PA

LETTER

Analyzing Ballot Order Effects When Voters Rank Candidates

Yuki Atsusaka

Assistant Professor, Hobby School of Public Affairs, University of Houston, Houston, TX 77204, USA.
URL: https://atsusaka.org

Corresponding author: Yuki Atsusaka; Email: atsusaka@uh.edu

Abstract

How does candidate order on the ballot affect voting behavior when voters rank candidates? I extend the analysis of ballot order effects to electoral systems with ordinal ballots, where voters rank candidates, including ranked-choice voting (RCV). First, I discuss two types of ballot order effects, including "position effects"—voters vote for specific candidates because of their ballot positions—and "pattern ranking"—voters rank candidates geometrically given their grid-style ballots. Next, I discuss experimental designs for identifying and estimating these effects based on ballot order randomization. Moreover, I illustrate the proposed methods by using survey and natural experiments based on mayoral and congressional RCV elections in 2022. I find that while voters seem less susceptible to specific ballot positions, ballot design can still impact voters' ranking behavior via pattern ranking. This work has several implications for ballot design, survey research, and ranking data analysis. First, it shows that pattern ranking may affect electoral outcomes in RCV and other systems even when ballot order is *fully randomized*. Consequently, it may be worth considering an alternative solution to ballot order effects, which does not solely depend on randomization or rotation. Second, similar effects may impact any survey research using ranking questions. Future research must investigate the statistical consequences of pattern ranking for survey research. Finally, ranking data allow researchers to study diverse quantities of interest while targeting many different substantive questions. However, this flexibility also implies that analyzing ranking data can be prone to arbitrary analysis.

Edited by: Dr. Jeff Gill

1. Introduction

How does candidate order on the ballot affect voting behavior and election results? Ballot order effects have important implications for ballot design, candidate selection, and intra-party competition in American and comparative politics (Ho and Imai 2008; Marcinkiewicz and Stegmaier 2015; Ortega Villodres 2008). Previous research suggests that candidates in the first and last positions tend to receive more votes than other candidates (Alvarez, Sinclair, and Hasen 2006), focusing on first-past-the-post (Ho and Imai 2008) and party-list proportional representation (Gulzar, Robinson, and Ruiz 2022).

In this study, I extend the analysis of ballot order effects to electoral systems with ordinal ballots, where voters rank candidates. These systems include ranked-choice voting (RCV), instant-runoff elections or alternative vote, single-transferable vote, and Borda count (Reilly 2001; Santucci 2021). Here, I focus on elections with *grid-style ballots*. Although several studies examine ballot order effects

under some of these systems, unfortunately, these analyses are limited to voters' first-choice votes (Curtice and Marsh 2014; King and Leigh 2009; Marcinkiewicz and Stegmaier 2015; Orr 2002; Ortega Villlodres 2008). In contrast, I seek to propose an approach to study ballot order effects on people's entire candidate rankings.

This letter makes several contributions. First, I discuss two types of ballot order effects, including "position effects"—voters vote for specific candidates because of their ballot positions—and "pattern ranking"—voters rank candidates geometrically given their grid-style ballots. Next, I discuss experimental designs for identifying and estimating these effects. Finally, I illustrate the proposed methods by using survey and natural experiments based on actual RCV elections in 2022. I find that while voters seem less susceptible to specific ballot positions, ballot design can still impact voters via pattern ranking—even when candidate order is fully randomized.

This work has several implications for ballot design, survey research, and ranking data analysis. First, it shows that pattern ranking may affect electoral outcomes in RCV and other systems. With fixed ballot order, as in many RCV elections in the United States, pattern ranking may be the most consequential. Experts have suggested that ballot order randomization may solve the problem.[1] However, this letter demonstrates that pattern ranking may still affect electoral results even when ballot order is *fully randomized*, which is often considered the best but practically challenging solution.[2] Consequently, it may be worth considering an alternative solution to ballot order effects, which does not depend on randomization or rotation, when voters rank candidates.

Second, similar effects may impact all survey research using ranking questions. When a subset of respondents offer ranking responses based on geometric patterns and not their underlying preferences, collected data cannot be used to study people's attitudes and behaviors without caution. Thus, future research must investigate the statistical consequences of pattern ranking for survey research. Atsusaka and Kim (2024), for example, examine how pattern ranking leads to measurement errors in ranking questions and propose design-based methods to address the problem.

