Voters' Cognitive Bias and Strategic Candidate Entry *

Xuan Li †

This version: December 2023

Abstract: I study whether voters' cognitive biases affect political candidates' entry decisions. Building off the insight that in down-ballot elections, voters tend to choose the first-listed candidate due to choice fatigue and the primacy effect, I conjecture that potential candidates with late-alphabet surnames, expecting positional disadvantages on an alphabetically ordered ballot, are less likely to run for office. Using within-state variation in ballot order rules and data on 341,156 candidates running for U.S. state legislatures from 1967 to 2022, I find that alphabetically ordered ballots have an impact on candidate entry, resulting in a 3.68 percentage-point decrease in the representation of late-alphabet candidates (equivalent to a 16.4% reduction). This shift in composition is primarily driven by a decline in late-alphabet candidates running for office, which exacerbates the overall shortage of candidates in state elections. Moreover, alphabetically ordered ballots may unintentionally impact minority candidate entry, due to these candidates' distinctive distribution of surname initials.

Keywords: Cognitive bias, ballot order rule, electoral institution, candidate entry

Word count: 8203

---

*I am indebted to Raymond Fisman for his continuous guidance, advice, and support. I would also like to thank Justin de Benedictis-Kessner, Jacob Brown, Michelle Dion, Oeindrila Dube, Siddharth George, James Feigenbaum, Claudio Ferraz, Martin Fiszbein, Andrey Fradkin, Justin Hong, Kevin Lang, Erzo F.P. Luttmer, Benjamin Marx, Sultan Mehmood, Dilip Mookherjee, Daniele Paserman, Maxwell Palmer, Spencer Piston, Guillem Riambau, Johannes Schmieder, Lu Sun, Yuzhao Yang, and all participants at the Development Reading Group (BU 2022), Toronto Political Behaviour Workshop (UoT 2022), EPOVB Conference (FSU 2023), Applied Micro Workshop (BU 2023), 2024 ASSA Annual Meeting, and Banff Empirical Microeconomics Conference (UCalgary 2024) for their helpful comments. All errors are my own.

†Department of Economics, Boston University (email: xuanli@bu.edu)

1 Introduction

Politicians are aware of the public's limited attention and strategically time unpopular measures during newsworthy events to distract both the media and the broader community (Durante and Zhuravskaya, 2018). While similar cognitive limitations/biases—which can have important electoral consequences—have been observed in voting behavior (Shue and Luttmer, 2009; Augenblick and Nicholson, 2016), there is little research exploring whether or how politicians strategically respond to voters' cognitive biases. In this paper, I provide evidence that politicians both recognize the existence of bias and are responsive to it, in the context of U.S. state legislature elections.

Due to choice fatigue and the primacy effect, voters without a strong preference are likely to vote for the first-listed candidate that they have no reason to oppose (Koppell and Steen, 2004; Faas and Schoen, 2006; Ho and Imai, 2008; Meredith and Salant, 2013; Van Erkel and Thijssen, 2016; Augenblick and Nicholson, 2016). This well-documented ballot order effect happens more often in low-information elections (Brockington, 2003; Grant, 2017). Interestingly, various ballot order rules may assign different positions to the same candidate, potentially granting them a positional advantage. On an alphabetically ordered ballot, candidates with early-alphabet surnames are much more likely to appear first and thus receive a sizable portion of windfall votes. A randomized or rotated ballot order, however, is unrelated to alphabetic order and is not expected to favor any particular candidates.

Anecdotally, candidates appear to understand the advantage of being listed first on the ballot, even to the point of attributing their losses to an alphabetically ordered ballot. For example, as the fourteenth candidate on a U.S. Senate primary ballot, Michael Schaefer lost to Benjamin Cardin, the first-listed candidate. Schaefer asserted that ballots arranged in alphabetical order pose a significant disadvantage to individuals like him with names towards the end of the alphabet. He took this matter to court, challenging Maryland's method of alphabetical ballot ordering (Edwards, 2014).[1] As candidates understand the ballot advantage, they seem to naturally adapt to ballot order rules to gain the advantage if permitted. For example, Illinois ranks candidates based on their application time, and candidates even camp out over night at the office of the Secretary of State to appear first on the ballot (Beazley, 2013).[2]

To the extent that candidates recognize the disadvantage of not being listed first, we may conjecture that prospective candidates with late-alphabet surnames, anticipating that disadvantage on an alphabetical ballot, may opt out of running for political office. To evaluate this hypothesis, I combine U.S. state legislative election candidate data (1967-2022) (Klarner, 2023) with data on primary election ballot order rules across 34 states, which use either alphabetically ordered or randomized/rotated ballots (Edwards, 2015).

Specifically, I evenly divide the 26 English surname initials into three groups: early-alphabet names (letters A-H), middle-alphabet names (letters I-R), and late-alphabet names (letters S-Z).[3] I find that states with alphabetically ordered ballots have much higher (lower) shares of candidates with early-alphabet surnames (late-alphabet surnames). This observed pattern doesn't seem to stem from differences in the underlying population. Moreover, the pattern remains the same when the identification exploits within-state variation in ballot order rules, i.e., moving from alphabetically ordered ballots to randomized/rotated ballots. In the preferred specification, with alphabetically ordered ballots, the share of candidates with late-alphabet surnames (early-alphabet surnames) decreases (increases) by 16.4% (6.9%). Additionally, I plot event-study graphs to show that in states that change from alphabetically ordered ballots to a random/rotated order, the composition of candidates converges to that of

[1]Schaefer v. Lamone, 2006 U.S. Dist. LEXIS 96855, 2006 WL 8456798 (United States District Court for the District of Maryland, November 30, 2006, Filed).

[2]See https://capitolnewsillinois.com for media coverage.

[3]While acknowledging the inherent arbitrariness in any classification, it's important to note that my primary conclusions remain robust and unaffected by this chosen categorization. Under two other specifications, the results remain the same: (1) dividing the 26 letters evenly into five surname groups, and (2) assigning letters numeric values according to their alphabetical positions (e.g., A is 1, B is 2, etc.).

states using the latter order. Overall, I interpret these results as indicating that prospective candidates' entry decisions anticipate the effects of voters' cognitive biases.[4]

The compositional change in the candidate pool is primarily due to a decrease in candidates with late-alphabet surnames running for office. This reduces the competitiveness of elections, highlighting a significant concern in modern U.S. state politics: the overall lack of candidates running for office. Moreover, the compositional change in the candidate pool can have other real consequences: I find that the composition of candidates is highly correlated with the composition of elected legislators. Furthermore, the choice of ballot order rules can potentially affect the state's minority representation. Minorities have different surname distributions in the U.S.: 43% of Whites have an early-alphabet surname, while 22% of them have a late-alphabet surname. However, among Asian Americans early-alphabet surnames are much less prevalent (around 31%) and late-alphabet surnames are much more prevalent (27%). If different race/ethnic groups respond similarly to the ballot order effect, it is expected that alphabetically ordered ballots will be associated with fewer Asian American candidates, thus altering the set of candidates that voters can choose from. Moreover, individuals who hold a higher social status, as indicated by their education level or income, are also more likely to possess a late-alphabet surname. Again, this raises the concern that alphabetically ordered ballots may adversely affect the composition of candidates in a nontrivial manner. Finally, I discuss the association between alphabetically-ordered ballots and state legislator behavior, in particular absences from roll-call votes.

This work mainly contributes to the literature on political selection, particularly research that emphasizes endogenous candidate entry (Black, 1972; Osborne and Slivinski, 1996; Besley and Coate, 1997; Dal Bó et al., 2017; Dal Bó and Finan, 2018; Gulzar, 2021). The existing literature tends to estimate the

[4]In Section 5.3, I delve into an alternative explanation: that the combined influence of the ballot order effect and incumbents' political persistence can alter the candidate composition, even in the absence of deliberate strategies by the candidates themselves. Moreover, I also discuss the role of party recruitment and the interpretation that parties internalize ballot order effects through strategic recruitment of candidates in Section 5.3.

benefits and costs of winning office (Diermeier, Keane and Merlo, 2005; Dal Bó, Dal Bó and Snyder, 2009; Fisman et al., 2015) and analyze how these two key parameters influence candidacy decisions (Caselli and Morelli, 2004; Messner and Polborn, 2004; Kotakorpi and Poutvaara, 2011; Brollo et al., 2013; Dal Bó, Finan and Rossi, 2013; Gagliarducci and Nannicini, 2013; Fisman et al., 2015; Gulzar and Khan, 2018; Pique, 2019). This is the first paper to demonstrate that voters' cognitive biases and their interplay with electoral rules play a surprisingly large role in political entry decisions. Specifically, I show that alphabetically ordered ballots discourage those who have late-alphabet names to run for office, and thus distort the composition of candidates in U.S. state legislative elections. More broadly, my results can be linked to the candidate's probability of winning office, a key parameter in standard political selection models. It is generally challenging to empirically evaluate the extent to which win probability affects the decision to run for office, since each citizen's probability is essentially unobservable and related to personal characteristics and broader political conditions. The rule-based shift from alphabetically ordered ballots to randomized/rotated ballots, introduces within-individual variation in the expected vote share for potential candidates as a function of surname initial, while essentially keeping the benefits and costs constant. My finding thus serve as an empirical validation of a core assumption in political selection models.

