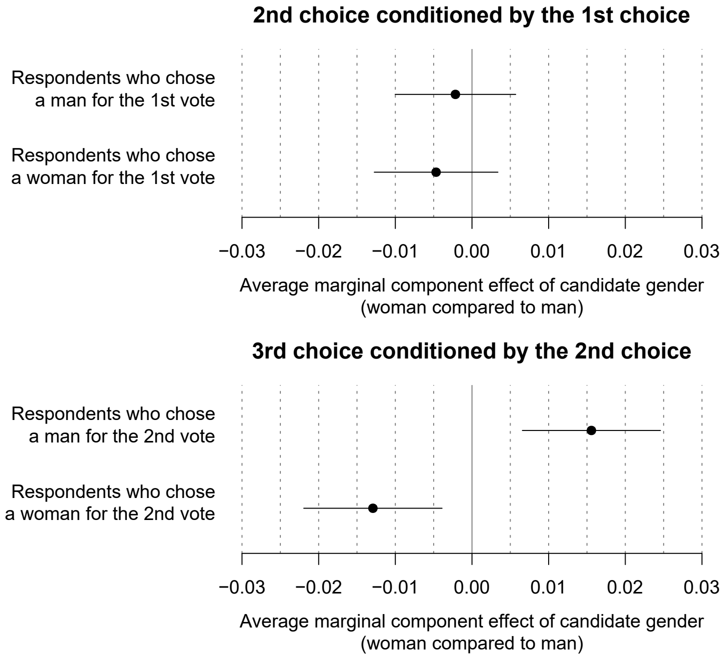

Figure 3: AMCE of candidate gender in second- and third-rank choices in the BV group depending on the gender of the candidate chosen as the previous rank

Note: Dots represent point estimates of the AMCEs for candidate gender (woman compared to man), and segments indicate the associated 95% confidence intervals.

To illustrate the results for H1 more intuitively, we have arranged the pattern of gender preferences for candidates chosen by respondents in the BV group in Table 1. The data clearly shows that, across all combinations of first- and second-rank choices, instances where the genders of the second- and third-rank choice candidates differed were consistently more frequent than their same-gender counterparts. For example, the sequence of choosing a male candidate followed by another male candidate, and then a female candidate (man-man-woman) was 2.3 percentage points more likely than the choice of three male candidates in a row (man-man-man). Similarly, the sequence of selecting a female candidate, then a male candidate, and another female candidate (woman-man-woman) occurred 1.1 percentage points more frequently than the sequence of choosing a female candidate, a male candidate, and then another male candidate (woman-man-man).

Secondly, Figure 4 presents the results for testing H2 displaying the AMCE of candidate gender for each rank of choice within the BV group. A notable distinction is evident between the first-rank choice and the subsequent second- and third-rank choices. Female candidates face

Table 1: Pattern in the gender order of the candidates chosen by the BV respondents

Candidate gender			Prob. (%)
1st	2nd	3rd	
Man	Man	Man	11.6
Man	Man	Woman	13.9
Man	Woman	Man	13.9
Man	Woman	Woman	12.8
Woman	Man	Man	11.6
Woman	Man	Woman	12.7
Woman	Woman	Man	12.1
Woman	Woman	Woman	11.5

a statistically significant disadvantage of 1.6 percentage points compared to male candidates in the first-rank choice. However, this disadvantage is not observed in the second- and third-rank choices (-0.3 and 0.1 percentage points, respectively). A bootstrapping analysis confirms that these differences in the AMCEs are not simply due to random chance: the 95% bootstrap percentile confidence interval of $|\beta^{(1)}| - |\beta^{(2)}|$ in Model (2) is [0.005, 0.018] (hereafter, values in brackets represent a 95% confidence interval), and that of $|\beta^{(1)}| - |\beta^{(3)}|$ is [0.005, 0.019]. These results provide support for H2.[19]

Finally, Figure 5 shows the AMCEs of candidate gender for both the SNTV and the BV groups.[20] The point estimate indicates that female candidates had a marginal advantage of 0.5 percentage points over male candidates in the SNTV group; however, this difference was not statistically significant. Conversely, in the BV group, female candidates were significantly less likely to receive votes by a margin of 0.9 percentage points. When Model (3) was estimated, δ was calculated to be -0.013 and reached statistical significance. These findings are contrary

[19] Focusing on other attributes, some attribute-levels exhibit slight variations in their effects depending on the rank of choices. Nevertheless, except for the gender attribute, all attributes retain their importance up to the third-rank choice. In other words, it is only the gender attribute where the observed effect in the first-rank choice dissipates in the second- and third-rank choices. We provide detailed estimations of marginal means for all attributes across every-rank choice in Appendix C.1.