Finally, ranking data allow researchers to study diverse quantities of interest while targeting many different substantive questions. However, this flexibility also implies that analyzing ranking data can be prone to arbitrary analysis and presentation (Atsusaka 2023). Thus, future research must investigate how to leverage ranking data while avoiding selective inference.

2. Framework and Quantities of Interest

Consider an experiment where subjects are exposed to a grid-style ballot with a particular candidate order. Figure 1, for example, displays a grid-style ranking ballot based on the 2022 U.S. Senate election in Alaska.[3] In this experiment, the treatment is the *order* in which candidates appear on the ballot (e.g., Chesbro → Kelley → Murkowski → Tshibaka). The outcome in this experiment is a *ranking* that subjects assign to the list of candidates, such as (2, 1, 3, 4). I consider two effects: *position effects* and *pattern ranking*.

2.1. Position Effects

First, I consider the causal effect of candidate j ($= 1, \ldots, J$) shown in a specific position $O_{ij} = t$ (treated) as opposed to another position $O_{ij} = t^*$ (control) for voter i ($= 1, \ldots, n$). To define quantities of interest, I rely on the potential-outcomes framework for ranking data considered in Atsusaka (2023).

Let voter i's potential rankings be $\mathbf{Y}_i(O_{ij} = t)$ and $\mathbf{Y}_i(O_{ij} = t^*)$, respectively (e.g., $\mathbf{Y}_i(O_{ij} = t) = (1,4,2,3)$). Let $\boldsymbol{O}_{i[-j]}$ be the ordering (position) of the other $J - 1$ candidates. Finally, let $g()$ be a summary function that maps rankings into scalar values.

---

[1] For example, see a report by the Tasmanian Electoral Commission at https://www.tec.tas.gov.au/info/Robson_Rotation_Paper.pdf and a summary by the MIT Election Lab at https://electionlab.mit.edu/research/ballot-order-effects.

[2] Pattern ranking may also be used as a form of vote-buying (Williams, Baltz, and Stewart III 2024).

[3] For this sample ballot, see https://www.elections.alaska.gov/election/2022/genr/FEDERAL_ONLY.pdf.

> **PLEASE NOTE: A candidate's designated affiliation does not imply that the candidate is nominated or endorsed by the political party or group or that the party or group approves of or associates with that candidate, but only that the candidate is registered as affiliated with the party or group.**
>
> ### Instructions:
> - Do not use red ink or a pencil to mark your ballot.
> - Rank as many or as few candidates as you like.
> - Completely fill in no more than one oval for each candidate or column.
> - For your 1st choice, fill in the oval in the 1st choice column.
> - For your 2nd choice, fill in the oval in the 2nd choice column.
> - For your 3rd choice, fill in the oval in the 3rd choice column, and so on.
> - **If you make a mistake, you can ask for a new ballot.**
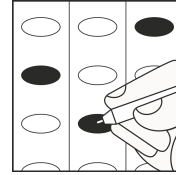
United States Senator	1st Choice	2nd Choice	3rd Choice	4th Choice
Chesbro, Patricia R. (Registered Democrat)	◯ 1	◯ 2	◯ 3	◯ 4
Kelley, Buzz A. (Registered Republican)	◯ 1	◯ 2	◯ 3	◯ 4
Murkowski, Lisa (Registered Republican)	◯ 1	◯ 2	◯ 3	◯ 4
Tshibaka, Kelly C. (Registered Republican)	◯ 1	◯ 2	◯ 3	◯ 4

Figure 1. Example of grid-style ordinal ballot.

One causal quantity of interest is the average treatment effect of position t for candidate j on voters' rankings with respect to counterfactual position t^* and the order of the remaining candidates $\boldsymbol{o}$. I call it the *conditional position effect*:

$$\tau_{jtt^*\boldsymbol{o}} = \mathbb{E}\left[g(\underbrace{\mathbf{Y}_i(O_{ij}=t, \boldsymbol{O}_{i[-j]}=\boldsymbol{o})}_{\text{potential ranking under the treatment}}) - g(\underbrace{\mathbf{Y}_i(O_{ij}=t^*, \boldsymbol{O}_{i[-j]}=\boldsymbol{o})}_{\text{potential ranking under the control}}) \,\Big|\, O_{ij}=t^*, \boldsymbol{O}_{i[-j]}=\boldsymbol{o} \right]. \quad (1)$$

While various summary functions can be used, the following application focuses on two functions for $g()$: (1) the marginal rank of candidate j (e.g., candidate A was ranked third) and (2) the indicator function that denotes whether candidate j was ranked or not (e.g., candidate B was not ranked).[4] Thus, I study the position effects on the *average rank of each candidate* and the *probability that each candidate is selected*, both of which are useful quantities in elections with ordinal ballots.