This paper also complements a distinguished literature on the ballot order effect (Taebel, 1975; Brockington, 2003; Koppell and Steen, 2004; Ho and Imai, 2008; Meredith and Salant, 2013; Augenblick and Nicholson, 2016; Jun and Min, 2017; Gulzar, Robinson and Ruiz, 2022). Most of this literature has focused on political selection–essentially the magnitude of candidates' electoral advantage of being at the top of the ballot–and its underlying mechanisms.[5] A handful of recent papers highlight the down-

[5] Gulzar, Robinson and Ruiz (2022) show that the ballot order effect may not be fully explained by voters' cognitive biases. It can, at least in part, be attributed to candidates strategically adjusting their campaign behavior when assigned a top position on the ballot. My findings complement Gulzar, Robinson and Ruiz (2022)'s work by demonstrating that potential candidates strategically respond to the (expected) ballot order even before entering the election. Despite the similarity in strategic responses, it is important to note that we are addressing two different research questions. Gulzar, Robinson and

stream electoral consequences of the ballot order effect, showing that it can alter the representation of elected officials and subsequent policy making (Shi and Singleton, 2023; Fischer, 2023). Closest to my research setting is the work of Edwards (2015), which demonstrates that alphabetically ordered ballots are associated with a higher proportion of elected legislators with early-alphabet surnames. Considering the evident similarities in our research settings, it's essential to highlight the unique features of my work. Rather than documenting and estimating the magnitude of the ballot order effect (as is done in Edwards (2015) and other existing work), this paper takes a step back and exhibits a novel perspective on the impact of the ballot order effect. It is the first to demonstrate that the ballot order effect can extend upstream: anticipating (dis)advantages from ballot position can influence the entry margin and pool of candidates in the first place. Incorporating the entry margin into the analysis of the ballot order effect suggests that this effect may be underestimated, as candidate selection occurs before researchers can directly observe the vote share of candidates. I further discuss how ballot order rules can potentially influence descriptive representation because surname distributions across racial/ethnic groups vary significantly. This adds to the growing body of literature that emphasizes the role of electoral institutions in shaping minority (under)representation (Trebbi, Aghion and Alesina, 2008; Ricca and Trebbi, 2022).

2 Setting: U.S. state legislative elections

I study alphabetical ballot order effects using U.S. state legislative elections. Despite the outsized power wielded by state legislatures, Americans know remarkably little about their state representatives, often making state legislative elections low-information contests – for example, Rogers (2017) finds that unpopular roll calls do not have severe electoral consequences for state legislators, partly due to the

Ruiz (2022) seek to understand the source of the ballot order effect beyond voters' cognitive biases, whereas my focus is on the impact of ballot order on candidates' decisions to enter the race.

lack of media coverage.[6]

The lack of attention to state contests may make them particularly susceptible to ballot order effects.[7] Critically from an identification perspective, the ballot order rule in legislative elections varies across states and changes over time, providing necessary variation to examine how ballot order effects influence candidate entry. Among the various ballot order rules, the alphabetically ordered ballot and the randomized/rotated ballot stand out as two of the most commonly employed. The randomized/rotated ballot provides an ideal setting for estimating the advantage of being listed first by introducing within-individual variation in ballot position across districts/precincts in the election. This does not allow for the identification of entry effects, since the randomization/rotation process leads to an average ballot advantage of zero across all candidates. However, an alphabetically ordered ballot, in principle, provides clear advantages to early-alphabet candidates and disadvantages to late-alphabet candidates. Based on primary and secondary sources, Edwards (2015) identifies 34 states that have ordered their primary election ballots either alphabetically or randomly/rotated among precincts (see Table 1).[8] Moreover, two states have transitioned from the alphabetically ordered ballot to the randomized/rotated ballot: Indiana in 1991 and New Hampshire in 2006.[9] Therefore, analyzing the timing of this shift and the distribution of candidates' surnames could offer insights into how potential candidates recognize and leverage opportunities to take advantage of voters' inattention.

[6]At the federal level, position-taking on roll calls has electoral consequences (Bovitz and Carson, 2006; Carson et al., 2010).

[7]Despite the importance of state politics, there is remarkably little research on how it is potentially affected by an inattentive electorate. Some earlier work does suggest that the electorate votes on matters other than policy or party concerns. For example, Koppell and Steen (2004)), using candidate name rotation across precincts to estimate the ballot order effect in the 1998 New York City Democratic primary. They find that in seven out of 71 contests, the advantage of being first on the ballot exceeds the margin of victory for the winner. As Meredith and Salant (2013) demonstrate, in other comparable large-scale low-information elections such as California city council and school board elections, first-listed candidates are likely to win by approximately five percentage points more often than expected when ballot order effects are absent.

[8]U.S. ballot order rules are highly decentralized: they not only differ between states but also within states, and they change over time (Miller, 2010). Thus, surveying ballot order rules in the U.S. poses a significant challenge. To streamline the analysis and avoid the added complexity of other ballot order rules, this paper primarily focuses on comparing the alphabetical ballot with an ideal counterfactual: the randomized/rotated ballot.

[9]It is also noteworthy that no state has changed its ballot order rules in the opposite direction, from randomized/rotated ballots to alphabetically ordered ballots.

Before proceeding to an overview of the data, I note that the state legislative candidate recruitment process underwent a profound decentralization due to a series of reforms implemented during the Progressive Era a century ago (Squire and Moncrief, 2019). In particular, the introduction of the primary election system widened the pool of potential candidates. According to Moncrief, Squire and Jewell (2001) 's survey on non-incumbent candidates for state legislatures, nearly one-third of candidates are "self-starters," and 45% of candidates are approached and encouraged to run by a local party committee (as it is ultimately the local party committee's responsibility to fill out the party's slate of candidates), while only one fifth are convinced by the local party to run for election. State legislative races are often not very competitive: 40% of state legislative races nationwide are uncontested, another 40% are contested but not competitive, and only the rest 20% are genuinely competitive. Local party leaders are more likely to engage in candidate recruitment and campaign support when elections in some districts are expected to be competitive but potentially winnable. Furthermore, state legislative races typically do not involve a great deal of campaign money and most of that money is only spent in a few of truly competitive districts. For instance, in 2016, the average raised by top-funded candidates in each state legislative race in New Hampshire and Indiana are $4,772 and $151,847, respectively. In addition, only 22 percent and 6 percent of legislative races in these two states are financially competitive, meaning the top fundraiser raised no more than twice as much as the second-top fundraiser.[10]

3 Data

3.1 Data source

Candidates running for down-ballot elections are drawn from State Legislative Election Returns (1967-2022) (Klarner, 2023). The dataset contains 341,156 candidates who ran for U.S. state legislatures from 1967 to 2022. 178 elections with only incumbent legislators are dropped. To eliminate duplicates,

[10] Source: https://www.followthemoney.org.

I only keep one observation for one candidate running for one post (i.e., State*Sen./Rep.*district*year). Based on this candidate-level data, I then calculate the share of candidates with early/middle/late-alphabet surnames for each post and construct the post-(surname)group-level data I will use in the main analysis.

3.2 Preliminary evidence

A preliminary analysis of how different ballot order rules correlate with different candidate compositions is provided in Panel A of Table 2. Compared with elections that randomize/rotate ballot order, elections that alphabetically order ballots have a greater share of candidates with early-alphabet surnames and a lower share of candidates with late-alphabet surnames. In terms of magnitude, elections with alphabetically ordered ballots have 7.8% more candidates with early-alphabet surnames and 9.8% fewer candidates with late-alphabet surnames.