[20] Estimates for other attributes' AMCE are provided in Appendix C.2.

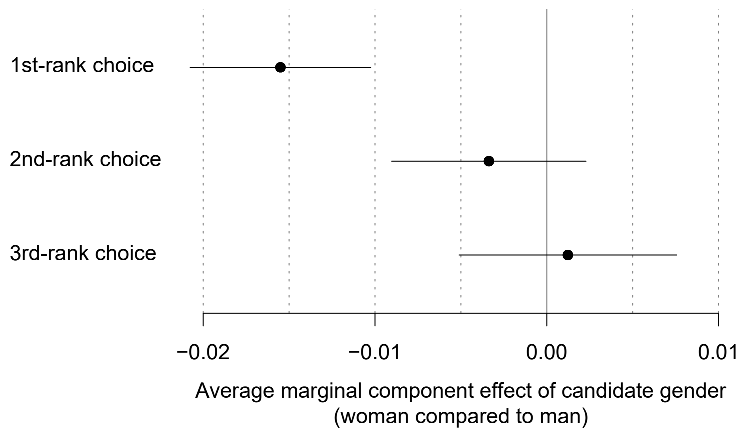

Figure 4: The AMCE of candidate gender in the first-, second-, and third-rank choices in the BV group

Note: Dots represent point estimates of the AMCEs for candidate gender (woman compared to man), and segments indicate the associated 95% confidence intervals.

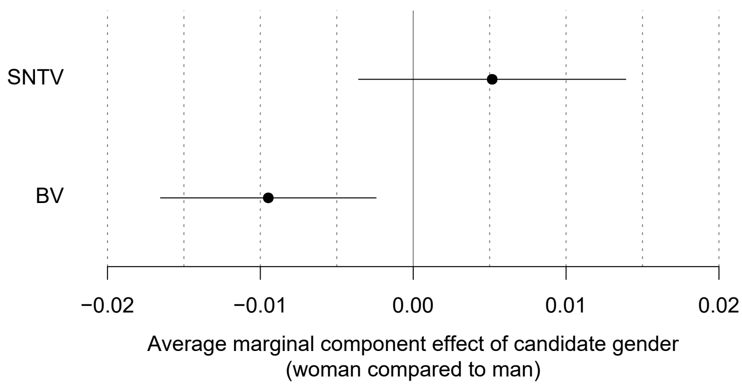

Figure 5: AMCE for candidate gender depending on electoral systems

Note: Dots represent point estimates of the AMCEs for candidate gender (woman compared to man), and segments indicate the associated 95% confidence intervals.

to what we had expected in H3; the difference was actually greater in the BV group than in the SNTV group.

The results of gender-balancing voting remained consistent even when we narrowed our analyses to male respondents and those scoring above the median on a hostile sexism scale.[21] These groups generally exhibited a preference for male candidates, with this trend being more pronounced in the BV group, although the interaction term in Model (3) did not reach statistical significance for male respondents. As hypothesized in H1, in their third-rank choice, male

[21] Detailed results can be found in Appendices C.3 and C.4.

respondents were clearly more likely to select candidates of a different gender compared to their second-rank choice. Similarly, respondents with sexist attitudes, even if they did not have a preference for female candidates, exhibited less avoidance of female candidates when choosing a male candidate as their second preference. In alignment with H2, while male respondents and those with sexist tendencies showed a strong preference for male candidates as their first-rank choice, this bias lessened in their subsequent choices.

In summary, the results of our conjoint experiment indicate that under the BV condition, respondents were more inclined to choose candidates of the opposite gender for their second and third choices, thereby supporting our hypotheses. However, female candidates had lower overall chances of being elected under the BV system compared to the SNTV system. This outcome was due to the tendency of respondents to more frequently select male candidates as their first choice, leading to a higher cumulative support for male candidates.