In many applications, however, researchers may wish to study the overall effect of position t for candidate j on voters' rankings *averaged over* all possible counterfactual position t^* *and* all possible ordering of the remaining candidates $\boldsymbol{o}$. I call it the *average position effect* of position t:

$$\tau_{jt} = \mathbb{E}\left[\tau_{jtt^*\boldsymbol{o}}\right]$$

$$= \sum_{\boldsymbol{o}}\left\{ \sum_{t^*} \mathbb{E}\left[g(\underbrace{\mathbf{Y}_i(O_{ij}=t, \boldsymbol{O}_{i[-j]}=\boldsymbol{o})}_{\text{potential ranking under the treatment}}) - g(\underbrace{\mathbf{Y}_i(O_{ij}=t^*, \boldsymbol{O}_{i[-j]}=\boldsymbol{o})}_{\text{potential ranking under the control}}) \,\Big|\, O_{ij}=t^*, \boldsymbol{O}_{i[-j]}=\boldsymbol{o} \right] \right.$$

$$\times \underbrace{\mathbb{P}(t, t^* | \boldsymbol{O}_{i[-j]}=\boldsymbol{o})}_{\text{conditional probability of candidate } j \text{ listed in } t\text{th and } t^*\text{th positions}} \Bigg\}$$

$$\times \underbrace{\mathbb{P}(\boldsymbol{O}_{i[-j]}=\boldsymbol{o})}_{\text{probability of a particular ballot order } \boldsymbol{o}}. \quad (2)$$

---

[4]The so-called primacy ($t = 1$) and recency ($t = J$) effects are based on the second function applied to categorical ballots.

The average position effect is the probability-weighted average of all possible conditional position effects.[5] Appendix A of the Supplementary Material shows that researchers can identify and estimate the average position effect by fully randomizing ballot orders. Here, one challenge is that the number of comparisons proliferates as the number of candidates increases. For example, with 10 candidates and a given treatment position t (e.g., $t = 1$) for a target candidate, researchers need to consider 9 counterfactual control positions t^* (e.g., $t^* \in \{2, 3, \ldots, 10\}$) and 9! ways to permute the other nine candidates. Accordingly, researchers need at least $9 \times 9! = 3{,}265{,}920$ control units to estimate the above effect.

To resolve this issue, Appendix A of the Supplementary Material shows that analysts can estimate the average position effect with considerably fewer control units by making two additional assumptions. Researchers can estimate the average position effect on candidate ranks, for example, for each position-candidate pair by regressing each candidate's rank (or any function of it) on a dummy variable denoting a specific ballot position.

2.2. Pattern Ranking

Pattern ranking refers to when voters provide ranked ballots by following specific geometric patterns independent of their preferences. A special case of pattern ranking—$(1, 2, \ldots, J)$—has been known as "donkey voting" (Orr 2002; Reilly 2001, 158). Figure 2 visualizes that pattern ranking generalizes the idea of donkey voting by accommodating many more geometric patterns, including what I call diagonal vote (Panel A, no angle), zigzag vote (Panel B, two or more major angles), and dogleg vote (Panel C, one major angle).

Appendix B of the Supplementary Material shows that when (1) ballot order is randomized and (2) no voter performs pattern ranking, raw responses (recorded rankings with respect to given candidate orders) follow a uniform distribution regardless of people's underlying preference. Thus, deviation from uniformity can be treated as potential evidence for pattern ranking. More formally, researchers can apply Pearson's χ^2 test to this problem with test statistic:

$$\chi^2 = \sum_{r=1}^{J!} \frac{\left(N\widehat{p}_r - N\frac{1}{J!}\right)^2}{N\frac{1}{J!}}, \tag{3}$$

where $\widehat{p}_r$ is the observed proportion of recorded ranking r $(= 1, \ldots, J!)$, $\frac{1}{J!}$ is the expected proportion of the same ranking under the null, and N is the number of observations. Here, one limitation is that applying the test to ranking data with more than five or six items becomes challenging. I leave future research to address the limitation.

