Strategic Voting and Ballot Order Effects

Mirko Fillbrunn*

Abstract. Substantial evidence suggests ballot order matters, e.g., candidates win 10% more elections just for being listed first on ballots. In this paper, I first establish new empirical patterns of ballot order effects, which I show are inconsistent with both purely behavioral and purely strategic voting models. Second, I build a new model with both behavioral and strategic voters, which gives rise to an added strategic response to the behavioral advantage of candidates listed first. Third, using Californian elections I find ballot order effects caused by strategic voters are empirically important. This suggests previous models overestimate the importance of behavioral voters. (JEL D03, D72)

Elections are fundamental to democracy but remain little understood, in particular when voters deviate from what is typically considered rational. This inadequacy has recently been evidenced by difficulties in predicting the 2016 “Brexit” referendum or the 2016 US presidential election, (Cohn et al. (2016); Cohn (2016)). In this paper, I study the puzzling phenomenon of elections that candidates listed first on ballots enjoy significant advantages such as winning 10% more elections just for being listed first (Meredith and Salant (2013)).

Order effects are pervasive and there is ample evidence that the order in which alternatives are presented can affect how a person chooses from these alternatives. For example, the ordering of questions on surveys may affect a respondent’s answers (Bertrand and Mullainathan (2001)); the first letter of a company’s name can influence its trading activity in stock markets (Jacobs and Hillert (2014)); and academic success may be correlated with an economist’s surname’s initial (Einav and Yariv (2006)). Similarly, many firms in the Yellow Pages have names starting with “AA” or “AAA”. Order effects in elections have been well known for several decades, for instance, a quote from a 1940 U.S. court case reads, “It is a commonly known and accepted fact that in an election [...] those whose names appear at the head of the list have a distinct advantage.”

Accordingly, the academic literature identifies

---

Date: June 7, 2017.

*Harvard Medical School and Massachusetts General Hospital, Department of Medicine, 50 Staniford Street, Suite 830, Boston, MA, 02114 (email: mkfillbrunn@gmail.com)

I greatly thank Jawwad Noor, who gives me insightful advice even on “empirical stuff”. I am grateful to Laurent Bouton, Randy Ellis, Sambuddha Ghosh, Dilip Mookherjee, Juan Ortner, Daniele Paserman, and Ben Solow. I also benefited from conversations with Levent Altinoglu, Sam Bazzi, Felipe Cordova, Michael Manove, Francisco Pino, Marc Rysman, and Fan Zhuo, and I thank all the participants of the SITE Psychology & Economics workshop, the EEA-ESEM meeting in Toulouse, the North American Summer Meeting of the Econometric Society, the Weihnachtskonferenz, and seminar participants at Boston University. I thank Marc Meredith and Yuval Salant for providing their data. All remaining errors were unnoticed by me.

1Elliot v. Secretary of State, 294 N.W. 171, 173 (Mich. 1940), as cited in Miller (2010).
ballot order effects in elections of considerable size (Ho and Imai (2008)), while they are also recognized by policy makers: in the United States, 22 states instituted laws to require some randomization of the ballot ordering of candidates (Miller (2010)).

But why do ballot order effects exist in the first place? This is an understudied question, as the majority of the literature has focused on establishing the existence of ballot order effects (Ho and Imai (2008); Miller and Krosnick (1998)) rather than their underlying cause (Meredith and Salant (2013)). Moreover, the standard explanations are purely behavioral in that voters’ choices are directly affected by ballot order, thereby ignoring potential strategic elements of elections.

To understand this peculiar phenomenon better, I first establish new empirical patterns of ballot order effects. I show that these patterns seem inconsistent with both a purely behavioral model in which voters gain extra utility when voting for first-listed candidates (Miller and Krosnick (1998)); and a purely strategic model in which voters are not intrinsically affected by ballot order but may use it to coordinate (Forsythe et al. (1993)). In contrast, I construct a novel model that is populated by both strategic and behavioral voters. In this model, there is an added strategic response to the behavioral advantage of candidates listed first, which I will explain more in detail below. This strategic response may then amplify the initially purely behavioral ballot order effects and thus gives rise to the possibility of strategic ballot order effects.\(^2\) I show that the distinct implications of this model are supported empirically and estimate it using data from Californian local elections. I find that ballot order effects caused by strategic voters are empirically important, accounting for half of all ballot order effects in the data. This suggests that previous models without strategic elements significantly overestimate the importance of behavioral voters in creating ballot order effects.

To provide more detail, I use a data set containing election outcomes of local Californian elections from 1995 to 2012 where voters elect city council members, school board members, etc. These elections differ in how many candidates run for office and for how many candidates voters may vote, for instance voters may vote for two seats on the school board. The special characteristics of this data set is that the ordering of candidates is randomized before the election, but every voter sees the same, randomly determined ballot ordering. This structure implies that all ballot positions should have the same vote share on average, so any deviation must be a ballot order effect.

I derive new empirical patterns by looking at how the benefit from being listed first varies with two meaningful features of elections: a candidate’s popularity and the number of votes per voter. These patterns may guide us to which model can explain the data. As popularity is, however, not directly observable, I construct a measure of popularity in terms of vote share percentiles, which I will take to indicate a certain level of popularity. For example,

\(^2\)This is reminiscent of models in finance that consider both noise traders and rational traders Campbell and Kyle (1993); Black (1986).
a candidate at the 90th vote share percentile is presumably more popular than a 60th vote
share percentile candidate.\textsuperscript{3}

To construct this measure, I first derive the empirical cumulative distribution function
(CDF) of vote shares across all elections in the data for each ballot position separately. I
then compare the CDF for candidates listed first on the ballot to the average of the CDFs of
candidates not listed first. Without any order effects, these two CDFs should be identical due
to the randomly determined ballot order. However, I find that in three-candidate elections
where voters may elect one candidate to win (\textit{single-vote elections}), these two CDFs are not
only different but follow a particular pattern:

- Pattern 1: the horizontal difference between the two CDFs (the difference in vote
  shares for a given vote share percentile) is lowest for low and very high vote share
  percentiles, and largest for intermediate candidates.

Pattern 1 may be interpreted as unpopular and very popular candidates benefitting the
least from being listed first, and intermediate candidates benefitting most. I find a similar
pattern in four-candidate single-vote elections.

Next, I repeat the analysis for three-candidate elections where voters may elect all but
one candidate in the election (\textit{multi-vote elections}) and find the following pattern:

- Pattern 2: the difference in CDFs is relatively constant, i.e., similar for all vote share
  percentiles.

Pattern 2 suggests that candidates benefit from being listed first independently of their
popularity. Again, a similar result holds for four-candidate elections, this time with with
three votes per voter.\textsuperscript{4}

I next ask what theories can explain these empirical patterns. I compare three different
types of models:

- a purely behavioral voting model,
- a purely rational voting model, and
- a model with both behavioral and strategic voting.

The behavioral model I consider assumes that voters gain extra utility when voting for
candidates listed first (Miller and Krosnick (1998)). This model is difficult to reconcile with
both of these empirical patterns simultaneously. While it may generate hill-shaped ballot
order effects (pattern 1), it does so for both single-vote elections and multi-vote elections,
and thereby cannot generate the relative flatness of ballot order effects in multi-vote elections
with respect to candidate quality at the same time (pattern 2).\textsuperscript{5}

\textsuperscript{3}This relationship holds true in my theoretical voting model as more popular candidates (in terms of pref-

\textsuperscript{4}There are not enough elections to conduct the same analysis with five candidates (59 elections) compared
to three-candidate elections (1,899) or four-candidate elections (1,030).

\textsuperscript{5}Search models as in (Meredith and Salant (2013); Simon (1955); Ho and Imai (2008)) typically imply a
similar relationship of popularity and ballot order effects for candidates listed first and second. However,
Next, consider the voting model that comprises both behavioral and strategic voters. In this model, behavioral voters always vote for the first-listed candidate. Strategic voters then take into account this advantage of the first-listed candidate due to behavioral voting, which may induce them to also vote for the first-listed candidate. To see this, consider for example an election where two candidates A and B are equally desirable for some voter, but another candidate C is likely to win. Suppose candidate A is listed first. Since candidate A gains additional votes from behavioral voters, voting for candidate A lowers the possibility of wasting a vote and is thus the preferred choice for the strategic voter.