I also plot the distribution of candidates' surnames by ballot ordering methods in Figure A1. The maximum distance between these distributions occurs at "Dy." In states employing random name ordering, only 23.3% of candidates have surnames up to "Dy" alphabetically. Conversely, in states with alphabetic ordering, this figure rises to 27.5%. A non-parametric Kolmogorov–Smirnov test suggests that these two distributions are unlikely to be equal, with a p-value of 0.059.

3.3 The population's distribution of surname initials

We might have a problem with the observation based on a simple comparison between elections with different ballot order rules: the results may just represent the fundamental difference between surnames among states with different ballot order rules. Those states that order ballots alphabetically may have a greater percentage of voters with early-alphabet surnames, which proportionally results in a greater percentage of candidates with early-alphabet surnames. In order to shed light on this concern, I calculate the share of surname initials in each sample state based on L2 voter data (2022).[11] It contains all registered voters in sample states in 2022, and has more than one hundred and twenty million observations.

Results are reported in Panel B in Table 2. It is worth noting that the L2 voter data only provides a snapshot of the electorate surnames in each state in 2022. Nevertheless, Panel B provides a useful benchmark to understand the differences between states using different ballot order rules. There are three key observations. First, in states (elections) that randomize/rotate ballot orders, candidates' surname distributions are close to the electorate's distributions. The share of the electorate with early-alphabet surnames (middle/late-alphabet surnames) is 41.86% (35.90%/22.24%), and the share of candidates with early-alphabet surnames is 41.99% (35.60%/22.40%). Second, candidates' surnames differ from the electorate's surnames in states (elections) that alphabetically order ballots. The difference is consistent with the advantage/disadvantage that alphabetically ordered ballots provide for candidates with early/late alphabet surnames. Third, a similar distribution of surname initials is observed in states with different ballot order rules. In comparison with Panel A's candidate pool, the difference in the electorate is mild and small. The results indicate that states with different ballot order rules do not have fundamentally different surname initial distributions. Nevertheless, in the next section, I utilize an identification strategy based on the existence of within-state variations in ballot order rules.

4 Identification

4.1 Identification specification

In order to investigate whether voter cognitive bias affects candidate entry decisions, I estimate the following specification for the office-year-group-level data containing all states (and corresponding

[11] L2 is a private data vendor: https://l2-data.com/. Population surnames can also be found in census records. However, census data with information on names is not publicly available.

periods after 1967) listed in Table 1[12]:

$$GroupShare_{o(s)tg} = \beta_1 G1_g * Rnd_{st} + \beta_2 G2_g * Rnd_{st} + \beta_3 G3_g * Rnd_{st} + \pi_{og} + \pi_{tg} + \epsilon_{otg}$$

where $GroupShare_{o(s)tg}$ is the share of candidates whose surname initial belongs to the group g in year t running for office o (office = State*Sen./Rep.*district) in state s.[13] Surnames are divided into three equal groups: early-alphabet (A, B, C, D, E, F, G, and H), middle-alphabet (I, J, K, L, M, N, O, P, Q, and R), and late-alphabet (S, T, U, V, W, X, Y, and Z). Rnd_{st} is a dummy variable indicating whether the order of the ballot is randomized or rotated in state s year t. $G1_g$ is a dummy variable indicating whether the group share is the share of candidates with early-alphabet surnames. Similarly, $G2_g/G3_g$ is a dummy variable indicating whether the group share is the share of candidates with middle/late-alphabet surnames. The coefficients of interest, β_1, β_2, and β_3, indicate how the share of candidates with early, middle, and late-alphabet surnames, respectively, respond to a change in the ballot order rule.[14] It would be reasonable to expect β_1 to be negative and β_3 to be positive if voters' cognitive bias influences candidate entry decisions.

Controlling for office-group fixed effects, π_{og}, is crucial to the analysis, which exploits the variation in ballot order within states: Indiana and New Hampshire changed the ballot order rules during the sample period.[15] Assumptions such as similarity between states that alphabetically order ballots and

[12]To be more specific, states that adopt one of the two ballot order rules during the sample period are included in the sample from the time of adoption onwards. For instance, Florida adopted an alphabetical ballot order in 1971, and thus Florida enters the sample in 1971.

[13]$GroupShare_{o(s)tg}$ is equal to the number of candidates whose surname initial belongs to the group g in year t running for office o divided by the total number of candidates in year t running for office o.

[14]For example, if the adoption of randomized/rotated ballots led to a change in the share of candidates with early-alphabet surnames from 0.45 to 0.42, then the value of β_1 would be -0.03. It's worth noting that the sum of β_1, β_2, and β_3 must be 0 since the total share of three surname groups always equals 1.

[15]Note that the state fixed effects are absorbed by office-group fixed effects. Given the extensive duration covered by the sample (from 1967 to 2022), it's plausible that the districts experienced changes over these years. To account for potential district alterations during this period, I present the analysis based on a State-Sen./Rep.-Year-Group panel in Column (3) of Table A2. The results are essentially the same.

states that randomize/rotate ballot order are no longer required. Instead, the identification requires that changes in the ballot order rule are not correlated with changes in surname distributions in the general population. I provide some evidence for this identification assumption in the following section. Also, I include year-group fixed effects, π_{tg}, to take into account the overall change in population surname distribution. Standard errors are clustered at the state-group level.[16]

Furthermore, I can take advantage of the staggered shift in the ballot order rule across states. Indiana switched from alphabetically ordered ballots to randomized/rotated ballots in 1991, but New Hampshire did not switch until 2006. Similar to the staggered DiD, I compare the outcomes of the state that changed earlier (Indiana) and the state that changed later (New Hampshire), before and after the change of the ballot order rule (1980-2000). This ensures that my treatment effects are estimated only by comparing the state switching into randomized/rotated ballots to a not-yet-switched state. The identifying assumption is further relaxed and relies on the exact timing of the change of ballot order rule being uncorrelated with changes in the distribution of the population's surnames. Lastly, I adopt an event-study specification to examine whether any pre-trends exist and to show the composition of candidates before and after the adoption of randomized/rotated ballots.

4.2 Immigration, racial/ethnic composition, and the ballot order rule

The main assumption is that states do not change their ballot order rules simply because the population's surname distribution changes. To the best of my knowledge, there have been no media reports indicating that states change their ballot order rules in response to shifts in the distribution of surnames among their population. The New Hampshire Supreme Court declared alphabetically ordered ballots unconstitutional in 2006, so the state changed its ballot order rule. Providing direct evidence to

[16]The main results do not change when performing wild cluster bootstrap inference clustered at the state level (Column (4) of Table A2).

support the identification assumption is difficult because there is no available data on the distribution of surnames in each state over the entire sample period.

Immigrants and minorities are likely to have different surname distributions, so changes in immigration or minority proportions could affect the distribution of surnames in the population. To check the relationship between immigration/minorities and the ballot order rule, I construct state-level data on immigration, non-European immigration, and population growth, measured every five years between 1975 and 2010, based on Burchardi, Chaney and Hassan (2018). I also construct state-level data on racial/ethnic composition, measured every ten years between 1970 and 2020, from the individual files of the Integrated Public Use Microdata Series (IPUMS) samples of five waves of the US census (1970, 1980, 1990, and 2000), along with two waves of the American Community Survey (2010 and 2020). Results are reported in Table A1. As neither immigration nor racial/ethnic composition is correlated with the ballot order rule, it is likely that changes in the ballot order rule are not associated with changes in surname distributions.

5 Results

5.1 Main results

The main results are reported in Table 3. I start the analysis with office-group fixed effects and year-group fixed effects in Column (1), which exploits the within-state variation in the ballot order rule. In this preferred specification, the shift from alphabetically ordered ballots to randomized/rotated ballots decreases the share of candidates with early-alphabet surnames by 6.4% (2.91 percentage points), while increasing the share of candidates with late-alphabet surnames by 16.4% (3.68 percentage points).[17]

In Column (2), I directly control for time-varying state controls to account for the population

[17] As shown in Table 2, in elections where ballots are ordered alphabetically, candidates with early-alphabet surnames make up 45.27% of the share, while candidates with late-alphabet surnames account for 20.21%.

changes in each state over time, including immigration, non-European immigration, population growth, and racial/ethnic composition. While the sample period is shortened to accommodate the availability of the control variables, the main finding remains unchanged. In Column (3), I compare the outcomes of the state that changed earlier (Indiana) to those in the state that changed later (New Hampshire), before and after Indiana's 1992 change of the ballot order rule (1980-2020). Coefficients are estimated less precisely, but the results are still consistent with the story that citizens with late-alphabet surnames are more likely to run for office when they are no longer at a positional disadvantage in the election.