3. Application I: Survey Experiments in Congressional and Mayoral RCV Elections

To illustrate the framework, I performed survey experiments via Lucid Marketplace from October 10 to November 7, 2022, before the Oakland mayoral election, the U.S. House of Representatives election in
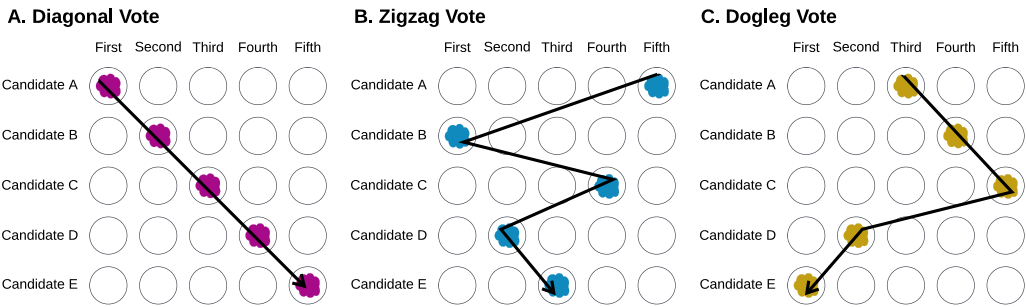

Figure 2. Examples of pattern ranking.

---

[5] The joint probability $\mathbb{P}(O_{ij} = t^*, \boldsymbol{O}_{i[-j]} = \boldsymbol{o})$ is factorized into $\mathbb{P}(O_{ij} = t^* | \boldsymbol{O}_{i[-j]} = \boldsymbol{o})\mathbb{P}(\boldsymbol{O}_{i[-j]} = \boldsymbol{o})$.

Alaska, and the U.S. Senate election in Alaska.[6] I used survey experiments because they allowed me to fully randomize ballot orders. This section examines ballot (item) order effects in the context of surveys. Thus, its results may directly apply to all ranking questions beyond RCV. In contrast, the next section seeks to offer more realistic estimates of ballot order effects in actual elections.

The survey presented a list of actual candidates to randomly sampled respondents from Oakland ($n = 258$) and Alaska ($n = 354$), respectively. The survey asked all respondents two questions. In "forced ranking" questions, which are used to study position effects on average ranks and pattern ranking, respondents must rank all candidates. In "optional ranking" questions, which are leveraged to identify position effects on candidate selection and pattern ranking, they could rank up to three candidates in
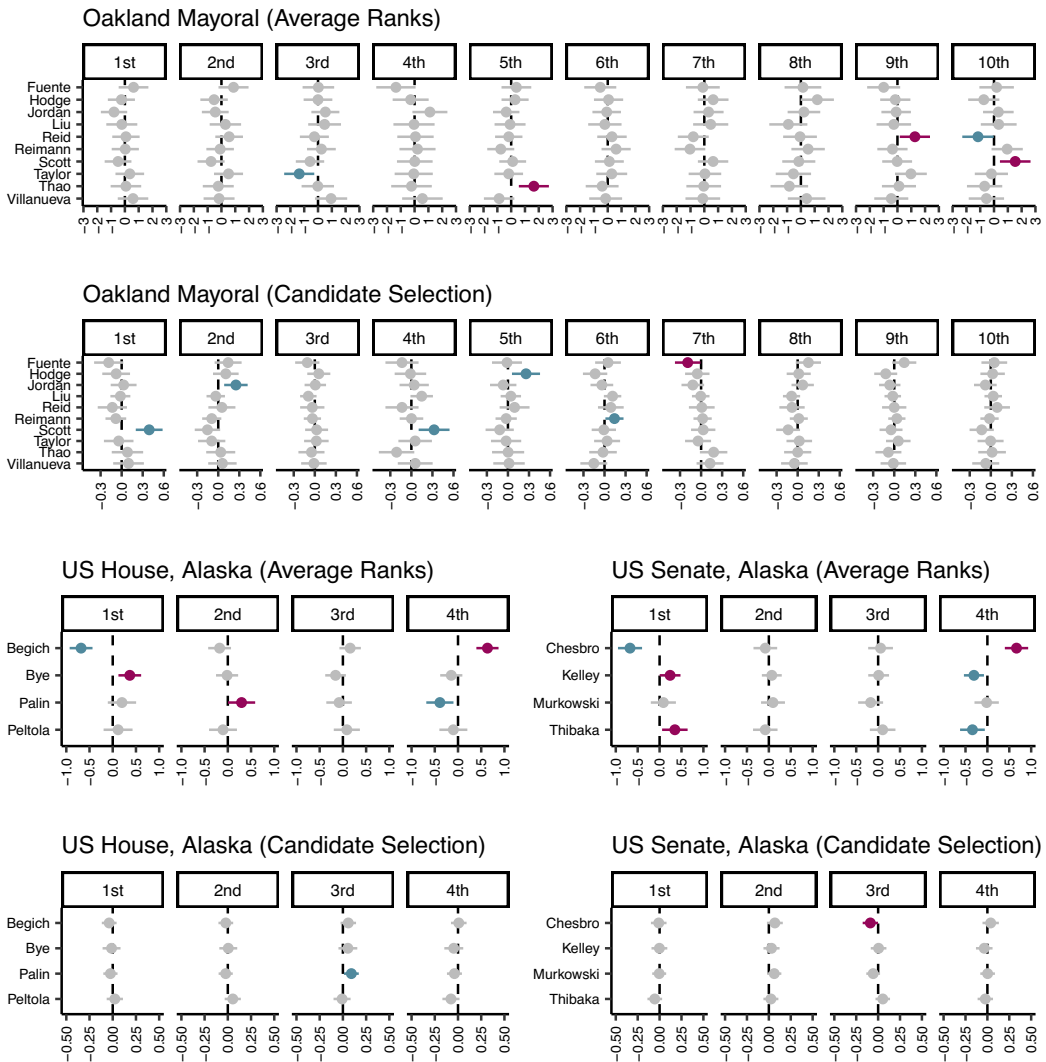
Figure 3. Average position effects.
Note: Blue = positive, red = negative, gray = nonsignificant.