A rough intuition for how my model explains the empirical patterns 1 and 2 is as follows: When voters only have one vote, the extra votes from behavioral voters can change whether strategic voters perceive the first-listed candidate as a potential contender and may lead to strategic order effects benefitting the first-listed candidate as seen above. However, this does not happen if the candidate is a sure loser or a sure winner as additional votes are unlikely to change their election chances, so there will be no strategic order effects for these candidates. But notice that these sure losers or sure winners are likely to be unpopular or very popular candidates, respectively, which will then lead to pattern 1.

In contrast, in elections with multiple votes per voter, the additional votes give strategic voters more freedom to break the trade-off between their preferences and a candidate’s probability of winning the election. This allows strategic voters to vote more according to their preferences, which are, by definition, order-independent. Thus, strategic order effects become smaller, implying that behavioral voters create most order effects in multi-vote elections, causing flat order effects as in pattern 2. In sum, the model suggests that pattern 1 is driven by strategic voters, whereas pattern 2 comes predominantly from behavioral voters.

To capture this intuition, I develop a theoretical model that extends the widely-used voting model of Myerson and Weber (1993) to allow for behavioral voters and elections with multiple votes and winners. I focus my analysis particularly on the interaction between behavioral voters and strategic voters and subsequently derive theoretical results to explain the data.

Lastly, a standard voting model without any behavioral voters can be rejected using elections with two candidates. In these elections, non-behavioral voters always vote for their preferred candidate, thus preventing any ballot order effects. But this contrasts with the significant ballot order effects that we find in two-candidate elections.

Next, having presented empirical support for the model combining behavioral and strategic voting, I estimate the empirical importance of the different sources of ballot order effects

---

I find patterns 1 and 2 only for candidates listed first but not for second candidates (relative to other candidates listed later). I consider that model in the appendix.

I do not specify the reason for why these voters vote for the first-listed candidate to keep the pressure of explaining any qualitative results on the strategic voters.
in that model. To do so, I make use of the fact that any curvature leading to pattern 1 must be due to strategic voters in this model. Using this to identify strategic order effects, I find that half of all ballot order effects in terms of actual votes are indeed strategic. In other words, strategic voters amplify the initially purely behavioral vote share advantage to roughly twice the size in single-vote elections with three or four candidates. A structural estimation approach using my theoretical model yields similar results.

An alternative measure of ballot order effects is how often candidates in different ballot positions win an election. In the data, candidates listed first enjoy substantial advantages and win around 12% more often than expected. Simulations of the model suggest that strategic voters, who vote only for likely winners, have a substantially larger impact on a candidate’s chances to win (five times as large) than behavioral voters and thus allow matching the data better.

These findings point to the potential empirical importance of strategic voting when considering ballot order effects and show how strategic complementarity in elections can naturally lead to a rational amplification of behavioral tendencies. It also supplies indirect evidence of strategic voting. In fact, the take on strategic voting here is different from the literature as I investigate the interaction between strategic voting and behavioral voting.

Furthermore, I also estimate the share of behavioral voters in two-candidate elections, in which there is no room for strategic considerations, so any ballot order effects must be due to behavioral voters. This provides an estimate of 2% of behavioral voters in two-candidate elections, which is comparable to the share of behavioral voters I estimate in elections with three or four candidates.

Turning to policy, ballot order effects are typically perceived as detrimental to society as they introduce a (potentially manipulable) systematic advantage for a particular candidate (Miller (2010)). However, my study considers the possibility that ballot order may serve as a coordination device and can thus in fact lead to Pareto-improvements in election outcomes. This provides a new perspective on the normative desirability of ballot order effects.

Nevertheless, in my data set, around 12% of the elections that are won by candidates listed first would have been won by some other candidate without ballot order effects. One can eliminate these ballot order effects by ordering each single ballot randomly for each voter. However, I show that randomization schemes with only a few different ballots (instead of all possible permutations) are sufficient to reduce order effects when strategic voting is present as even a low number of different ballots is enough to distort the strategic coordination. Such cheaper randomization methods may be particularly useful in elections where electronic voting is not available or feasible, for example when voting is done via vote-by-mail or absentee voting.
1. Literature Review

My study is related to three strands of the literature: Firstly, it is connected to the literature that investigates the empirical extent of ballot order effects and the psychological motivation behind it. Parts of my data set stem from Meredith and Salant (2013) which provides evidence for the existence of ballot order effects in Californian local elections. Their study also tests the implications of a theory in which voters are unsure about their favorite candidate (Satisficing), a prominent theory to explain ballot order effects. It shows that a simple model of Satisficing cannot alone explain the empirical pattern based upon a test comparing the advantages of the second-listed candidate and the third-listed candidate. Related, Ho and Imai (2008) examines Californian state elections and finds that while ballot order effects are not significant in general elections, they are influential in primaries and large enough to have changed the election winner in around 12% of primary elections. Their paper considers a cognitive model similar to Satisficing. The study Miller and Krosnick (1998) investigates elections in Ohio and finds ballot order effects of around 2.5% vote share percentages. A study by Augenblick and Nicholson (2016) looks at how choice fatigue affects voters when they are faced with multiple elections on a single ballot. Exploiting variation in the ballot position of elections for different voters, they find that voter fatigue has significant effects both on abstention and subsequently on the election outcome.

My paper is further connected to the literature on strategic voting. On the theory side, I derive my theoretical model from Myerson and Weber (1993), which derives voting equilibria as the set of solutions in which voters believe that only the top two contenders have a reasonable chance of winning. Like mine, their model does not use rational expectations. I extend this model by allowing for behavioral voters and elections with multiple votes per voter. This enables me to look at the interaction between these two types of voters and investigate what theoretical implications this interplay has. Contrarily, Myerson (1998, 2000) defines a completely rational model of voting games, in which the uncertainty about the election outcome stems from uncertainty about the number of voters who participate in the election and which turns out to have comparable properties as my model.

The only other paper to connect strategic voting to ballot order effects, to the best of my knowledge, is the experimental paper Forsythe et al. (1993). Their paper tests how strategic voters coordinate in elections and it allows for the possibility that strategic voter use ballot order as a coordination device. However, it concludes that the relationship is not strong enough to be identified in their experimental setting. Moreover, my model differs in that I allow for behavioral voters and the interplay between different types of voters is the focus of my paper.

My paper differs from other studies in the empirical literature in that I am concerned with the interaction of strategic voting and behavioral voting, whereas most studies are interested in either one of the two aspects separately. However, there are various empirical
studies that identify strategic voting by looking at elections that provide different incentives to vote strategically, which is similar to the identification strategy I use. For example, Fujiwara (2011) uses a regression discontinuity in the vote system assignment in Brazil to show that third place candidates are more often deserted in plurality elections than in runoff-elections, as hypothesized by strategic voting models. Similarly, Cox (1994) considers Japanese elections, in which voters have one vote, but candidates run for more than one seat simultaneously. It derives testable implications of the equilibrium results for these multi-winner elections, namely that trailing and leading candidates will be deserted, and finds that the data indeed support this strategic voting prediction. Kawai and Watanabe (2013) directly estimates the share of strategic voters in Japanese general elections using a structural estimation. It uses variation of preferences along observable characteristics and election outcomes both within and across counties to identify strategic voting in their paper and finds a large share of voters, roughly three out of four voters, to be strategic. Likewise, Degan and Merlo (2009) considers the possibility of identifying non-sincere voting in individual-level data and shows that such identification of strategic voting requires multiple observations of voter behavior. With regard to my model, I observe a given voter only once in my data (or at least I cannot assign them to more than one election). However, I can still identify non-sincere voting as voting exhibits a correlation between vote share and ballot order which cannot be explained by sincere voting alone.