4.2 Robustness checks

We further conducted several robustness checks to confirm that the results we presented are not merely artifacts of the selectability of female and male candidates on the list in our conjoint experiment.[22] First, regarding H1, to rule out the possibility that our results were influenced by the fact that selecting a male (or a female) candidate limits the availability of same-gender candidates in subsequent choices, we estimated an additional model while accounting for the selectability of male and female candidates for $r = 3$:

$$y_{ijk}^{(3)} = \alpha + \beta x_{ijk} + \gamma_1 v_{ij}^{(2)} + \gamma_2 s_{ij}^{(3)} + \gamma_3 v_{ij}^{(1)} + \delta_1 x_{ijk} v_i^{(2)} + \delta_2 x_{ijk} s_{ij}^{(3)} + \delta_3 x_{ijk} v_i^{(1)} + \boldsymbol{\lambda}' \boldsymbol{w}_{ijk} + \varepsilon_{ijk}. \quad (4)$$

We incorporated two new variables, $s_{ij}^{(3)}$ and $v_{ij}^{(1)}$, into Model (1). The variable $s_{ij}^{(r)}$ denotes the

[22] Robustness checks for H2 and H3 were conducted as per our pre-analysis plan. Although robustness checks for H1 were also preregistered, we later realized that the registered methods were inadequate for addressing concerns about selectability. Consequently, we developed a non-preregistered robustness check ex-post for H1. The detailed results of these checks are available in Appendix C.5.

number of remaining female candidates when respondent i made their r-th rank choice during the j-th task, directly reflecting the selectability of male and female candidates. We also controlled for $v_{ij}^{(1)}$, a dummy variable indicating the choice of a male candidate as the first-rank choice. This modification accounts for the fact that the more male (or female) candidates a respondent had already selected before the third-rank choice, the fewer high-quality male (or female) candidates, considering other attributes, would be available for selection. To minimize the influence of these variables on our primary parameter of interest, δ_1, we introduced interaction terms between these variables and a dummy variable representing the gender of the candidate.

The estimation results of Model (4) confirmed that our conclusion for H1 remained valid even after incorporating these control variables. The AMCE of candidate gender for the third-rank choice was significantly influenced by the gender of the candidate chosen as the second-rank (the point estimate of δ_1 was -0.029, and the 95% confidence interval was $[-0.042, -0.017]$). We present simulated AMCEs for scenarios where two men and two women were available for the third-rank choice in Figure 6 (i.e., we set $s_{ij}^{(3)} = 2$). The top panel illustrates the scenario where respondents selected a male candidate as their first-rank preference (i.e., $v_{ij}^{(1)} = 0$), while the bottom panel depicts the reverse situation. Consistent with Figure 3, a clear pattern emerges: respondents tended to choose a woman if they had selected a male candidate in their previous choice, and vice versa if their prior choice was a female candidate.

Second, our test for H2 could also be subject to the criticism that respondents had a wider range of options for their first-rank choice compared to the subsequent choices. This potential bias might mechanically diminish the effect of candidate gender in the second-rank and third-rank choices. To mitigate this concern, we reanalyzed the data, restricting our observations in the BV group to tasks where the number of male and female profiles were equal. In these scenarios, respondents with a strong preference for male (female) candidates had the opportunity to allocate all their votes to male (female) candidates. We found that although the confidence interval of $\beta^{(1)} - \beta^{(2)}$ included zero ($[-0.022, 0.003]$), the confidence interval of $\beta^{(1)} - \beta^{(3)}$ did not include zero ($[-0.030, -0.002]$). This result reinforces our finding that the disadvantage of female candidates was less pronounced in the third-rank choice compared to the first-rank choice.

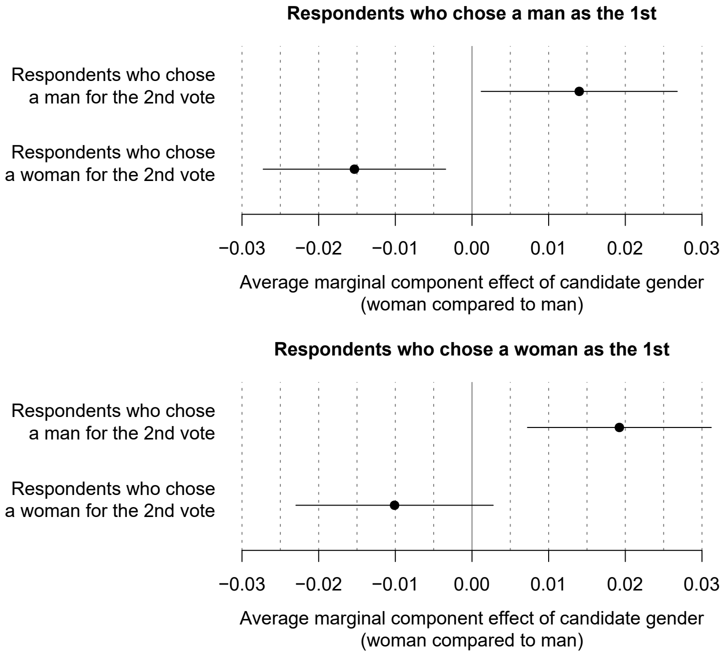

Figure 6: AMCE of candidate gender in the third-rank choice within the BV group, depending on the gender of the candidate chosen as the previous rank, when the number of selectable women is two