A potential limitation of using the candidate data based on Klarner (2023) is that it may not contain the universe of primary candidates. Therefore, an analysis based solely on Klarner (2023) cannot dismiss the possibility that the change in candidate composition is a consequence of the ballot order effect of the primaries themselves. First, it is worth mentioning that a considerable portion of the state legislative elections in the U.S. are uncontested. For instance, in the November 8, 2022 election, nearly 41% of the 6,278 state legislative seats were uncontested. The presence of uncontested elections limits the degree to which candidate composition can be attributed to the ballot order effect of primaries. Second, in order to demonstrate that the alteration in the composition of candidates is primarily a result of strategic entry in primary races, I replicate Column (3) by utilizing a newly constructed dataset, which contains information regarding all candidates running for state legislative primaries in Indiana and New Hampshire, by digitizing the *Election Report of Indiana (1980-2000)* and web-scraping the website https://nh.electionstats.com/.[18] Reassuringly, the coefficients in Column (4) are very close to those in Column (3), suggesting that the compositional effect is mainly due to strategic entry into primary races.

In the last column, I shift the dependent variable from the share of candidates for each surname

[18]The information available on https://nh.electionstats.com/ is sourced from *The State of New Hampshire Manual for the General Court*. It is worth noting that New Hampshire did not disclose the results of its primary elections for state House of Representatives seats until the year 1986. As a result, the number of observations in Column (4) is lower than that in Column (3).

group to the actual number of candidates in each group, without normalization by the total number of candidates running in a given contest. I thus show that the compositional effect is driven by a higher number of candidates with late-alphabet surnames putting themselves forward for political office. More specifically, while there has been a growing trend towards political participation, there appears to be a disproportionate increase in the number of candidates with late-alphabet surnames. This is relevant for encouraging competitive races, which is at the core of democratic accountability and a concern in particular for state-level politics in the U.S. In recent state legislative elections (November 8, 2022), 2,560 of the 6,278 state legislative seats were uncontested.[19] There is growing concern about the declining number of candidates running for elections, and public perception of candidate quality has been deteriorating. Only 26% of people rate the quality of political candidates as very or somewhat good.[20] To the extent that the widely-used alphabetically ordered ballot discourages the entry of late-alphabet candidates, it potentially weakens the competitiveness of state elections and further limits the accountability of elected legislators.[21]

To investigate whether any pre-trends exist, I adopt an event-study specification to show more transparently the composition of candidates before and after the adoption of randomized/rotated ballots in Indiana and New Hampshire. Given the four-year election cycle, I combine four years as a time window to make the estimate more precise. For the analysis, I include all states (and corresponding periods after 1976) listed in the right column of Table 1, and the baseline group comprises states that have employed randomized/rotated ballots for 13 years or more.[22] Results are reported in Figure A2. I find that the share of candidates belonging to different surname groups in Indiana and New Hampshire quickly converges to the share in states using randomized/rotated ballots, following the change of ballot

[19]Source: https://ballotpedia.org/.

[20]Source: https://www.pewresearch.org/politics

[21]In line with the accountability argument, Konisky and Ueda (2011) find that state legislators who are elected in uncontested contests have higher rates of roll-call vote absenteeism and introduce fewer bills.

[22]Always-alphabetical states are not included.

order rule.

5.2 Robustness checks and heterogeneity

Any approach to dividing the 26 letters of the English alphabet into groups is arbitrary to some extent. In the main specification, I classify the 26 letters into three evenly divided surname groups. I exploit two other specifications in this subsection to show that the finding is not driven by any particular classification.

Specification 1: A finer classification of 26 letters. I evenly divide the 26 letters into five surname groups: group I (A, B, C, D, and E), group II (F, G, H, I, and J), group III (K, L, M, N, O, and P), group IV (Q, R, S, T, and U), and group V (V, W, X, Y, and Z). These five surname groups accounted for 26.4%, 19.4%, 26.2%, 19.4%, and 8.5% of the population in 2010, respectively. Results are reported in Column (1) of Table A2. The share of candidates with top-of-the-alphabet (group 1) surnames substantially decreases when randomized/rotated ballots are used. The negative effect gradually becomes smaller and statistically insignificant for candidates with surnames belonging to group 2 and group 3. Consistent with the baseline results, the share of candidates with an end-of-alphabet surname increased. The share of candidates with a surname belonging to the last surname group only slightly increased because four of the five surname letters only represent a very small portion of the population: V, X, Y, and Z account for 1.75%, 0.04%, 0.63%, and 0.57%, respectively.

Specification 2: The average letter position. I employ a more parsimonious approach and examine how the average letter position changes in response to the ballot order rule. Specifically, I assign each letter a numeric value equal to its position in the alphabet (1 to A, 2 to B, 3 to C, etc.). Then, I calculate the average letter position in each contest and test whether the randomized/rotated ballots are associated with later letter positions. Although this approach obscures heterogeneous responses from candidates with different surnames, it circumvents the arbitrary classification of 26 letters. As shown in Column (2) of Table A2, the average letter position is 0.447 later with randomized/rotated ballots than with alphabetically ordered ballots, which is consistent with the main story.

Heterogeneity: The presence of incumbents. Incumbents typically benefit significantly from their incumbency advantage in re-elections (Gelman and King, 1990; Prior, 2006; de Benedictis-Kessner, 2018). This phenomenon is particularly pronounced within state legislatures, where incumbents secure reelection more than 90% of the time (Niemi et al., 2006). Consequently, the order of candidates is anticipated to have a reduced impact on elections featuring incumbents, while it is expected to exert a more significant influence in contests without incumbents. In line with this expectation, in Table A3, I have observed that the influence of alphabetically ordered ballots on the proportion of candidates with late-alphabet surnames, in the absence of an incumbent, is nearly 2.7 times greater than that in the presence of an incumbent.

5.3 Alternative interpretations

In this section, I discuss two alternative interpretations: political persistence of incumbents, and parties' strategic recruitment.

Alternative interpretation 1: Political persistence of incumbents. Due to the ballot order effect, alphabetically ordered ballots tend to favor the election of state legislators with surnames that appear early in the alphabet (Edwards, 2015). Concurrently, incumbents frequently pursue reelection (Wasserman, 2023). This means that, even without specific knowledge or strategies on the part of the candidates, there can be a natural shift toward a higher proportion of early-surname candidates in the long run. Yet this phenomenon is unlikely to drive the main results for two reasons. First, as observed in Figure A2, the composition of candidate names in rule-changing states rapidly aligns with the distribution found in always-non-alphabetical states. Second, in Column (3) of Table A3, I exclude candidates who have previously secured the same office within the past four years. The results support the notion that strategic responses by candidates are the primary driving forces behind the changing composition of candidates.

Alternative interpretation 2: Parties' strategic responses. Given the role of local party committees in candidate recruitment (see Section 2), parties might internalize ballot order effects and adapt their recruitment strategies by targeting a different pool of potential candidates. I cannot dispositively rule out this possibility. Regardless, however, my findings show that political agents – whether individual or party – respond to ballot order effects, and we thus observe that ballot order rules affect candidate composition. In practice, however, it is unlikely that parties play a significant role in explaining the strategic candidate entry that I document. First, candidates are more likely to be recruited by local party committees when elections are expected to be competitive, but only 20% of state legislative elections nationwide are competitive. Furthermore, competitive elections tend to have higher awareness among voters, so that ballot order effects are expected to be relatively small and thus less likely to explain the large effect observed.

6 Discussion on consequences

In section 5.1, I demonstrate that alphabetically ordered ballots reduce the number of candidates with late-alphabet surnames and can thus lower the competitiveness of elections, given the shortage of candidates running for state legislative positions. In this section, I further discuss potential consequences of the compositional change in the candidate pool: how the choice of electoral rules may affect minority candidate entry, and its association with the legislator behavior.

6.1 Political entry and political selection

Edwards (2015) show that alphabetically ordered ballots are associated with a higher share of elected state legislators with early-alphabet surnames. There are two prominent channels: (1) changes in the probability of winning office, holding the composition of candidates constant; and (2) changes in the composition of candidates, holding the probability of winning office constant. I show that these two channels jointly explain Edwards (2015)'s finding, while approximately 65% of compositional change in elected officials can be attributed to the change in the candidate pool.