Lastly, my paper is also related to various studies on behavioral biases in political environments. Bisin et al. (2015) investigates how rational actors (politicians) respond to voters suffering from self-control problems, leading to an amplification of the behavioral bias of voters. This amplification mechanism is similar to the one I study as it comes from rational agents (in my case, rational voters) responding to the behavioral tendencies of voters. Levy and Razin (2015) considers voters who are oblivious of a possible correlation between their information sources (correlation neglect) and thus put too much weight on information gained from others rather than their own political convictions. It shows that this correlation neglect can, under certain conditions, in fact lead to better information aggregation. In an empirical context, Ortoleva and Snowberg (2015) investigates the connection between overconfidence in one’s opinion and various political variables of interest and finds that overconfidence has indeed strong predictive power with regard to these variables.

2. **Empirical Patterns of ball order effects**

In this section, I present my data set and define ballot order effects formally. I show that ballot order effects exhibit certain regularities that can help us identify their sources.
2.1. Data. I use data from the California Elections Data Archive (CEDA), a publicly available collection of local Californian elections outcomes. These elections determine county supervisors, city council members, etc., in California’s 58 counties and more than 1,100 school and community college districts. The data set comprises 9,905 different elections from the years 1995 to 2012. Elections are typically held in March or June and in November, together with the accompanying state-wide elections, during even years and only in November during odd years. Elections vary in the number of candidates \( m \) running for office and the number of possible winners and votes per voter, which I denote both by \( w \) as voters can nearly always cast as many votes as there are winners in my data set. All elections are non-partisan, meaning that candidates are not allowed to state their party affiliation if they have any.

The California Secretary of State (CSS) office determines the ordering of candidates on all ballots roughly 80 days before Election Day (but after candidates decided to enter). The CSS office does so by randomly drawing a new alphabet and then ordering candidates by their names. For example, if a new alphabet would be drawn to be A C B instead of A B C, “Charlie” would be listed before “Brown”. This determination ensures randomly determined ballot orderings, which implies that ballot order is uncorrelated with any characteristics of the candidates.\(^8\) All voters see the same ballot in an election and are thus faced with the same ballot ordering. My data on ballot ordering for the years 1995 to 2008 come from Meredith and Salant (2013) where applicable and I collect additional alphabets for the years 2008-2012.\(^9\) I assign ballot ordering manually according to these publicized alphabets.\(^10\)

In the data, slightly more than half of all elections are school board elections (5,843), a third are city council elections (3,377), and the remaining elections (853) are for county or city offices such as mayor, community service district directors, etc. From the raw data set, I exclude elections for which I did not have the ordering of candidates (4,102),\(^11\) the election data were incomplete (46), the number of potential winners did not coincide with the number of votes per voter (228) or the number of votes per voter was equal or higher than the number of candidates (188). I also exclude runoff elections as they provide potentially different incentives for voters (1,034)\(^12\) (Bouton (2013); Bouton and Gratton (2015); Solow

\(^8\)In fact, this is not perfect randomization. For example, candidates named Smith will never be between two Adamses, but this randomization failure would only then be important if candidates responded to it, which is unlikely due to the small chance of it changing the ordering.

\(^9\)My assignment coincides for all but one election with Meredith and Salant (2013) for the overlapping year 2008.

\(^10\)Meredith and Salant (2013) shows that assigning ballot order manually matches the actual ballot order 97% of the time for the San Bernardino County, where discrepancies stem from confusion about what constitutes the last name of the candidate.

\(^11\)The numbers refer to the amount of elections dropped after the previous step was executed.

\(^12\)My data set does not explicitly denote which elections are runoff elections but rather indicates in first round elections whether a candidate proceeded to the second round or potentially could have had. This allows me to easily distinguish first rounds of runoff elections from plurality elections. The second round of runoff elections can then be inferred by comparing the set of candidates in two-candidate elections to elections of the previous election cycle.
(2016)) as well as elections that cross county borders (249). Every ballot contains multiple electoral races. I argue that this reduces the influence of any single election on whether voters turn out, which I use as justification to assume exogenous turnout in my theoretical model later on. However, salient elections might influence a voter’s turnout decision, so I drop mayoral races as they may drive voters to the polls.\footnote{I use month and year dummies in most estimations to pick up differences in turnout due to seasonality but I do not use their interactions. However, including these interactions has typically only marginal effects on the regression results.}

I run my baseline regressions excluding the 5\% of elections with the lowest total number of votes cast (668), i.e., elections with less than 667 total votes. This restriction is motivated by the theoretical model of strategic voting that I use, which I discuss in more details below. In a nutshell, if elections are sufficiently large and the chance of a decisive vote diminishes, the choice of strategic voters can be characterized by relative preferences rather than utilities. Robustness checks suggest that alternative restrictions provide qualitatively similar results.

Below I present a few summary statistics in table 1. Half of all elections are open seat elections, i.e., without an incumbent running for office, and roughly every third candidate is an incumbent, while elections have on average 27,023 votes (median 10,082 votes). Figure 2 displays the number of elections in my data set for the types of elections I use most frequently.

<table>
<thead>
<tr>
<th></th>
<th>count</th>
<th>mean</th>
<th>sd</th>
<th>min</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>County elections</td>
<td>9237</td>
<td>0.056</td>
<td>0.23</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>City elections</td>
<td>9237</td>
<td>0.37</td>
<td>0.48</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>School elections</td>
<td>9237</td>
<td>0.58</td>
<td>0.49</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4 year term</td>
<td>9237</td>
<td>0.96</td>
<td>0.19</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Incumbent</td>
<td>42995</td>
<td>0.32</td>
<td>0.47</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Open seat</td>
<td>9237</td>
<td>0.50</td>
<td>0.50</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Odd year elections</td>
<td>9237</td>
<td>0.29</td>
<td>0.45</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Presidential election cycle</td>
<td>9237</td>
<td>0.34</td>
<td>0.47</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Votes for candidates</td>
<td>42995</td>
<td>5805.4</td>
<td>15882.7</td>
<td>2</td>
<td>613376</td>
</tr>
<tr>
<td>Total election votes</td>
<td>9237</td>
<td>27023.0</td>
<td>56593.1</td>
<td>668</td>
<td>872335</td>
</tr>
<tr>
<td>Year</td>
<td>9237</td>
<td>2004.1</td>
<td>5.35</td>
<td>1995</td>
<td>2012</td>
</tr>
<tr>
<td>Number of candidates</td>
<td>9237</td>
<td>4.42</td>
<td>2.38</td>
<td>2</td>
<td>54</td>
</tr>
<tr>
<td>Number of votes/winners</td>
<td>9237</td>
<td>2.06</td>
<td>0.86</td>
<td>1</td>
<td>15</td>
</tr>
</tbody>
</table>

Table 1. Summary statistics

2.2. Total ballot order effects. I now define (ballot) order effects (BOE) formally and provide a first look at them in the data. Suppose my data contain \(|D|\) elections where each
ballot position $k$ is associated with a vote share $\tau_{kj}$ in election $j$. Let $\bar{\tau}_k$ be the average vote share of ballot position $k$ across all $|D|$ elections in the data

$$\bar{\tau}_k = \frac{1}{|D|} \sum_{j=1}^{\bar{\tau}_{kj}}.$$

Without any ballot order effects and randomly determined ballot order, all ballot positions should on average have the same vote share when the number of elections $|D|$ grows large. This follows directly from the law of large numbers and the independence of ballot order from any candidate characteristics. Thus, define (unconditional or total) order effects as the difference between the first position’s average vote share and that of the other ballot positions

$$b = \bar{\tau}_1 - \frac{1}{m-1} \sum_{k=2}^{m} \bar{\tau}_k.$$

The data suggest that only the first-listed candidate has a systematic advantage, so I restrict attention to the first candidate’s advantage.