Note: Dots represent point estimates of the AMCEs, and segments indicate the associated 95% confidence intervals.

Third, regarding the test of H3, one might argue that comparing the effect of candidate gender between the SNTV and BV groups is not equitable, as respondents in the BV group, who must choose three candidates, might feel more compelled to vote against their gender preferences than those in the SNTV group. To address this issue, we performed two additional analyses. We firstly implemented a robustness check for H3 analogous to the one we performed for H2. In this analysis, we replicated the main analysis using data comprising observations from tasks with an equal number of men and women profiles. The results showed that the disadvantage of being a woman remained significant in the BV group, and the difference in the AMCEs between the groups was significant (-0.021 [-0.037, -0.006]). Secondly, we analyzed the data using the "1st dummy" as the dependent variable for the BV group, instead of the "choice dummy." The "1st dummy" indicates whether a candidate was chosen as the first-ranked preference. As illustrated in Figure 4, the AMCE of being a woman was more negative for the first-rank choice compared to the entire BV process. Moreover, the interaction term in Model (1) was estimated to be significantly negative (-0.021 [-0.031, -0.011]), suggesting that respondents exhibited stronger gender bias in their initial choices than in subsequent choices.

5 Conclusion

Extensive research has examined the role of PR and quota systems in increasing the number of women representatives in legislative bodies. However, less is known about how different majoritarian electoral systems might affect women's representation, particularly given that these systems vary in the number of votes each voter can cast. Previous studies suggest that when voters are allowed to cast votes for more than one candidate, they may attempt to diversify their choices. Yet, the voter behavior underpinning this argument remains empirically untested. To address this gap, we conducted a preregistered survey experiment aimed at exploring differences in voting behavior in majoritarian multi-member districts. Specifically, we compared the BV system, which permits voters to cast multiple votes, with the SNTV system, where voters are restricted to casting only a single vote.

We hypothesized that under the BV system, voters would be incentivized to "split" their tickets by selecting candidates of different genders, especially in nonpartisan contexts such as local elections. To test this gender-balancing voting hypothesis, we conducted a conjoint experiment in which respondents were randomly assigned to one of two municipal assembly election scenarios: the SNTV condition, where they could choose only one candidate from a set of six randomly presented candidates, or the BV condition, where they could select three candidates from the same set.

Consistent with our hypothesis, we found that respondents in the BV condition tended to choose a candidate of a different gender in their second and third choices. This pattern indicates a trend toward achieving gender balance by splitting votes across their selections. Furthermore, even among male and sexist respondents, who typically are less likely to select female candidates, there was a noticeable trend under the BV condition to seek gender balance in their second and third choices. Our robustness checks confirmed that this tendency was not merely a result of respondents being mechanically constrained from choosing a male (or female) candidate, even if they wished to do so.

Contrary to our hypothesis, however, female candidates in the BV condition were at an overall disadvantage compared to those in the SNTV condition. This was mainly due to the first-rank choice in the BV condition, which demonstrated a more pronounced preference for male candidates than the selections in the SNTV condition. This preference prevailed even though the candidate selected in the SNTV condition and the first-rank candidate selected in the BV condition should theoretically hold equal significance for voters. This finding suggests that voters are more inclined to favor male candidates when afforded the opportunity to choose more than one candidate. While further research is needed to see if these results generalize to other countries, they appear to contradict the conclusions of a recent meta-analysis of candidate-choice experiments by Schwarz and Coppock (2022), which found a general preference for female candidates over male candidates. This discrepancy may stem from the fact that the meta-analysis assumed scenarios where respondents choose only one candidate.

Our experimental results, supporting the existence of gender-balancing voting under the BV system, seem to align with the phenomenon observed in Japan's 1946 election, which was marked by a significant number of female winners in the absence of a solid party system. Nevertheless, our study found that voters under the BV system showed an overall bias in favor of male candidates compared to those in the SNTV system. Therefore, the unique outcomes of the 1946 Japanese election are likely attributable to other factors, such as the novelty of female candidates in the first elections after the enactment of women's suffrage.