---

[6]The replication data and materials are available in Atsusaka (2024).

Oakland and four candidates in Alaska, reflecting their election laws. The order of the two questions was randomized.[7] Appendix C of the Supplementary Material summarizes the survey design.

3.1. Position Effects

Figure 3 visualizes the average position effects on average ranks (upper panels) and candidate selection probabilities (lower panels) for all candidate-position combinations in the three elections. The blue (red) results represent positive (negative) effects, denoting *higher* (*lower*) average ranks and selection probabilities, while the gray results suggest nonsignificance.

I find that about 91% (240/264) of the average position effects are statistically insignificant. Although some effects are statistically significant, I find almost no consistent pattern among them. Indeed, 5% of all effects can be significant purely by chance. The results are conservative in the sense that we should observe *less* significant effects (stronger evidence) if I correct statistical inference for multiple comparisons. The result suggests that respondents were hardly susceptible to average position effects in the experiments.

3.2. Pattern Ranking

Panel A of Figure 4 plots the empirical distributions of all recorded rankings in the U.S. House (left) and U.S. Senate (right) races in Alaska. The panel shows the results from the forced (filled) and optional (non-filled) ranking questions. There are 4! = 24 different ways to rank, and each recorded ranking (given a particular candidate order) should be selected with probability 1/24 = 0.042 without pattern ranking. In Panel B, I visualize all 24 rankings, which are categorized into either the diagonal vote (solid line), zigzag vote (dashed line), or dogleg vote (dotted line).[8]

Panel A shows nonuniform distributions of recorded rankings, implying the presence of pattern ranking. Some geometric patterns, notably (1, 2, 3, 4), are selected significantly more often than others. This means that although some of the ballots reflect people's genuine preferences, many of them may be based solely on a diagonal line. Pearson's χ^2 test rejects the null hypothesis that each of the four distributions is uniform at the $\alpha = 0.05$ level. Moreover, in Tables F1 and F2 in the Supplementary Material, I find that the dogleg vote is more prevalent than the zigzag vote in almost all cases.

Panel B of Figure 4 provides additional evidence for pattern ranking—31% and 37% of respondents offered the same recorded ranking across the two questions, even though only 3% (House) and 1.6% (Senate) of respondents saw the same candidate order. Namely, many people selected the same geometric pattern even though they were exposed to *different orders* of candidates. While it requires additional information to estimate the proportion of pattern rankers (Atsusaka and Kim 2024), the results offer partial evidence for pattern ranking.