Table 3 provides a first look at some statistics of ballot order effects in the data. I regress whether a candidate won the election on an indicator of whether a candidate is listed first (and dummies for election types $(m,w)$). The first candidate on the list wins 7.43% (0.70) more elections than other ballot positions, which is a 12.3% increase in winning chances (1). I also run a similar regression using vote shares as outcome. On average, the first-listed candidate has a 3.25% (0.22) higher vote share than other ballot positions (2). These results are similar to previous estimates of unconditional order effects (Miller and Krosnick (1998)). In contrast, the second or third candidate on the ballot does not gain from the ballot ordering in elections with three or four candidates.

2.3. Conditional ballot order effects. Total ballot order effects are of limited use when determining the source of ballot order effects. Thus, I now look at how ballot order effects change with the “popularity” of the first candidate on the ballot and with how many seats a voter can vote for, fixing the number of candidates in the election. I use these new empirical findings to distinguish between the different voting models in the next section. More graphs, including those with confidence bounds, can be found in the appendix.
Instead of looking at the unobservable popularity of candidates, I compare how order effects vary with a candidate’s vote share percentile relative to other candidates in the data that are in the same ballot position. To do so, I first compute for each ballot position separately the empirical cumulative distribution function (CDF) of vote shares over all elections in the data. For instance, the solid line in figure 2.1 is the CDF $Q_1$ of the first ballot position. I similarly construct the average of the CDFs of all other positions, $Q_k$, $k > 1$, which is the dashed line in that figure. Now, without any ballot order effects, the two CDFs should be identical. However, I find they are not the same. I thus define conditional ballot order effects as the horizontal difference between the CDF of the first candidate and the average CDF of all the other candidates. Formally, denote by $\tau_k(q)$ as the vote share of ballot position $k$ at vote share percentile $q$ so that

$$q = Q_k(\tau_k(q)).$$

Then, the definition of ballot order effects $b(q)$ conditional on vote share percentile $q$ is

$$b(q) = \tau_1(q) - \frac{1}{m-1} \sum_{k=2}^{m} \tau_k(q).$$

Conditional order effects $b(q)$ are mapped in figure 2.2 for three-candidate elections with vote share percentiles on the x-axis for single-vote elections where voters may cast one vote (solid line) and multi-vote elections in which voters may cast a cast for all but one candidate. However, I find they are not the same. I thus define conditional ballot order effects as the horizontal difference between the CDF of the first candidate and the average CDF of all the other candidates. Formally, denote by $\tau_k(q)$ as the vote share of ballot position $k$ at vote share percentile $q$ so that

$$q = Q_k(\tau_k(q)).$$

Then, the definition of ballot order effects $b(q)$ conditional on vote share percentile $q$ is

$$b(q) = \tau_1(q) - \frac{1}{m-1} \sum_{k=2}^{m} \tau_k(q).$$

Conditional order effects $b(q)$ are mapped in figure 2.2 for three-candidate elections with vote share percentiles on the x-axis for single-vote elections where voters may cast one vote (solid line) and multi-vote elections in which voters may cast a cast for all but one candidate. However, I find they are not the same. I thus define conditional ballot order effects as the horizontal difference between the CDF of the first candidate and the average CDF of all the other candidates. Formally, denote by $\tau_k(q)$ as the vote share of ballot position $k$ at vote share percentile $q$ so that

$$q = Q_k(\tau_k(q)).$$

Then, the definition of ballot order effects $b(q)$ conditional on vote share percentile $q$ is

$$b(q) = \tau_1(q) - \frac{1}{m-1} \sum_{k=2}^{m} \tau_k(q).$$

Conditional order effects $b(q)$ are mapped in figure 2.2 for three-candidate elections with vote share percentiles on the x-axis for single-vote elections where voters may cast one vote (solid line) and multi-vote elections in which voters may cast a cast for all but one candidate.

14Meredith and Salant (2013) also consider how order effects with vote share percentiles, meanwhile for different reasons.
Figure 2.1. Data, Order Effects by vote share percentile, 3-candidate election, one vote per voter, one winner

(dashed line). Figure 2.3 provides the same graph for four-candidate elections. Without ballot order effects it must hold that $|D| \to \infty \Rightarrow \bar{\tau}_k(q) = \bar{\tau}_{k'}(q)$ for all $k, k'$.

In my voting model, such vote share percentiles directly relate to popularity as more popular candidates are (expected to) garner a higher vote share and are thus assigned a higher vote share percentile. In other words, candidates at the lower end of vote share percentiles are (expected to be) less popular than those at high vote share percentiles. Tracing the size of conditional ballot order effects along vote share percentiles can thus tell us something about the effect of a candidate’s popularity on ballot order effects.

My empirical findings can then be summarized for three-candidate and four-candidate elections as follows:

**Finding 1:** Order effects are lowest for candidates with low or very high popularity (vote share percentiles) in single-vote elections.

**Finding 2:** Order effects are relatively constant (with respect to popularity or vote share percentiles) in multi-vote elections.

These findings are reflected in figures 2.2, 2.3, and 2.4. There are multiple ways of showing this statistically. For example, order effects at the 75th vote share percentile are significantly
higher than at either extreme in three-candidate single-vote elections, whereas no statistical
differences can be found in multi-vote elections. I provide a table of all conditional order
effects estimated by vote share percentile in the appendix, as well as the graphs with 95%
confidence intervals that are not shown here.

Relatedly, the difference between conditional order effects in single-vote and multi-vote
elections seem to exhibit an inverted U-shape. To see this, I regress the vote share $\tau_{kj}$ of
ballot position $k$ in election $j$ on an interaction of

- an indicator of whether candidate $k$ is listed first in election $j$, $F_{kj}$
- a candidate’s vote share percentile $q_{kj}$ or squared vote share percentile $q_{kj}^2$, and
- a group indicator $g_j$ (single-vote vs. multi-vote elections),

which then comes out to the following regression equation,

$$
\tau_{kj} = C_{kj} + F_{kj} + q_{kj} + q_{kj}^2 + g_j + F_{kj} \cdot q_{kj} + F_{kj} \cdot q_{kj}^2 + g_j \cdot q_{kj} + g_j q_{kj}^2 + \beta_1 F_{kj} \cdot q_{kj} \cdot g_j + \beta_2 F_{kj} \cdot q_{kj}^2 \cdot g_j + \varepsilon_{kj}.
$$
Control variables $C_{kj}$ include dummies for election types, incumbency interacted with election type, whether the election is an open seat election (i.e., no incumbent) and year, month, and county dummies.\(^{15}\) I estimate confidence intervals using the t-percentile bootstrap method with 2,000 repetitions.

Table 4 presents the results. Specification (1) through (3) run the above regression for elections with three candidates, (4) and (5) for four-candidate elections. The point estimates of $\beta_1$ and $\beta_2$ suggest a hill shape in all specifications.\(^{16}\) As can be seen from figure 2.2, the difference in three-candidate elections is significantly hill-shaped for higher vote share percentiles - the estimates of $\beta_1$ and $\beta_2$ imply a downturn at around the 75-th vote share percentile (3).\(^{17}\)

\(^{15}\)For this regression, I drop elections that are held in May as there are only very few of these in my data set.

\(^{16}\)To test whether the estimated effects are due to the construction of vote share percentiles, I manually assign a random ballot ordering (different from the one in the data) and rerun the regressions. However, these estimates yield insignificant results, supporting the validity of the results presented here.

\(^{17}\)This regression uses only for upper half of vote share percentiles. The quadratic function using $\beta_1$ and $\beta_2$ has a downturn around 50%, which is then at 75% for the upper half.
Figure 2.4. Data, Difference of order effects between single-vote elections and multi-vote elections by vote share percentile with 95% confidence bounds, 3-candidate elections.

Table 4. Empirical patterns of ballot order effects.