The implications of our findings also extend beyond explaining the outcomes of Japan's 1946 election. In various countries, multiple-vote BV systems are used not only in political elections, like local contests, but also in situations where management board members are selected. Our results suggest that in these contexts, voters are incentivized to cast gender-balanced ballots, potentially influencing the composition of the elected bodies. Importantly, while our experimental results indicate that other candidate attributes—such as age and education—exert a greater influence on respondents' voting choices than candidate gender, no balance-voting behavior was observed for these attributes. Our findings reveal that balance-voting within the BV system is distinctly significant in terms of candidate gender and does not extend to other attributes such as age, educational attainment, hometown, and political experience (detailed results of these analyses are presented in Appendix C.7).

Finally, our study highlights significant distinctions in the influence of majoritarian electoral systems on women's representation in legislative bodies, a topic that has garnered limited attention in existing research. We observed that variations in voting patterns — specifically, whether voters can choose only one candidate or multiple candidates in a multi-member district — affect the election of women. However, our conjoint experiment did not incorporate party labels as an attribute of the candidates. The presence of party labels might diminish the effect of candidate gender in elections. Moreover, voters may not always cast their votes for their most preferred candidate due to strategic voting, influenced by factors such as candidates' likelihood of winning, which could significantly alter election outcomes (see Cox 1997). Future research should take into account these complexities of majoritarian systems when exploring the role of candidate gender in voting behavior within such contexts.

References

Alesina, Alberto, Howard Rosenthal et al. 1995. *Partisan Politics, Divided Government, and the Economy*. Cambridge University Press.

Atkinson, Mary Layton and Jason Harold Windett. 2019. "Gender Stereotypes and the Policy Priorities of Women in Congress." *Political Behavior* 41: 769–789.

Audinga, Baltrunaite, Alessandra Casarico, Paola Profeta, Giulia Savio et al. 2019. "Let the Voters Choose Women." *Journal of Public Economics* pp. 104085–104085.

Badas, Alex and Katelyn E Stauffer. 2019. "Voting for Women in Nonpartisan and Partisan Elections." *Electoral Studies* 57: 245–255.

Bauer, Nichole M and Martina Santia. 2022. "Going Feminine: Identifying How and When Female Candidates Emphasize Feminine and Masculine Traits on the Campaign Trail." *Political Research Quarterly* 75(3): 691–705.

Carey, John M. and Matthew Soberg Shugart. 1995. "Incentives to Cultivate a Personal Vote: A Rank Ordering of Electoral Formulas." *Electoral Studies* 14(4): 417–439.

Catalinac, Amy. 2016. "From Pork to Policy: The Rise of Programmatic Campaigning in Japanese Elections." *The Journal of Politics* 78(1): 1–18.

Catalinac, Amy. 2018. "Positioning Under Alternative Electoral Systems: Evidence from Japanese Candidate Election Manifestos." *American Political Science Review* 112(1): 31–48.

Cox, Gary W. 1997. *Making Votes Count: Strategic Coordination in the World's Electoral Systems*. Cambridge University Press.

de la Cuesta, Brandon, Naoki Egami and Kosuke Imai. 2022. "Improving the External Validity of Conjoint Analysis: The Essential Role of Profile Distribution." *Political Analysis* 30(1): 19–45.

de Quidt, Jonathan, Johannes Haushofer and Christopher Roth. 2018. "Measuring and Bounding Experimenter Demand." *American Economic Review* 108(11): 3266–3302.

Dhima, Kostanca, Sona N Golder, Laura B Stephenson and Karine Van der Straeten. 2021. "Permissive Electoral Systems and Descriptive Representation." *Electoral Studies* 73: 102381.

Evans, Heather K and Jennifer Hayes Clark. 2016. ""You Tweet Like a Girl!": How Female Candidates Campaign on Twitter." *American Politics Research* 44(2): 326–352.

Fiorina, Morris P. 1996. *Divided Government*. Allyn & Bacon.

Gender Equality Bureau Cabinet Office. 2023. "Reiwa 4 Nendo Josei no Seisaku Hoshin Kettei Sankaku Jokyo Shirabe [Survey of Women's Participation in Policy Making in Fiscal Year 2022]." https://www.gender.go.jp/research/kenkyu/sankakujokyo/2022/index.html, accessed on December 20, 2023.