4. Application II: A Natural Experiment from Alaskan RCV Elections

Finally, I present an alternative strategy to study pattern ranking. In Alaska's federal and gubernatorial races, candidates are alphabetically ordered in the first state house district and rotated in subsequent districts.[9] For example, consider four candidates: A, B, C, and D. Then, District 1 has ballot order {A, B, C, D}, District 2 {B, C, D, A}, District 3 {C, D, A, B}, and District 4 {D, A, B, C}. This rotation continues until District 40.

This "alphabetization-rotation" procedure may give us a unique chance to analyze RCV in Alaska as a natural experiment. Here, districts are the units of analysis. The outcome is the proportion of voters providing a particular candidate ranking.[10] The binary treatment denotes whether voters see a given ballot order or not.

---

[7]Appendix E of the Supplementary Material shows that, overall, the survey data predict the relative popularity of each candidate.

[8]With optional ranking, I only use respondents who ranked four or three candidates (I filled in the rank of the last candidate). Analyzing only respondents who ranked four candidates led to the same substantive conclusion.
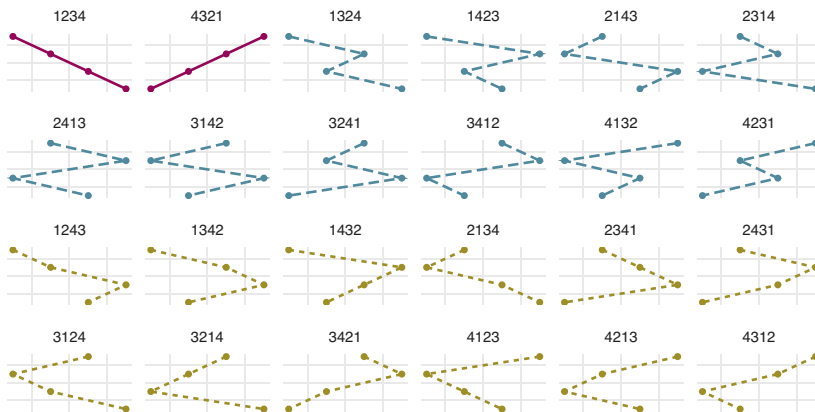
[9]Appendix G1 of the Supplementary Material discusses how Alaska's procedure differs from the so-called Robson Rotation.

[10]Here, I focus on a subset of voters who rank four candidates.