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Vote Share</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>β₁ Linear term</td>
<td>12.38</td>
<td>23.54</td>
<td>13.96</td>
</tr>
<tr>
<td>Bts 95% Confidence Interval</td>
<td>[-0.52,25.05]</td>
<td>[4.56,42.17]</td>
<td>[2.12,26.84]</td>
</tr>
<tr>
<td>β₂ Quadratic term</td>
<td>-9.23</td>
<td>-21.32</td>
<td>-14.47</td>
</tr>
<tr>
<td>Bts 95% Confidence Interval</td>
<td>[-20.72,2.52]</td>
<td>[-38.77,-5.04]</td>
<td>[-26.70,-2.71]</td>
</tr>
<tr>
<td># of candidates</td>
<td>3</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Vote share percentiles</td>
<td>[50%, 100%]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>5,175</td>
<td>3,501</td>
<td>3,804</td>
</tr>
</tbody>
</table>

I now discuss three different voting models. While all of these models give rise to ballot order effects, they differ in their underlying cause. I first construct a purely behavioral
model. Here, voters gain additional utility when voting for first-listed candidates but they do not have any strategic considerations while voting. Next, I test a standard, purely rational model of voting with strategic voters but no behavioral voters. In this model, ballot order effects only exist if strategic voters use ballot ordering to coordinate between multiple candidates. The last model combines both (simplified) behavioral voting and strategic voting. In this model, ballot order effects are due to behavioral voters and strategic voters, but also due to an added strategic response to the existence of behavioral voters.

In my analysis, I first present the purely behavioral model and then turn attention to the model comprising both behavioral and strategic voting before finally investigating the purely rational voting model.

3.1. Environment. My model is designed to reflect the election design observed in the data. The set-up is identical across all models, while the set of potential types is the only feature that differs between models and is thus responsible for the distinct empirical predictions.

An election consists of $m$ candidates and $w$ possible winners with $m > w \geq 1$. Denote the $m$ candidates in the election by $C = \{c_1, ..., c_m\}$. These $m$ candidates are listed according to a randomly determined ballot ordering $y$ which is the same for all voters in an election. There are $N$ voters who participate in the election. Voters may cast up to $w \geq 0$ votes (including 0, i.e., abstaining) but no more than one vote for any candidate. Denote by $a_i$ an action (i.e., a set of votes) of voter $i$, where $a_{ik} = 1$ if voter $i$ votes for candidate $k$ and $a_{ik} = 0$ if not, and it must hold that $\sum_k a_{ik} \leq w$. Every vote counts equally and the $w$ candidates with the most votes win the election, while ties are broken with equal probability. All voters vote simultaneously without prior communication. Voting is costless which implies that turnout is exogenous and the set of voters that turn out are randomly drawn from the pool of potential types. A voter’s type is private information, but the distributions of types are publicly known.

3.2. Purely behavioral model. I first consider a purely behavioral model in which voters vote according to their preferences and have thus no strategic motivations. However, ballot

---

18Note that $w > 1$ is possible, for example in school board elections, when two seats on the board are up for the taking.

19I assume that candidates enter independently, as my study is concerned with voting behavior and I thus abstract away from strategic candidate behavior.

20I assume, as is common in the literature, that $N$ is sufficiently large so we can make use of the law of large numbers if necessary.

21Throughout I assume that voters can give out as many votes as there are possible winners. This is simply because my data set mostly consists of such elections.

22I justify this by that there are multiple elections on each ballot, which arguably reduces the importance of any given election on the turnout decision. I exclude mayoral elections as these might be salient enough to be significantly related with turnout.
order effects exist because voters gain extra utility when voting for the first-listed candidate while having otherwise order-independent preferences.\textsuperscript{23}

Let voter’s \(i\) utility \(u_i : C \to \mathbb{R}^m\) be a vector consisting of voter’s \(i\) gains from having some candidate win the election.\textsuperscript{24} Let \(\eta > 0\) be the extra utility that voters enjoy when voting for the first-listed candidate, which I assume independent of voter and election. Voters choose an action \(a_i \in \{0, 1\}^m = A\) (i.e., vote) to maximize their overall utility \(U(a_i)\) given by

\[
U(a_i) = \sum_{a_{ik}=1} u_i^y(c_k),
\]

where \(u_i^y(c) = u_i(c) + \eta\) if \(y(c) = 1\) (the first-listed candidate) and \(u_i^y(c) = u_i(c)\) otherwise.

A numerical example of how ballot order effects may look like in such a model is given in figure 3.2. Let me now provide some intuition for this graph: note first that candidates benefit from being listed first if they get a vote when listed first but not otherwise. In single-vote elections, this means candidates benefit only if they are close to being the most-preferred candidate of a voter in terms of the order-independent utility \(u_i\). In that case, they would get a vote when listed first due to the extra utility but not otherwise. Similarly, in two-vote elections, candidates listed first need to be close to being the second-most-preferred candidate of a voter to benefit from their ballot position. This implies that, everything else equal, candidates who benefit most from being listed first in single-vote elections have, on average, a higher mean utility than those that benefit most in two-vote elections. This gives rise to the shape of the numerical example as in figure 3.2.

This model is able to replicate the empirical pattern of ballot order effects in single-vote elections, namely small order effects at either extreme of vote share percentiles. However, it predicts a similar shape of ballot order effects in multi-vote elections, small effects at either extreme, which contradicts empirical finding 2 of relatively constant order effects in these elections. I analyze this model formally in the appendix and provide sufficient conditions for the above intuition to hold analytically. I then test the model predictions in the data and show that it indeed differs from what we observe empirically. Thus, such a model of behavioral voting seems unable to explain our findings.

3.3. Strategic and behavioral model. In this model, there are two types of voters, behavioral and strategic voters.\textsuperscript{25}

\textsuperscript{23}This model has similar implications as a model in which voters vote for the first-listed candidate when indifferent.

\textsuperscript{24}I assume that voters are not indifferent between candidates.

\textsuperscript{25}I do not model sincere voters, i.e., voters who always vote according to order-independent preferences. These voters never create any ballot order effects and are thus not directly important for the qualitative analysis here. However, I consider sincere voters later on when I estimate the model.
Strategic voting and ballot order effects

Figure 3.1. Purely behavioral model, numerical example

Behavioral voting is more simplified than in the previous model. Behavioral voters always vote for the first-listed candidate and distribute their other votes randomly among the other candidates.\footnote{Formally, let $u^b_i : C \times Y \to \mathbb{R}^m$ be the order-dependent utility of a voter for candidates. A behavioral voter then chooses an action (a set of votes) $a_i \in \{0, 1\}^m = A$ to maximize $U(a_i)$

$$U(a_i) = \sum_{a_{ij} = 1} u^b_{ij},$$

where the $j$-th entry of $a_i$ equals one if voter $i$ votes for candidate $j$ and the sum of votes cannot exceed $w$, $\sum_{j=1}^m a_{ij} \leq w$. Further, $u^b_i(c) = \infty$ if $y(c) = 1$ and $u^b_i(c) = \epsilon_c$ if $y(c) \neq 1$ where $\epsilon_c$ are finite, positive random numbers drawn from i.i.d. continuous distributions.}

Let $\lambda$ be the probability of a voter being behavioral.\footnote{I assume that $\lambda$ is fixed throughout all elections. Simulations show that a purely behavioral model where the share of behavioral voters differs by election cannot easily be reconciled with the data.}

A strategic voter maximizes an expected utility that combines an order-independent utility and beliefs about the election outcome. Denote by $p(c_k|a_i)$ the belief that candidate $c_k$ wins given voter $i$ choosing action $a_i$ (i.e., a set of votes). The expected utility can then be written as

$$\tilde{U}_i(a_i) = \sum_{k \leq m} p(c_k|a_i) \cdot u_i(c_k).$$
Strategic voters only care about how their votes change their expected utility. Thus, let $p_i(c_k, c_j)$ be the perceived pivot probability of the event in which candidates $c_j$ and $c_k$ would be ranked in position $w$ and $w + 1$, respectively, and

- candidates $c_k$ and $c_j$ have the same number of votes or
- $c_k$ has one vote less than $c_j$.

Note that only when these events occur, a single vote may decide the election. I assume that $p_i(c_k, c_j) = p_i(c_j, c_k)$ as is common in the literature. I further assume that the pivot probability of any three-way or higher-order tie is perceived to be zero. Henceforth, when I refer to pivot probabilities I mean a function of a candidate pair. We need to use probabilities because the composition of the $N$ voters who end up voting is determined by chance.