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.

Gonzalez-Eiras, Martin and Carlos Sanz. 2021. "Women's Representation in Politics: The Effect of Electoral Systems." *Journal of Public Economics* 198: 104399.

Hainmueller, Jens, Daniel J. Hopkins and Teppei Yamamoto. 2014. "Causal Inference in Conjoint Analysis: Understanding Multidimensional Choices via Stated Preference Experiments." *Political Analysis* 22(1): 1–30.

Hughes, Melanie M, Pamela Paxton, Amanda B Clayton and Pär Zetterberg. 2019. "Global Gender Quota Adoption, Implementation, and Reform." *Comparative Politics* 51(2): 219–238.

Jenke, Libby, Kirk Bansak, Jens Hainmueller and Dominik Hangartner. 2021. "Using Eye-Tracking to Understand Decision-Making in Conjoint Experiments." *Political Analysis* 29(1): 75–101.

John, Sarah, Haley Smith and Elizabeth Zack. 2018. "The Alternative Vote: Do Changes in Single-Member Voting Systems Affect Descriptive Representation of Women and Minorities?" *Electoral Studies* 54: 90–102.

Kedar, Orit. 2006. "How Voters Work Around Institutions: Policy Balancing in Staggered Elections." *Electoral Studies* 25(3): 509–527.

Krook, Mona Lena. 2018. "Electoral Systems and Women's Representation." *The Oxford Handbook of Electoral Systems* 175.

Lacy, Dean and Philip Paolino. 1998. "Downsian Voting and the Separation of Powers." *American Journal of Political Science* 42(4): 1180–1199.

McGing, Claire. 2013. "The Single Transferable Vote and Women's Representation in Ireland." *Irish Political Studies* 28(3): 322–340.

Mebane, Walter R. 2000. "Coordination, Moderation, and Institutional Balancing in American Presidential and House Elections." *American Political Science Review* 94(1): 37–57.

Mebane, Walter R and Jasjeet S Sekhon. 2002. "Coordination and Policy Moderation at Midterm." *American Political Science Review* 96(1): 141–157.

Milita, Kerri, John Barry Ryan and Elizabeth N Simas. 2014. "Nothing to Hide, Nowhere to Run, or Nothing to Lose: Candidate Position-Taking in Congressional Elections." *Political Behavior* 36: 427–449.

Mummolo, Jonathan and Erik Peterson. 2019. "Demand Effects in Survey Experiments: An Empirical Assessment." *American Political Science Review* 113(2): 517–529.

Ono, Yoshikuni and Barry C. Burden. 2019. "The Contingent Effects of Candidate Sex on Voter Choice." *Political Behavior* 41(3): 583–607.

Ono, Yoshikuni and Hirofumi Miwa. 2023. "Gender Differences in Campaigning Under Alternative Voting Systems: Analysis of Election Manifestos." *Politics, Groups, and Identities* 11(5): 1203–1211.

Profeta, Paola and Eleanor F Woodhouse. 2022. "Electoral Rules, Women's Representation and the Qualification of Politicians." *Comparative Political Studies* 55(9): 1471–1500.

Santucci, Jack and Jamil Scott. 2021. "Do Ranked Ballots Stimulate Candidate Entry?" *Available at SSRN 3956554* .

Schaffner, Brian F. 2005. "Priming Gender: Campaigning on Women's Issues in US Senate Elections." *American Journal of Political Science* 49(4): 803–817.

Schaffner, Brian F. 2022. "Optimizing the Measurement of Sexism in Political Surveys." *Political Analysis* 30(3): 364–380.

Schwarz, Susanne and Alexander Coppock. 2022. "What Have We Learned About Gender from Candidate Choice Experiments? A Meta-Analysis of Sixty-Seven Factorial Survey Experiments." *The Journal of Politics* 84(2): 655–668.

Somusho Jichi Gyosei Kyoku Senkyo Bu. 2021. "Heisei 31 Nen 4 Gatsu Shikko Chiho Senkyo Kekka Shirabe [Outcomes of Local Elections in April 2019]." https://www.soumu.go.jp/main_content/000786840.pdf, accessed on May 2, 2023.

Tripp, Aili Mari and Alice Kang. 2008. "The Global Impact of Quotas: On the Fast Track to Increased Female Legislative Representation." *Comparative Political Studies* 41(3): 338–361.