A. Proportion of Observed Rankings
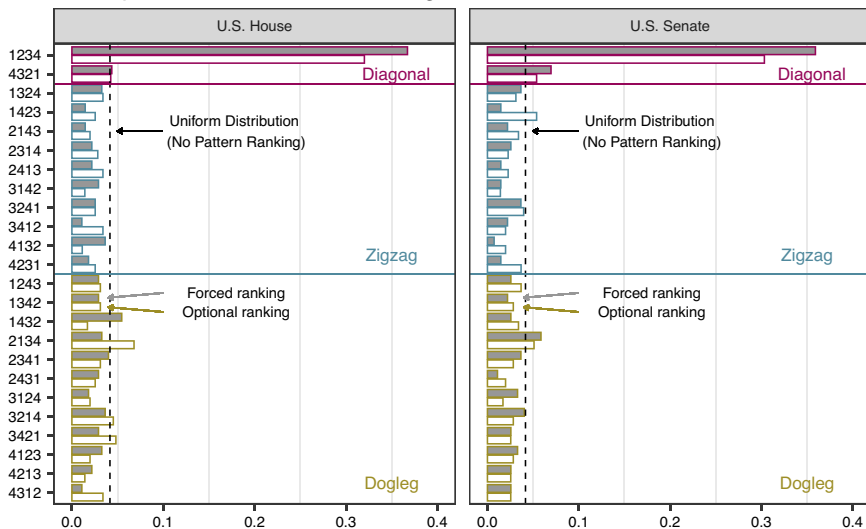
B. Types of Pattern Ranking

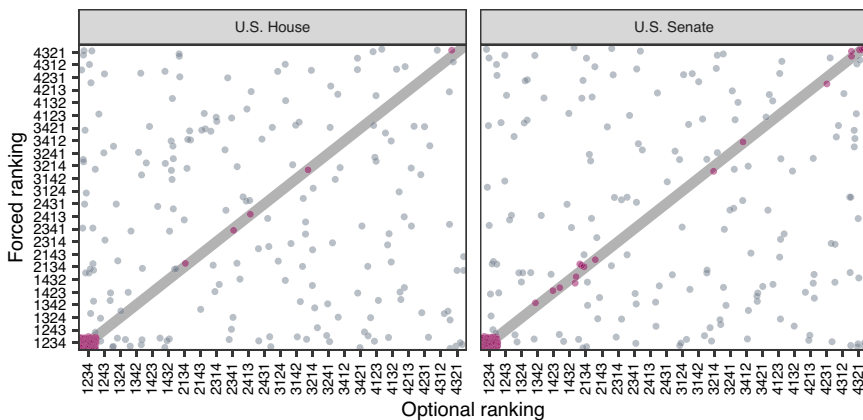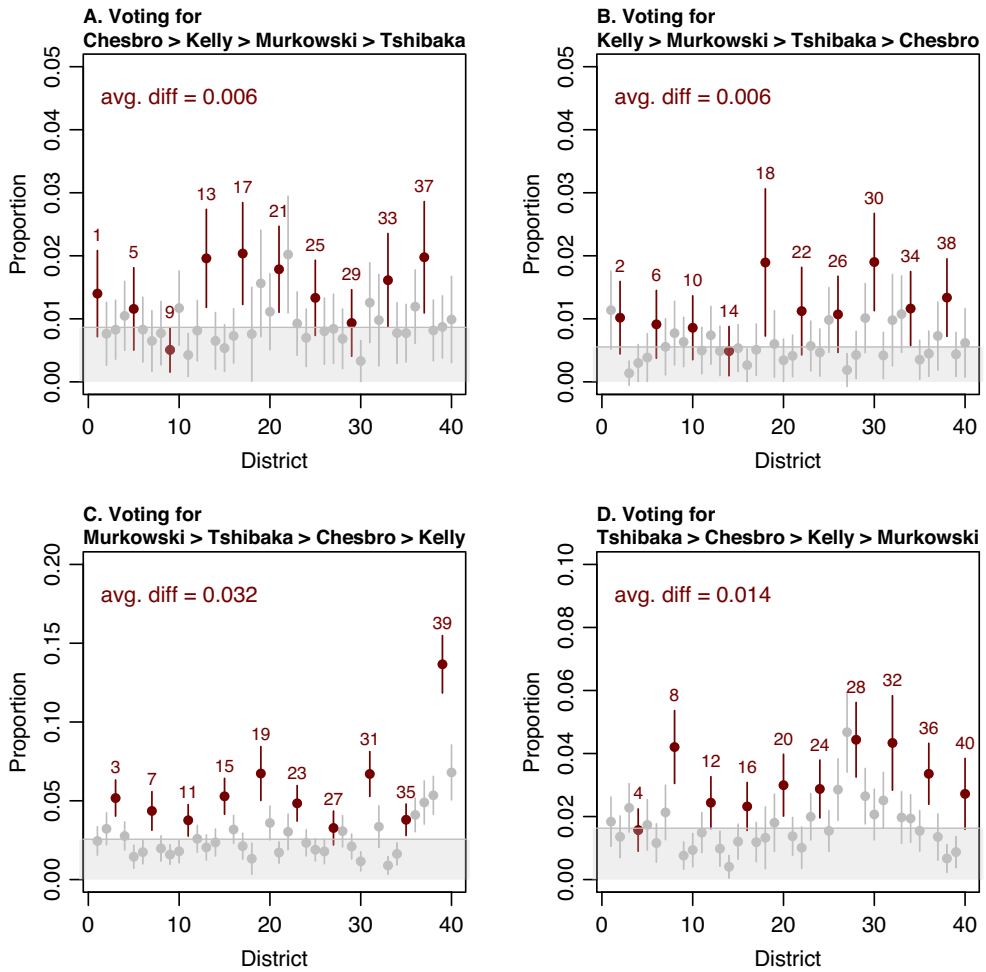## C. Observed Rankings in Forced and Optional Questions

Figure 4. Suggestive evidence for pattern ranking.

Figure 5. Analysis of donkey voting (U.S. Senate election in Alaska).
Note: Numbers index treated districts. Horizontal lines are control average values.

Suppose that a target outcome is the proportion of candidate ranking (B, C, D, A). Suppose also that researchers wish to study the prevalence of donkey voting. Then, Districts 2, 6, ..., 38 become "treated" districts, where performing donkey voting is the same as casting ranked ballot (B, C, D, A), while others are "control" districts. *How many voters provided the target candidate ranking per donkey voting?* One possible identification strategy is to compare the outcomes between the treated and control districts. The key assumption is that the proportion of voters who prefer (B, C, D, A) is comparable across the treatment status. Then, the difference in the proportions can be solely attributable to donkey voting.