The maximization problem of voters can be rewritten using pivot probabilities: in elections with a single vote per voter $w = 1$, voter $i$ chooses $c_k$ given pivot probabilities $p$ to maximize $\tilde{u}_i(c_k|p)$

$$\tilde{u}_i(c_k|p) = \sum_{c_j \in C} p_i(c_k, c_j) [u_i(c_k) - u_i(c_j)].$$

This generalizes to elections with more than one vote $w > 1$, where strategic voters vote for a $w$-subset of candidates to maximize $\tilde{u}_i(a_i|p)$

$$\tilde{u}_i(a_i|p) = \sum_{a_{ij}=1} \sum_{c_j \in C} p_i(c_k, c_j) [u_i(c_k) - u_i(c_j)] = \sum_{a_{ij}=1} \tilde{u}_i(c_k|p),$$

This encompasses strategic abstention if the number of candidates voted for is less than $w$.

3.3.1. Voting equilibrium. My definition of a voting equilibrium extends the commonly used model of Myerson and Weber (1993) to elections with multiple winners and votes per voter and augments it with behavioral voters. As in that model, I do not use rational expectations when specifying the equilibrium. Instead, I only require that beliefs satisfy some consistency conditions with respect to the expected ex-post vote share ranking as outlined below. This is to capture the intuition that voters do not go through potentially complicated pivot probability and utility computations but rather use the expected vote share ranking of candidates to proxy which candidates may win. Moreover, despite my model’s simplicity, most results derived here only depend on typical strategic voting characteristics such as the abandoning of weak candidates and the milder restrictions of strategic voting in elections with multiple votes. This suggests that the qualitative results are robust to the choice of the specific voting model as long as they exhibit these features. Lastly, such a condensed model facilitates exposition and presents more clearly the connection between the theoretical results and the empirical analysis.

---

28 This is approximately satisfied if the number of voters becomes sufficiently large, Myerson and Weber (1993).
29 This is analogous to Myerson and Weber (1993).
To define a voting equilibrium, let me provide a couple of definitions first. Recall that the uncertainty in our model comes from Nature randomly drawing a set of voters to participate in the election. Then, for a given draw \( s \) of voters, let \( t_k(s) \) be the vote share of candidate \( k \) given the votes \( a_i \) for every voter

\[
t_k(s) = \frac{\sum_{i \leq N} a_{ik}(s)}{\sum_{l=1}^{m} \sum_{i \leq N} a_{il}(s)} \cdot w.
\]

I multiply vote shares by \( w \) so that if a candidate gets a vote from every voter, her vote share equals 100\%.

The (objective) uncertainty about the composition of voters in the election induces an election to be a distribution over election outcomes \( t_j(s) \), one for each random draw \( s \). In contrast, there will be only one expected vote share \( \tau(k) = E[t_s(k)] \), where the expectation is being taken over the possible draws of voter compositions.

The upcoming consistency conditions between pivot probabilities and the equilibrium outcome are defined with respect to the expected vote share \( \tau \) rather than with respect to any specific election outcome \( t_k(s) \). In what follows, I therefore often refer to the expected vote share \( \tau \) as “vote share” and omit the term “expected”.

The first consistency condition states that voters must believe that for every winning candidate there is another candidate with whom the candidate has a non-zero pivot probability, i.e., a non-zero probability of tying for the last winning spot.

**Definition 1. (Uncertainty condition):**

If \( \tau(c_j) \) is among the \( w \)-highest vote shares \( \tau \), there exists some \( c_k \in \mathcal{C} \) such that for all \( i \)

\[
p_i(c_j, c_k) > 0.
\]

In effect, this condition implies that there exists some uncertainty in the election, i.e., it is not pre-decided. This implies that voters will not vote for their least preferred candidates just because their vote does not matter (i.e., elimination of weakly dominated strategies). It also implies that strategic voters will vote sincerely, i.e., according to their preferences, in elections with two candidates, which in turn prevents any ballot order effects from strategic voters in these elections.

The next condition is analogous to Myerson and Weber (1993) who only consider single-winner elections, and I extend it to multi-winner elections. Note that the number of winners and the number of votes are interchangeable in my study. Now, to explain the intuition of the next condition, consider an election with three candidates \( A, B, \) and \( C \) and in which only one candidate may win. Suppose that \( \tau_A > \tau_B > \tau_C \), i.e., \( A \) and \( B \) are expected to

\[30\)Note that vote shares can exceed 100\% in multi-vote elections with \( w > 1 \) if some voters abstain and one candidate garners more than \( \frac{1}{w} \) of all votes.
rank first and second respectively. Voters are uncertain about who will win, but it seems natural to believe that the two candidates most likely to tie for first place are A and B. Thus, Myerson and Weber (1993) requires that in equilibrium the pivot probability for these two candidates is larger than any other pivot probability. Reflecting that they are concerned with large elections, Myerson and Weber (1993) further assumes that the pivot probability \( p(A, B) \) is in fact one magnitude larger than any other pivot probability. Consequently, voters only maximize utility over these two candidates in equilibrium.

I want to carry over this intuition to the case when there are multiple winners. To do so, I now define the notion of competing candidates and illustrate it using examples for elections in which there are either a single winner or multiple ones. A formal definition is given after the examples.

For our example election with a single winner and \( \tau_A > \tau_B > \tau_C \), consider candidate A who is most likely to win the election. Among the remaining candidates, the candidate with whom she is most likely to tie for first place is candidate B. Therefore, we will say A’s competing candidate is B. Similarly, B’s competing candidate is A because the scenario in which she ties with A for the win is more likely than the scenario in which she ties with C. Lastly, C’s competing candidate is also A - although candidate C is more likely to tie with B than A, she is more likely to tie for a win with candidate A than B.\(^{31}\)\(^{32}\)

Next suppose that in the election with \( A, B, \) and \( C \) and \( \tau_A > \tau_B > \tau_C \), two candidates may win. Now what matters to voters is not whether a candidate places first or second but rather second or third. In this election, the two candidates most likely to tie for second place are candidates B and C, implying that B’s competing candidate is C, and C’s competing candidate is B. By a similar logic to before, A’s competing candidate is C. Notice that the definition of competing candidates is not necessarily symmetric.

To capture this formally, define \( r : C \times \tau \to \{1, ..., m\} \) to be the ranking of candidates in terms of their expected vote share \( \tau \). If \( r(c_j) \) is less than \( r(c_k) \), it means that \( c_j \) has a higher expected vote share than \( c_k \). Suppose there are no ties in vote share rankings.\(^{33}\)

Then define for any candidate \( c_k \) the competing candidate \( opp(c_k) \) by

\[
opp(c_k) = c_j \quad \text{s.t.} \quad \begin{cases} 
  r(c_j) = w, & \text{if } r(c_k) > w, \\
  r(c_j) = w + 1, & \text{if } r(c_k) \leq w. 
\end{cases}
\]

Figure 3.2 presents the previous two examples.

Now, to preserve the nature of the condition in Myerson and Weber (1993), I require that in equilibrium, voters perceive a candidate’s pivot probability with the competing candidate

---

\(^{31}\)Notice that the scenario where C ties B for the win involves both B and C beating A, which seems less likely.

\(^{32}\)In the appendix I explicitly state assumptions that exclude the possibility of correlations between pivot probabilities originating from preferences.