I first replicate Edwards (2015)'s finding using my preferred specification in Column (1) of Table 4: randomized/rotated ballots are associated with a lower share of elected state legislators with early-alphabet surnames and a higher share of elected state legislators with late-alphabet surnames. Furthermore, I directly control the composition of candidates in Column (2). The coefficient is 0.939, suggesting that the composition of elected politicians highly reflects the composition of candidates in these down-ballot elections. Moreover, the coefficients of the interaction terms decrease substantially. A back of the envelope calculation suggests that approximately 65% of compositional changes in elected officials can be explained by changes in the pool of candidates. Nevertheless, interaction terms remain important economically, indicating that changes in the probability of winning office (i.e., ballot order effects *per se*) also matter, but in a less significant way.

6.2 Minority candidate entry

Minorities are underrepresented in U.S. politics. Electoral rules can be part of the explanation: for example, the at-large rule and single-member district rule were strategically chosen to maximize minorities' underrepresentation (Trebbi, Aghion and Alesina, 2008; Ricca and Trebbi, 2022). Since racial and ethnic groups tend to have different surnames, might ballot order rules also contribute to minorities' underrepresentation?

Derived from a dataset of 162,253 surnames, each having a minimum of 100 occurrences and collectively representing 90 percent of the US population in the 2010 Census (Comenetz, 2016), I calculated the distribution of surname initials by race/ethnicity. As can be seen in Panel A in Table 5, 43.00% of Whites have an early-alphabet surname and 22.22% of Whites have a late-alphabet surname. Racial minority groups feature a slightly lower percentage of both early-alphabet surnames (41.53%) and late-alphabet surnames (20.91%). The heterogeneity across minority groups, of course, is obscured by studying all minorities together. The share of Hispanics with an early-alphabet surname is similar to the White share (43.29%), but the share with a late-alphabet surname is much lower (17.16%). More strikingly, Asian Americans' distribution of surname initials differs significantly from Whites' distribution: only 30.95% have early-alphabet surnames, while 26.67% have late-alphabet surnames.

In a simulation exercise conducted by Edwards (2014), he demonstrates that in a two-candidate race, Asian American candidates, upon entering the election, are six percentage points less likely to be listed first when competing against a white opponent. This outcome makes them less likely to receive windfall votes, assuming that different races/ethnicities react similarly to the ballot order effect. My finding further implies that alphabetically ordered ballots could even discourage Asian American candidate entry in the first place.[23] Token together, alphabetically ordered ballots may contribute to the underrepresentation of Asian American in politics (Wong and Ramakrishnan, 2023). Nevertheless, it is important to acknowledge that Indiana and New Hampshire – two states that have changed ballot order rules – are predominantly white. This demographic composition inevitably results in a scarcity of minority candidates, particularly Asian American candidates, available to conduct empirical analysis with sufficient statistical power.[24]

[23]The implication also relies on the assumption that the ballot order effects has a uniform effect across different groups, which could be a strong assumption, as voters are likely to have their own racial preferences. For instance, Asian American voters may vote for the first-listed Asian American candidate, even if they are not listed first on the ballot. Investigating whether the effects of ballot order are heterogeneous across racial/ethnic groups is beyond the scope of this paper. However, Fischer (2023) observes that the ballot order effect, if anything, is even more pronounced for Hispanic candidates: a top-tier Hispanic candidate on the ballot is associated with a 7.8 percentage-points increase in Hispanic school board representation. Moreover, a weaker (and more reasonable) assumption is that voters have no racial preference other than for their own race. Under this alternative assumption, voter's cognitive bias still plays a role in (minority) candidate entry, albeit to a lesser extent, as long as a contest does not have candidates from all racial/ethnic groups.

[24]Furthermore, the absence of information on candidates' ethnicity requires using predictions based on first and/or last names to determine their ethnicity, adding additional noise to the analysis.

Furthermore, the choice of ballot order rules has the potential to produce consequential downstream effects on policy outcomes. Fischer (2023) exploits randomized ballots in California and shows that Hispanic candidates' rankings on ballots are highly correlated with Hispanic school board representation, but uncorrelated with other characteristics of the district. More importantly, Hispanic school board members invest more in high-Hispanic schools, resulting in better academic performance and lower teacher turnover. While Fischer (2023) essentially leverages within-state election-level variation in minority ballot positions, which therefore does not necessarily have an aggregate impact, it is crucial to acknowledge that the commonly-used alphabetically ordered ballot could systematically impact minority candidates in a more nuanced manner. The link between minority ballot position and policy outcomes shown in Fischer (2023), combined with my findings that (1) alphabetical order deters the entry of late-alphabet candidates and (2) minorities have distinctive surnames, raises the potential for downstream policy effects of ballot order rules that disproportionately impact minority communities.

Lastly, I assess whether surname distribution is heterogeneous along other key (observable) dimensions. Results are reported in Table 5, from Panel B to Panel E. Surname distributions are very similar in terms of gender. However, there is a clear trend showing that individuals with higher education levels, as well as higher household income/wealth, are less likely to have an early-alphabet surname. For instance, individuals with a graduate degree are 2.8% less likely to have an early-alphabet surname and 9.6% more likely to have a late-alphabet surname, compared to those without a high school diploma (p-values<0.01). These associations naturally raise concerns that alphabetically ordered ballots may potentially discourage some highly qualified candidates from pursuing a career in politics.

6.3 Legislator behavior

Legislators' primary duty is to craft legislation and vote on it; however, roll-call vote absences are common. While abstentions can be strategic in some cases (Jones, 2003), it is generally regarded as a basic form of legislator shirking (Rothenberg and Sanders, 2000), which has been shown to be influenced by electoral incentives, such as reelection eligibility and competitiveness of the election (Konisky and Ueda, 2011; Fouirnaies and Hall, 2022). There are several channels that ballot order method could influence absenteeism rates. First, we observe that alphabetically ordered ballots discourage candidates with late-alphabet surnames from running for office (Column (5) in Table 3), thus reducing competition. Furthermore, the differential impact of the ballot order effect may also result in legislators with late-alphabet surnames abstaining less than their counterparts with early-alphabet surnames in states employing alphabetically ordered ballots. A behavioral adjustment by legislators and/or a change in the composition of elected legislators may result in more accountable legislators with late-alphabet names. However, we do not anticipate different absenteeism patterns among legislators in states using randomized/rotated ballots.

I test these predictions with U.S. State Legislative Roll-Call Dataset (Clark et al., 2009). The publicly available version of this dataset digitizes more than 3.4 million state legislative roll-call voting records across all 99 chambers for the 1999-2000 legislative session. During this period, ballot order does not change, so I simply compare absenteeism rates between states using alphabetical or randomized ballots, while allowing variation based on legislator surnames. The final sample includes 65 chambers, and Table A4 reports the result. I find the result aligns with my predictions. The absenteeism rate of legislators in states with alphabetical ballots is almost double that of legislators in states with randomized/rotated ballots. Given that this is a cross-sectional result, it raises concerns of other unobserved differences across states. To alleviate this concern, I further examine the differences among legislators with different surnames, which are less likely to be correlated with other state-level characteristics. The absenteeism rate of legislators is consistent across surnames in states with randomized/rotated ballots. By contrast, the absenteeism rate in states with alphabetical ballots is significantly different: Legislators with late-alphabet surnames have a 2.12 percentage-point lower absenteeism rate than legislators with early-alphabet surnames, a 12 percent drop. These results provide evidence, albeit suggestive, that ballot order method may contribute to legislator shirking behavior.

7 Conclusion

My paper shows that the pool of candidates entering state legislature elections is influenced by alphabetically ordered ballots, implying that potential candidates are strategically exploiting voters' cognitive bias.

Alphabetically ordered ballots are widely used since they are easily implemented and seem fair. However, given the ballot order effects in low-information elections, alphabetically ordered ballots disproportionately favor candidates with early-alphabet surnames and alter the set of candidates offered to voters. Hence, randomized/rotated ballots are recommended to eliminate the disadvantage against candidates with late-alphabet surnames. In general, when electoral rules are made, we should take into account voters' cognitive biases and candidates' strategic responses.

Underrepresentation of late-surname candidates does not seem to pose a serious threat. However, this underrepresentation is driven by the lower number of candidates with late-surnames and thus reduces the overall number of candidates. This, in turn, dampens the competitiveness of state legislative elections, which is a serious concern in modern U.S. state politics. Furthermore, as documented in section 6.2, if surname distributions are heterogeneous along some key dimensions, alphabetically ordered ballots could potentially dampen descriptive representation, including minority representation.

References

Augenblick, Ned, and Scott Nicholson. 2016. "Ballot position, choice fatigue, and voter behaviour."

The Review of Economic Studies, 83(2): 460–480.