Figure 5 illustrates this idea by plotting control and treated (with district numbers) districts with respect to four candidate ranking profiles in the 2022 U.S. Senate election. The horizontal lines represent the average values for the control districts. The graph illustrates that the treated districts, on average, have higher proportions than the control districts. I find that about 0.6%–3% of actual ballots may be attributable to donkey voting. Appendix F of the Supplementary Material presents similar results for the 2022 U.S. House and Alaska gubernatorial elections, providing suggestive evidence that donkey voting may be consistently present in actual RCV elections. Of course, this remains a demonstrative analysis, and future research must scrutinize the validity

of the identification strategy and its applications to position effects and other forms of pattern ranking.

5. Limitations and Future Research

One notable limitation of this study is that the above analyses require voters' full rankings, except for the analysis of position effects on candidate selection. To analyze pattern ranking, for example, this work only examined respondents (in Figure 4) and voters (in Figure 5) who ranked four candidates. This is an important limitation because, in many contests, voters may rank a single or at least not all candidates. Thus, it is critical that future research extend the proposed methods to partially ranked data. Potential extensions include the partial identification of position effects, where we may leverage principal stratification and extreme value bounds, and the analysis of "partial pattern ranking," where we focus on three or four digits in people's entire candidate rankings. Despite this and other limitations (Appendix H of the Supplementary Material), this letter makes several contributions to the study of ballot order effects when voters rank candidates.

References

Alvarez, R. M., B. Sinclair, and R. L. Hasen. 2006. "How Much Is Enough? The 'Ballot Order Effect' and the Use of Social Science Research in Election Law Disputes." *Election Law Journal* 5 (1): 40–56.

Atsusaka, Y. 2023. "Causal Inference with Ranking Data: Application to Blame Attribution in Police Violence and Ballot Order Effects in Ranked-Choice Voting." Preprint. https://doi.org/10.48550/arXiv.2207.07005

Atsusaka, Y. 2024. "Replication Data for: Analyzing Ballot Order Effects When Voters Rank Candidates." Harvard Dataverse, V1. https://doi.org/10.7910/DVN/AJXRCV

Atsusaka, Y., and S. S. Kim. 2024. "Addressing Measurement Errors in Ranking Questions for Social Sciences." *OSF Preprints*. https://doi.org/10.31219/osf.io/3ys8x

Curtice, J., and M. Marsh. 2014. "Confused or Competent? How Voters Use the STV Ballot Paper." *Electoral Studies* 34: 146–158.

Gulzar, S., T. S. Robinson, and N. A. Ruiz. 2022. "How Campaigns Respond to Ballot Position: A New Mechanism for Order Effects." *The Journal of Politics* 84 (2): 1256–1261.

Ho, D. E., and K. Imai. 2008. "Estimating Causal Effects of Ballot Order from a Randomized Natural Experiment: The California Alphabet Lottery, 1978–2002." *Public Opinion Quarterly* 72 (2): 216–240.

King, A., and A. Leigh. 2009. "Are Ballot Order Effects Heterogeneous?" *Social Science Quarterly* 90 (1): 71–87.

Marcinkiewicz, K., and M. Stegmaier. 2015. "Ballot Position Effects under Compulsory and Optional Preferential-List PR Electoral Systems." *Political Behavior* 37 (2): 465–486.

Orr, G. 2002. "Ballot Order: Donkey Voting in Australia." *Election Law Journal* 1 (4): 573–578.

Ortega Villlodres, C. 2008. "Gender and Party Duopoly in a Small State: Ballot Position Effects under the Single Transferable Vote in Malta, 1947–2008." *South European Society and Politics* 13 (4): 435–456.

Reilly, B. 2001. *Democracy in Divided Societies: Electoral Engineering for Conflict Management*. Cambridge: Cambridge University Press.

Santucci, J. 2021. "Variants of Ranked-Choice Voting from a Strategic Perspective." *Politics and Governance* 9 (2): 344–353.

Williams, J., S. Baltz, and C. StewartIII. 2024. "Votes Can Be Confidently Bought in Some Ranked Ballot Elections, and What to Do about It." *Political Analysis*. https://doi.org/10.1017/pan.2024.4