\(^{33}\)I consider ties in vote share rankings in the appendix.
to be one magnitude larger than with any other candidate. Thus, formally, for any $M \geq 1$ define the following condition:\footnote{Similar to our model, the equilibrium notion in Myerson and Weber (1993) requires this condition for arbitrarily large $M$. So does Myerson (2000) when the number of voters becomes sufficiently large (which is also one possible interpretation of this condition). Contrarily, Kawai and Watanabe (2013) assume a weaker condition requiring only $M = 1$.}

\textbf{Definition 2. (M-pivot probability condition):}

For all candidates $c_k \in C$, 
\[
p(c_k, opp(c_k)) \geq M \cdot p(c_k, c), \text{ for any candidate } c \in C \setminus \{c_k, opp(c_k)\}.
\]

If there are multiple competing candidates of a candidate, then 
\[
p(c_k, opp(c_k)) = p(c_k, opp(c_k))' \text{ for competing candidates } opp(c_k) \text{ and } opp(c_k)'.
\]

Recall that in the first example, $B$ was the competing candidate of $A$ because $A$ is more likely to tie for first place with $B$ than with $C$. The M-pivot probability condition requires that pivot probabilities respect this ordering. Moreover, the condition implies with $M$ becoming arbitrarily large that $p(A, B) \gg p(A, C)$, meaning that the pivot probabilities differ by a order of magnitude, similar to Myerson and Weber (1993). Similarly, $A$ was the competing candidate of $C$, which then implies that $p(C, A) \gg p(C, B)$. In the example with two winners, following a similar logic, it needs to hold that $p(B, C) \gg p(A, C)$ and $p(A, C) \gg p(B, C)$.

The two conditions introduced above are defined by inequalities, so there are multiple sets of pivot probabilities of $p$ that may satisfy these. Hence, to avoid unnecessary multiplicity of equilibria, I define an equilibrium in terms of the actions chosen by the types in the game rather than by the combination of actions and pivot probabilities. This reduces the set of

---

**Figure 3.2.** Competing candidates in three-candidate elections, single-vote and multi-vote.
equilibria in the elections that are relevant for our analysis (single-vote elections and multi-vote elections with \( m = w + 1 \)), as in those, for sufficiently large \( M \), equilibrium actions in response to the set of pivot probabilities will converge towards a unique set.\(^{35}\)

\textbf{Voting equilibrium:}

The voting profile \( a \) is a voting equilibrium if there exist pivot probabilities \( p_i \) such that for each voter type \( i \)

- voters choose actions \( a \) to maximize their expected utility given their pivot probabilities \( p_i \), and
- \( p_i \) satisfies:
  - the uncertainty condition and
  - the M-pivot probability condition for all \( M > \bar{M} \),
    where \( \bar{M} = \max_i \max_{(j,k)} (\min_{(j,k)} |u_{ij} - u_{ik}| (m - 1)) \).

The existence of such voting equilibria is proven in the appendix.

To provide an example of an equilibrium, consider an election with three candidates \( A, B, \) and \( C \) and let there be one vote per voter and one potential winner. Suppose that candidate \( A \) is listed first and consider a set of expected vote shares \( \tau \) for which \( \tau_A > \tau_B > 0 \) and \( \tau_C = 0 \). Let \( \tau_A \) be the proportion of voters who prefer \( A \) over \( B \) (which includes behavioral voters) and similarly let \( \tau_B \) be the proportion of those voters who prefer \( B \) over \( A \). Define pivot probabilities such that \( p(A,B) = 1 \) and \( p(A,C) = 0 \) and \( p(B,C) = 0 \). These pivot probabilities satisfy the equilibrium conditions as is easy to check. Then, all voters, both strategic and behavioral, will only vote for the top-two candidates, which thus generates the expected vote shares \( \tau \) given by \( \tau_A \) and \( \tau_B \). Therefore, the corresponding action profile is an equilibrium, and in this equilibrium the respective vote shares and pivot probabilities are given by \( \tau \) and \( p \).

3.3.2. \textit{Strategic ballot order effects}. In my model, behavioral voters always create ballot order effects because they always vote for the first-listed candidate. I will now show how strategic voters may create ballot order effects (\textit{strategic order effects}) and how these effects vary with a candidate’s popularity. The discussion here conveys the intuition, while a formal analysis can be found in the appendix.

Let us first consider single-vote elections, in which voters give out one vote and only one candidate may win the election. Specifically, consider an election with three candidates \( A, B, \) and \( C \) and consider a strategic voter who prefers \( A \succ B \succ C \). Let the strategic voter’s votes \( a \) be \( a = (a_A, a_B, a_C) \) where \( a_k = 1 \) if the strategic voter votes for candidate \( k \) and \( a_k = 0 \) if not. This voter’s contribution to order effects of candidate \( k \) is then

\[ b(k) = a_k(k \text{ listed first}) - a_k(k \text{ not listed first}). \]

\(^{35} \)In these elections, the M-pivot probability condition implies an ordering of pivot probabilities sufficient to uniquely pin down votes using only preferences rather than cardinal utilities.
Recall that in equilibrium, in single-vote elections, strategic voters maximize their utility only between the top-two candidates given the expected vote share ranking as this is the candidate pair with the highest pivot probability. For notational ease, say that in single-vote elections, the two-highest ranked candidates with respect to the expected vote share $\tau$ are in the running. I now consider multiple cases that differ in whether the first-listed candidate is in the running and show that order effects caused by strategic voters may exist when behavioral voters change whether a candidate is considered in the running. Intuitively, we would expect a candidate that is out of the running to have a low vote share and thus be among the candidates with a low vote share percentile; conversely, candidates who are in the running regardless of being listed first are those in the high vote share percentiles.

(Out/Out): Consider candidate A and suppose candidate A is not in the running both when listed first and when not listed first (“low vote share percentile”). This means that our strategic voter will not vote for candidate A in either case. Thus, her vote does not depend on the ballot ordering,

$$a(A \text{ listed first}) = a(A \text{ not listed first}) = (0, 1, 0)$$

as $A \succ B \succ C$. Therefore,

$$b(A) = a_A(A \text{ listed first}) - a_A(A \text{ not listed first}) = 0,$$

so there are no rational order effects.

(In/In): Next, let candidate A be in the running both when listed first and when not listed first (“high vote share percentile”). This means that our strategic voter will vote for her favorite candidate A in both cases,

$$a(A \text{ listed first}) = a(A \text{ not listed first}) = (1, 0, 0).$$

Therefore, her vote will again not depend on the ballot ordering

$$b(A) = a_A(A \text{ listed first}) - a_A(A \text{ not listed first}) = 0,$$

which again implies that there will be no rational order effects.

(In/Out): But now suppose that candidate A is not in the running when not listed first, but in the running when listed first (“intermediate vote share percentiles”). This means that behavioral voters provide enough support to push candidate A to be among the top-two candidates. In this case, the strategic voter does not vote for candidate A when A is not listed first as A is out of the running,

$$a(A \text{ not listed first}) = (0, 1, 0)$$

as $A \succ B \succ C$, but does so when A is listed first

$$a(A \text{ listed first}) = (1, 0, 0).$$
as $A > B > C$. This means her vote depends on the ballot ordering,

$$b(A) = a_A(A \text{ listed first}) - a_A(A \text{ not listed first}) = 1.$$  

We may thus observe rational order effects for these candidates. Note that the case Out/In is not possible because being listed first can only improve your vote share in our model. Figure 3.3 summarizes these cases.

In summary, in single-vote elections, behavioral order effects are independent of a candidate’s popularity (flat) while strategic order effects are small for very popular and unpopular candidates but stronger for candidates with intermediate popularity. This would explain the empirical pattern in single-vote elections of the previous section.

Next, consider elections where voters can vote for all candidates but one, i.e., $w = m - 1$. Remember that strategic voters base their vote on whether they like a candidate and on whether that candidate is likely to win. In these multi-vote elections, the additional votes give strategic voters more freedom to break the trade-off between these motives and thus allow them to vote more in line with their preferences. And since their preferences are order-independent, strategic order effects become smaller. Therefore, behavioral voters will create most order effects in multi-vote elections, leading to flat order effects as observed in the data.