Beazley, Mary Beth. 2013. "Ballot design as fail-safe: an ounce of rotation is worth a pound of litigation." *Election Law Journal*, 12(1): 18–52.

Besley, Timothy, and Stephen Coate. 1997. "An economic model of representative democracy." *The Quarterly Journal of Economics*, 112(1): 85–114.

Black, Gordon S. 1972. "A Theory of Political Ambition: Career Choices and the Role of Structural Incentives." *American Political Science Review*, 66(1): 144–159.

Bovitz, Gregory L., and Jamie L. Carson. 2006. "Position-Taking and Electoral Accountability in the U.S. House of Representatives." *Political Research Quarterly*, 59(2): 297–312.

Brockington, David. 2003. "A Low Information Theory of Ballot Position Effect." *Political Behavior*, 25(1): 1–27.

Brollo, Fernanda, Tommaso Nannicini, Roberto Perotti, and Guido Tabellini. 2013. "The Political Resource Curse." *American Economic Review*, 103(5): 1759–96.

Burchardi, Konrad B, Thomas Chaney, and Tarek A Hassan. 2018. "Migrants, Ancestors, and Foreign Investments." *The Review of Economic Studies*, 86(4): 1448–1486.

Carson, Jamie L., Gregory Koger, Matthew J. Lebo, and Everett Young. 2010. "The Electoral Costs of Party Loyalty in Congress." *American Journal of Political Science*, 54(3): 598–616.

Caselli, Francesco, and Massimo Morelli. 2004. "Bad politicians." *Journal of Public Economics*, 88(3-4): 759–782.

Clark, Jennifer Hayes, Tracy Osborn, Jonathan Winburn, and Gerald C. Wright. 2009. "Representation in U. S. Legislatures: The Acquisition and Analysis of U. S. State Legislative Roll-Call Data." *State Politics Policy Quarterly*, 9(3): 356–370.

Comenetz, Joshua. 2016. "Frequently occurring surnames in the 2010 census." *United States Census Bureau*, 1–8.

Dal Bó, Ernesto, and Frederico Finan. 2018. "Progress and Perspectives in the Study of Political Selection." *Annual Review of Economics*, 10(1): 541–575.

Dal Bó, Ernesto, Pedro Dal Bó, and Jason Snyder. 2009. "Political dynasties." *The Review of Economic Studies*, 76(1): 115–142.

Dal Bó, Ernesto, Frederico Finan, and Martín A. Rossi. 2013. "Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service*." *The Quarterly Journal of Economics*, 128(3): 1169–1218.

Dal Bó, Ernesto, Frederico Finan, Olle Folke, Torsten Persson, and Johanna Rickne. 2017. "Who Becomes A Politician?*." *The Quarterly Journal of Economics*, 132(4): 1877–1914.

de Benedictis-Kessner, Justin. 2018. "Off-Cycle and Out of Office: Election Timing and the Incumbency Advantage." *The Journal of Politics*, 80(1): 119–132.

Diermeier, Daniel, Michael Keane, and Antonio Merlo. 2005. "A Political Economy Model of Congressional Careers." *The American Economic Review*, 95(1): 347–373.

Durante, Ruben, and Ekaterina Zhuravskaya. 2018. "Attack When the World Is Not Watching? US News and the Israeli-Palestinian Conflict." *Journal of Political Economy*, 126(3): 1085–1133.

Edwards, Barry C. 2015. “Alphabetically ordered ballots and the composition of American legislatures.” *State Politics & Policy Quarterly*, 15(2): 171–191.

Edwards, Barry Clayton. 2014. “Race, Ethnicity, and Alphabetically Ordered Ballots.” *Election Law Journal*, 13(3): 394–404.

Faas, Thorsten, and Harald Schoen. 2006. “The importance of being first: Effects of candidates’ list positions in the 2003 Bavarian state election.” *Electoral Studies*, 25(1): 91–102.

Fischer, Brett. 2023. “No Spending without Representation: School Boards and the Racial Gap in Education Finance.” *American Economic Journal: Economic Policy*, 15(2): 198–235.

Fisman, Raymond, Nikolaj A Harmon, Emir Kamenica, and Inger Munk. 2015. “Labor supply of politicians.” *Journal of the European Economic Association*, 13(5): 871–905.

Fouirnaies, Alexander, and Andrew B Hall. 2022. “How do electoral incentives affect legislator behavior? Evidence from US state legislatures.” *American Political Science Review*, 116(2): 662–676.

Gagliarducci, Stefano, and Tommaso Nannicini. 2013. “DO BETTER PAID POLITICIANS PERFORM BETTER? DISENTANGLING INCENTIVES FROM SELECTION.” *Journal of the European Economic Association*, 11(2): 369–398.

Gelman, Andrew, and Gary King. 1990. “Estimating incumbency advantage without bias.” *American journal of political science*, 1142–1164.

Grant, Darren. 2017. “The ballot order effect is huge: evidence from Texas.” *Public Choice*, 172(3): 421–442.

Gulzar, Saad. 2021. “Who enters politics and why?” *Annual Review of Political Science*, 24: 253–275.

Gulzar, Saad, and Muhammad Yasir Khan. 2018. "Barriers to political entry: experimental evidence from local government elections in Pakistan." *Unpublished manuscript. Available at https://www. dropbox. com/s/gi0fadgu8fs6coj/political_entry. pdf.*

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

Ho, Daniel E, and Kosuke 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.

Jones, David R. 2003. "Position Taking and Position Avoidance in the U.S. Senate." *The Journal of Politics*, 65(3): 851–863.

Jun, Byunghill, and Heechul Min. 2017. "What creates heterogeneity in ballot order effects? Evidence from Korea's local elections of education superintendent." *Electoral Studies*, 46: 1–14.

Klarner, Carl. 2023. "State Legislative Election Returns, 1967-2022."

Konisky, David M., and Michiko Ueda. 2011. "The Effects of Uncontested Elections on Legislator Performance." *Legislative Studies Quarterly*, 36(2): 199–229.

Koppell, Jonathan GS, and Jennifer A. Steen. 2004. "The Effects of Ballot Position on Election Outcomes." *Journal of Politics*, 66(1): 267–281.

Kotakorpi, Kaisa, and Panu Poutvaara. 2011. "Pay for politicians and candidate selection: An empirical analysis." *Journal of Public Economics*, 95(7): 877–885.

Meredith, Marc, and Yuval Salant. 2013. "On the causes and consequences of ballot order effects." *Political Behavior*, 35(1): 175–197.

Messner, Matthias, and Mattias K. Polborn. 2004. “Paying politicians.” *Journal of Public Economics*, 88(12): 2423–2445.

Miller, Laura. 2010. “Election by lottery: ballot order, equal protection, and the irrational voter.” *NYUJ Legis. & Pub. Pol'y*, 13: 373.

Moncrief, Gary F, Peverill Squire, and Malcolm E Jewell. 2001. *Who runs for the legislature?* Prentice Hall.

Niemi, Richard G, Lynda W Powell, William D Berry, Thomas M Carsey, and James M Snyder Jr. 2006. “Competition in state legislative elections, 1992–2002.” *The marketplace of democracy: Electoral competition and American politics*, 53–73.

Osborne, Martin J, and Al Slivinski. 1996. “A model of political competition with citizen-candidates.” *The Quarterly Journal of Economics*, 111(1): 65–96.

Pique, Ricardo. 2019. “Higher pay, worse outcomes? The impact of mayoral wages on local government quality in Peru.” *Journal of Public Economics*, 173: 1–20.

Prior, Markus. 2006. “The Incumbent in the Living Room: The Rise of Television and the Incumbency Advantage in U.S. House Elections.” *The Journal of Politics*, 68(3): 657–673.

Ricca, Federico, and Francesco Trebbi. 2022. “Minority Underrepresentation in U.S. Cities.” National Bureau of Economic Research Working Paper 29738.

Rogers, Steven. 2017. “Electoral Accountability for State Legislative Roll Calls and Ideological Representation.” *American Political Science Review*, 111(3): 555–571.

Rothenberg, Lawrence S., and Mitchell S. Sanders. 2000. “Severing the Electoral Connection: Shirking in the Contemporary Congress.” *American Journal of Political Science*, 44(2): 316–325.

Shi, Ying, and John D. Singleton. 2023. "School Boards and Education Production: Evidence from Randomized Ballot Order." *American Economic Journal: Economic Policy*, 15(1): 438–72.

Shue, Kelly, and Erzo F. P. Luttmer. 2009. "Who Misvotes? The Effect of Differential Cognition Costs on Election Outcomes." *American Economic Journal: Economic Policy*, 1(1): 229–57.

Squire, Peverill, and Gary Moncrief. 2019. *State legislatures today: Politics under the domes.* Rowman & Littlefield.

Taebel, Delbert A. 1975. "The effect of ballot position on electoral success." *American Journal of Political Science*, 519–526.

Trebbi, Francesco, Philippe Aghion, and Alberto Alesina. 2008. "Electoral rules and minority representation in US cities." *The Quarterly Journal of Economics*, 123(1): 325–357.

Van Erkel, Patrick FA, and Peter Thijssen. 2016. "The first one wins: Distilling the primacy effect." *Electoral Studies*, 44: 245–254.

Wasserman, Melanie. 2023. "Gender Differences in Politician Persistence." *The Review of Economics and Statistics*, 105(2): 275–291.

Wong, Janelle S., and Karthick Ramakrishnan. 2023. "Asian Americans and the Politics of the Twenty-First Century." *Annual Review of Political Science*, 26(Volume 26, 2023): 305–323.

Table 1: Ballot Order Rules

Alphabetically ordered ballots		Randomized/rotated ballots	
State	Year	State	Year
Alabama	1923	Alaska	1949
Delaware	1915	Arkansas	1969
Florida	1971	California	1975
Georgia	1933	Idaho	1970
Hawaii	I960	Indiana	1991
Indiana	1945-1991	Kansas	1967
Louisiana	1952	Minnesota	1981
Maine	1954	Montana	1971
Maryland	1957	North Dakota	1971
Massachusetts	1894	Nebraska	1960
Nevada	1891	New Hampshire	2006
New Hampshire	1979-2006	New Mexico	1970
Rhode Island	1947-1994	Ohio	1971
South Carolina	1996	Oklahoma	1974
Tennessee	1972	Oregon	1953
Vermont	1912	Texas	1971
		Washington	1966
		West Virginia	1991
		Wisconsin	1970
		Wyoming	1971

Source: Table 1 in Edwards (2015).
Notes: Massachusetts lists incumbents first and then other candidates alphabetically.

Table 2: Preliminary Evidence and Balance

	(1) Share of A-H	(2) Share of I-R	(3) Share of S-Z
	Panel A: Candidates		
Elections with alphabetically ordered ballots	45.27	34.51	20.21
Elections with randomized/rotated ballots	41.99	35.60	22.40
Difference	3.28***	-1.09***	-2.19***
Observations	92,349	92,349	92,349
	Panel B: Registered Voters in 2022		
States with alphabetically ordered ballots	43.02	35.27	21.70
States with randomized/rotated ballots	41.86	35.90	22.24
Difference	1.16**	-0.63	-0.54
Observations	33	33	33

Notes: The unit of observation in Panel A is at the office-year level, while Panel B is at the state level. Data used in Panel A is from *State Legislative Election Returns, 1967-2022* (Klarner, 2023), and Panel B from L2 Voter Data in 2022. Observations are weighted by the number of candidates in office × year cells in Panel A, and by the number of registered voters in state cells in Panel B. Robust standard errors (in parentheses) are reported. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 3: Main Results

VARIABLES	(1) Group Share	(2) Group Share	(3) Group Share	(4) Group Share	(5) # of Cand.
Group1*Random	-0.0291***	-0.0271***	-0.0372	-0.0417	0.0896
	(0.00651)	(0.00659)	(0.0352)	(0.0366)	(0.151)
Group2*Random	-0.00770*	-0.000396	-0.0271	-0.0180	0.138
	(0.00456)	(0.00959)	(0.0339)	(0.0369)	(0.111)
Group3*Random	0.0368***	0.0275**	0.0644**	0.0597*	0.182**
	(0.00475)	(0.0126)	(0.0313)	(0.0321)	(0.0767)
Observations	275,718	184,809	10,461	8,601	275,718
R-squared	0.234	0.289	0.390	0.444	0.310
State Controls	.	X	.	.	.
Office-Group FEs	X	X	X	X	X
Year-Group FEs	X	X	X	X	X
Sample	Full	Full	IN & NH	IN & NH	Full
Sample Period	Full	1976-2010	1980-2000	1980-2000	Full
Data Source	Klarner (2023)	Klarner (2023)	Klarner (2023)	All Primary Candidates	Klarner (2023)

Notes: All columns where the group share is the dependent variable have observations weighted by the number of candidates in office x year cells. In Columns (3) and (4), the sample only includes Indiana and New Hampshire, and the sample period is from 1980 to 2000. The data used in Column (4) was collected by the author, which contains all candidates running for state legislative primaries. Standard errors are clustered at the state-group level, except for Columns (3) and (4), which are clustered at the district-group level due to concerns about the small number of clusters at the state-group level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 4: Composition of Elected Legislators

	(1)	(2)
VARIABLES	Group Share (Legislators)	
Group1*Random	-0.0420***	-0.0155**
	(0.0112)	(0.00661)
Group2*Random	-0.00256	0.00144
	(0.00637)	(0.00409)
Group3*Random	0.0445***	0.0141***
	(0.00779)	(0.00345)
Group Share (Candidates)		0.939***
		(0.00353)
Observations	268,479	268,479
R-squared	0.259	0.692
Office-Group FEs	X	X
Year-Group FEs	X	X

Notes: In all columns, observations are weighted by the number of elected legislators in office × year cells. Standard errors are clustered at the state-group level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 5: Surname Distribution

	Share of A-H	Share of I-R	Share of S-Z
Population	42.19	35.68	22.12
	Panel A: Race/Ethnicity		
White	43.00	34.77	22.22
Minority	41.53	37.56	20.91
Black	42.89	33.70	23.41
Asian American	30.95	42.38	26.67
Hispanic	43.29	39.55	17.16
	Panel B: Gender		
Male	42.21	35.64	22.15
Female	42.23	35.70	22.07
	Panel C: Education		
Less than HS Diploma	42.92	36.52	20.56
HS Diploma	42.68	35.18	22.14
Bach Degree	41.98	35.65	22.37
Grad Degree	41.70	35.77	22.53
	Panel D: HH Income		
$1000-49999	42.62	35.34	22.04
$50000-99999	42.37	35.48	22.15
$100000-149999	41.95	35.97	22.07
$150000-199999	41.61	36.26	22.13
$200000+	41.45	36.33	22.22
	Panel E: HH Net Wealth		
$1-99999	42.69	35.06	22.25
$100000-249999	42.13	35.58	22.29
$250000-499999	41.83	35.80	22.37
$499999+	41.59	35.88	22.53

Source: Calculations made by the author based on all registered voters in U.S. from L2 voter data (2022), expect for Panel A, which is based on the 2010 Census (Comenetz, 2016).

A Online Appendix

Table A1: Immigration, Racial Composition, and the Ballot Order Rule

	(1)	(2)
VARIABLES	Random	
Immigration (1000s)	0.000148	
	(0.000546)	
Non-European Immigration (1000s)	-0.000221	
	(0.000523)	
Population Change (1000s)	5.35e-05	
	(5.66e-05)	
Share of White		0.421
		(0.324)
Share of Black		0.375
		(0.645)
Share of Asian		0.332
		(0.359)
Share of Hispanic		-0.409
		(0.508)
Observations	226	183
R-squared	0.956	0.945
State FEs	X	X
Year FEs	X	X
Sample Period	1975-2010	1970-2020

Notes: Standard errors are clustered at the state level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A2: Robustness Checks

	(1)	(2)	(3)	(4)
VARIABLES	Group Share	Average Letter Position	Group Share	
GroupI*Random	-0.0252***			
	(0.00644)			
GroupII*Random	-0.0121			
	(0.0119)			
GroupIII*Random	0.00605			
	(0.00421)			
GroupIV*Random	0.0292***			
	(0.00492)			
GroupV*Random	0.00201			
	(0.00269)			
Random		0.447**		
		(0.0556)		
Group1*Random			-0.0296***	-0.0291*
			(0.00463)	(0.00657)
Group2*Random			-0.000631	-0.00770
			(0.00429)	(0.00460)
Group3*Random			0.0302***	0.0368*
			(0.00265)	(0.00480)
Observations	459,530	91,906	4,860	275,718
State-Sen./Rep.-Group FEs	.	.	X	.
Office-Group FEs	X	.	.	X
Year-Group FEs	X	.	X	X
Office FEs	.	X	.	.
Year FEs	.	X	.	.
Unit	Office-Year-Group	Office-Year	State-Sen./Rep.-Year-Group	Office-Year-Group
S.E. Cluster Level	State-Group	State (WCB)	State-Group	State (WCB)

Notes: In all columns, observations are weighted based on the count within the unit cell. Standard errors are clustered at the state-group level, with the exceptions of Columns (2) and (4). In these columns, standard errors are clustered at the state level, and wild-bootstrap p-values are reported. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A3: Discussion on Incumbency

	(1)	(2)	(3)
VARIABLES		Group Share	
Group1*Random	-0.0284***	-0.0195***	-0.0288***
	(0.00906)	(0.00558)	(0.00632)
Group2*Random	-0.0177**	0.00229	-0.00448
	(0.00746)	(0.00798)	(0.00690)
Group3*Random	0.0461***	0.0172***	0.0333***
	(0.00318)	(0.00534)	(0.00312)
Observations	82,980	117,249	202,191
R-squared	0.254	0.182	0.154
Office-Group FEs	X	X	X
Year-Group FEs	X	X	X
Sample	Elections w/o Incumbents	Elections w/ Incumbents	Full

Notes: The sample period is from 1972 to 2022. The group share in this table does not include candidates who have previously held the same office within the last four years. In all columns, observations are weighted based on the count within the unit cell. Standard errors are clustered at the state-group level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A4: Legislator Behavior

VARIABLES	(1) Absenteeism
Group2 (β_2)	0.00920
	(0.00635)
Group3 (β_3)	0.00712
	(0.00509)
Group1*Alphabet (β_4)	0.0853***
	(0.0244)
Group2*Alphabet (β_5)	0.0751***
	(0.0213)
Group3*Alphabet (β_6)	0.0641***
	(0.0241)
Baseline Group Mean (Group 1 in Randomized Ballot States)	0.0907
Observations	2,115,214
Party FEs	X
Data Source	Clark et al. (2009)
$H_0 : \beta_6 - \beta_4 = 0$	-0.0212
(P-value)	(0.0464)

Notes: The unit of observation is at bill-legislator level. $Absenteeism$ is a dummy variable indicating whether a legislator abstains from voting by choosing neither "yea" nor "nay" during a legislative process. Standard errors are clustered at the chamber level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

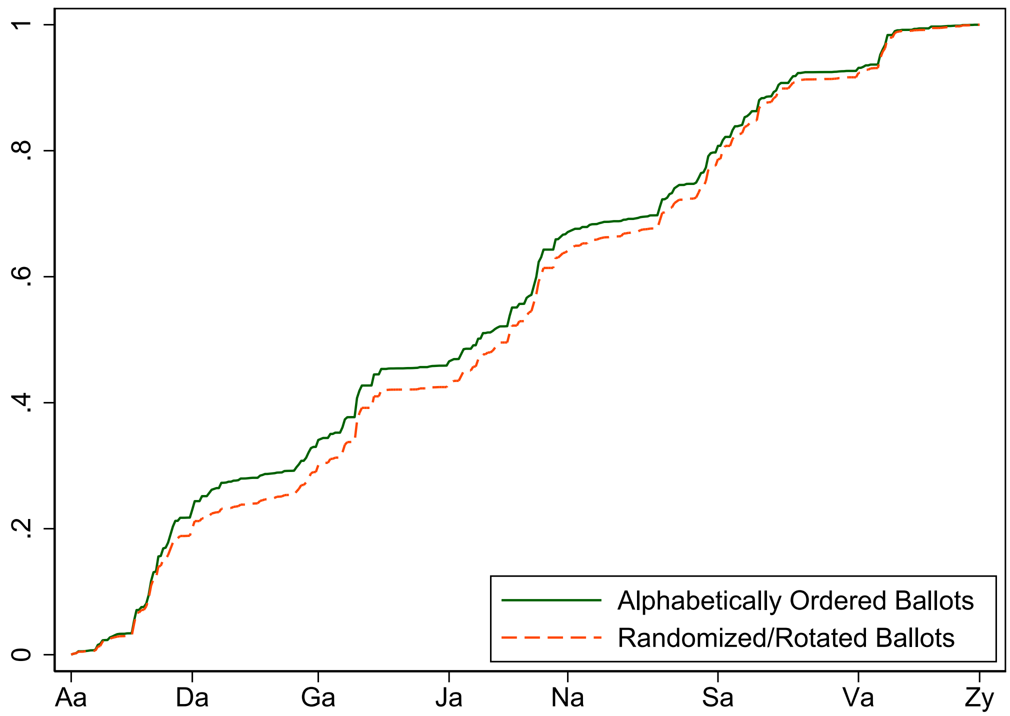

Figure A1: The Distribution of Candidates' Surnames by Ballot Ordering Methods

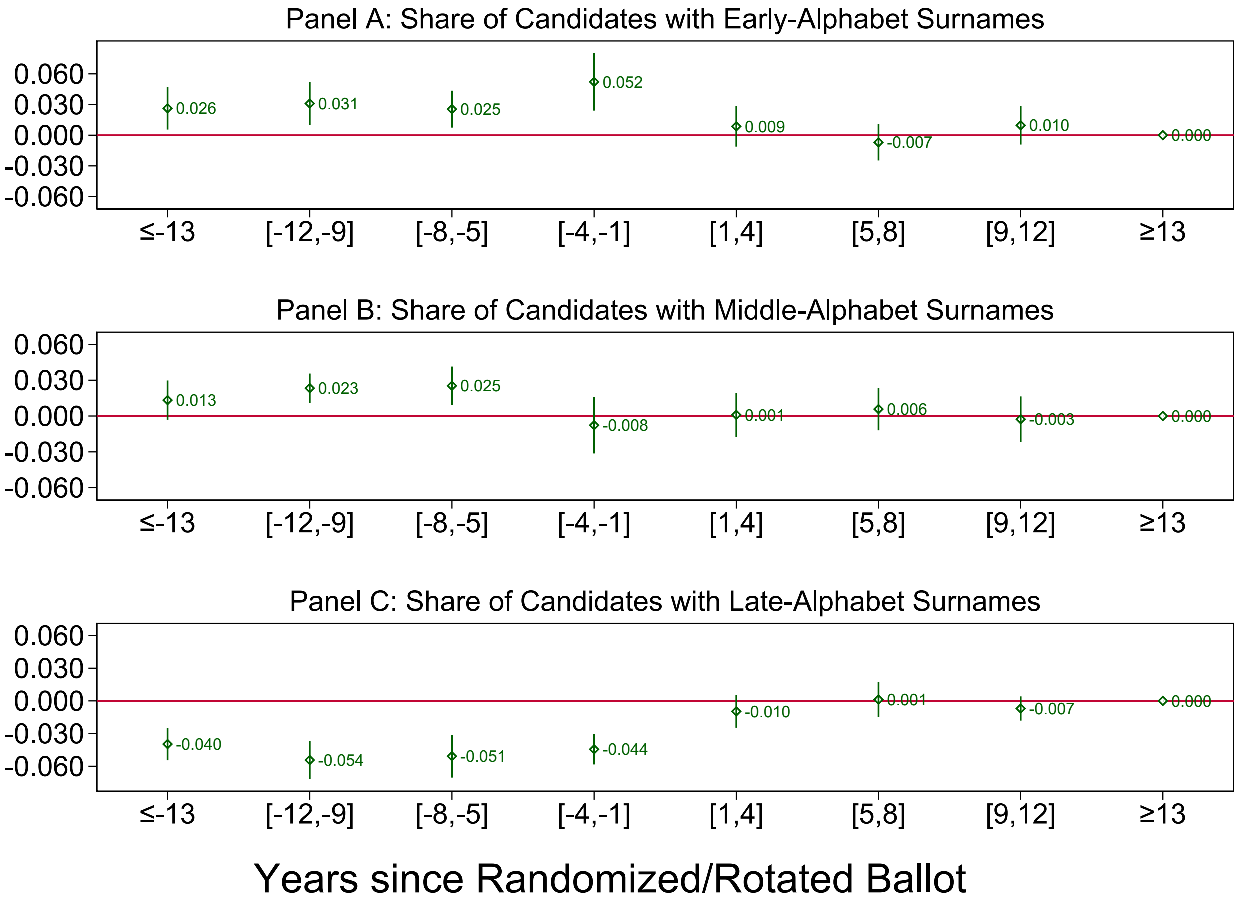

Figure A2: Event Study Plots

Notes: The figure reports the results using the event study specification. The baseline group includes states that have used randomized/rotated ballots for 13 years or more. The graphs depict 95 percent confidence intervals, based on standard errors clustered at the state level.