Nevertheless, it is still possible for strategic ballot order effects to exist in these multi-vote elections: For example, consider an election with three candidates $A$, $B$, and $C$, and suppose that when $B$ is not listed first, we have that $\tau_A > \tau_B > \tau_C$. In this case, voters first vote for their preferred candidate among $B$ and $C$, and then give out their second vote for $A$ if $A > C$. Next, suppose that when $B$ is listed first it holds that $\tau_B > \tau_A > \tau_C$, so behavioral voters provide enough support for $B$ to push her to first place. In that case, voters first vote for their preferred candidate among $A$ and $C$, and then vote for $B$ if $B > C$. While candidates $A$ and $B$ are compared to candidate $C$ in both cases, candidate $C$ is once compared to candidate $A$ and once to candidate $B$. Thus, candidate $C$ may receive a different amount of votes dependent on the ballot ordering, ergo strategic order effects exist.\(^\text{36}\) Nevertheless,

\(^{36}\)Note that for this case to happen, the vote shares of $A$ and $B$ must be sufficiently close. This in turn means that the difference in being compared to $A$ or $B$ will also be fairly similar. This can be used to show that strategic order effects in multi-vote elections are generally bounded above by the share of behavioral voters.
I show formally in the online appendix that this effect vanishes when preferences become sufficiently homogeneous (while strategic order effects in single-vote elections still exist).

In summary, in multi-vote elections, behavioral order effects are flat as function of a candidate’s popularity while strategic order effects are, under some conditions, close to zero. This would explain the empirical patterns in multi-vote elections of the previous section.

Figure 3.4 presents a numerical illustration of the intuition presented before. To generate this figure, I simulate elections as described further below in the next subsection. It showcases the characteristics of strategic ballot order effects: for low enough vote share percentiles, ballot order effects in single-vote elections are zero and rise at vote share percentiles of around a third. From thereon, the advantage continues to rise until vote share percentiles of around two thirds and decreases afterward. In contrast, ballot order effects in multi-vote elections are relatively more constant with only a small drop at low vote share percentiles.

### 3.4. Fully strategic voting model

Lastly, I consider a voting model in which there are no behavioral voters but instead only strategic voters. Thus, all ballot order effects must be due to strategic voters in this model.

The test of this model is straightforward. In elections with two candidates, strategic voters always vote for their preferred candidates if the election is not pre-decided and thus do not create order effects. However, as seen below in table 3.5, there are significant order effects in elections with two candidates, thus contradicting a fully strategic voting model. In other words, some behavioral voters are necessary to create the empirical patterns of ballot order effects. Lastly, this also provides an estimate of around 2% of the share of behavioral voters in two-candidate elections when considering the mixed model with both strategic and behavioral voters of the previous subsection.

---

37 In this simulation, the share of behavioral voters is lower in multi-vote elections than in single-vote elections. This is necessary to generate the overlap of order effects for low vote share percentiles - other, purely behavioral models that I consider require this as well. This is because behavioral voters give out all their w votes in multi-vote elections, so a given share of behavioral voters creates larger order effects in single-vote elections. Alternatively, one could generate this overlap in my model with identical shares of behavioral voters if they were to only give out one vote even if they had multiple votes.

38 The little spike in single-vote elections around 0.35 is because candidates not listed first cannot have a vote share less than $\lambda/(1 - \lambda)$ as otherwise the first candidate will overtake them (if not already ranked first). This spike in the figure is thus due to non-first vote shares jumping from zero to $\lambda/(1 - \lambda)$.

39 Again, it is not necessary to consider sincere voters for the qualitative analysis in this section as sincere voters would never create order effects with order-independent preferences.

40 If strategic voters believed the election to be pre-decided, they might use ballot order to determine their irrelevant vote - although this behavior bears some resemblance to what one might call “behavioral”, which would make the model not fully strategic.
Figure 3.4. Num. example of order effects in model, 3-candidate elections.

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vote Share</td>
<td>Vote Share</td>
<td>Vote Share</td>
<td>Vote Share</td>
<td>Vote Share</td>
<td>Vote Share</td>
</tr>
<tr>
<td>First</td>
<td>2.301</td>
<td>2.066</td>
<td>2.063</td>
<td>1.990</td>
<td>1.990</td>
</tr>
<tr>
<td></td>
<td>(0.637)</td>
<td>(0.566)</td>
<td>(0.570)</td>
<td>(0.566)</td>
<td>(0.566)</td>
</tr>
<tr>
<td>Incumbent</td>
<td>12.30</td>
<td>12.49</td>
<td>16.27</td>
<td>16.27</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.498)</td>
<td>(0.509)</td>
<td>(0.649)</td>
<td>(0.649)</td>
<td></td>
</tr>
<tr>
<td>Full term</td>
<td>0.00784</td>
<td>-0.00545</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0516)</td>
<td>(0.0531)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Open seat</td>
<td>8.217</td>
<td>8.212</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.327)</td>
<td>(0.328)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time &amp; Year dummies</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Office dummies</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>3,976</td>
<td>3,976</td>
<td>3,976</td>
<td>3,976</td>
<td></td>
</tr>
<tr>
<td>R-squared</td>
<td>0.007</td>
<td>0.176</td>
<td>0.178</td>
<td>0.230</td>
<td></td>
</tr>
</tbody>
</table>

Clustered (by election) robust standard errors in parentheses

Figure 3.5. Ballot order effect in two-candidate elections.
4. Empirical evaluation

I now estimate the model consisting of both behavioral and strategic voters and quantify the impact of strategic voters on the advantage of candidates listed first. My empirical analysis can be broken down in three parts, with my main results as follows:

1) Strategic order effects account for half of all ballot order effects in terms of vote shares.

2) Strategic order effects have a 5 times larger impact on who wins the election compared to behavioral order effects.
   - For example: 13% behavioral voters without strategic order effects are equivalent to only 3.2% when strategic voters also create order effects.
   - To compare: I estimate 2% behavioral voters in two-candidate elections

3) A policy to distribute two differently ordered ballots substantially reduces order effects (only) when strategic order effects are present.

4.1. Estimation of strategic order effects - vote shares. I first estimate the size of strategic order effects in terms of vote shares, i.e., how many more votes does the first-listed candidate get from strategic voters just for being listed first. To estimate this, I make use of that behavioral voters in my model do not create a relationship between vote share percentiles \( q \) and conditional order effects \( b(q) \). In contrast, strategic voters create positive order effects for high vote share percentiles. I jointly estimate the threshold at which conditional order effects begin to increase and the resulting size of that increase.

To do so, I first run a regression of vote share on various controls such as incumbency, month, year, and county dummies, whether the election is an open seat election (no incumbent), and whether the office is for 4 years or shorter. After partialling out the effects of these controls on vote share \( \tau_{kj} \), I compute conditional order effects \( b(q) \) using the adjusted vote shares, discretizing vote share percentiles \( q \) into 100 bins.

I then run a non-linear least-squares estimation to estimate a cutoff \( \bar{q} \) at which there is a significant difference between the curvature of two quadratic functions fitted to the curve of conditional order effects. The idea is that strategic and behavioral order effects are responsible for the two separate quadratic functions (compare with figure 3.4)

\[
b(q) = (a_0 + a_1 q + a_2 q^2) \cdot 1\{q < \bar{q}\} + (a'_0 + a'_1 q + a'_2 q^2) \cdot 1\{q \geq \bar{q}\}.
\]

Given the estimate of \( \bar{q} \), I then compute the difference \( \beta \) in order effects between vote share percentiles below and above this cutoff using the regression equation

\[
b(q) = \beta \cdot 1\{q \geq \bar{q}\} + \varepsilon.
\]

\( ^{41} \) I only do this for the dependent variable. I do not partial out the effects on the independent variable (being first or not) which is binary. However, as expected, no regressor comes up significant when regressing ballot position on the controls.
My estimate of strategic order effects is then $\beta \cdot (1 - \bar{q})$.\(^{42}\) I estimate standard errors by bootstrapping this procedure 1,000 times.

Table 4.1 lists the results. I estimate the cutoff to be around the 44th or 56th vote share percentile. Strategic order effects are 1.49% (0.68) in my main specification (1), which accounts for 45% (24.4) of all order effects. I find a share of 1.8% (1.21) behavioral voters. The estimates change marginally when using smaller elections with at least 196 votes (2) or when including time and county dummies (3). Including incumbency dummies reduces the point estimate of strategic ballot order effects to around a 37% (29.6) share of all order effects. These estimates are similar to the ones obtained with a structural approach I present below.

<table>
<thead>
<tr>
<th>VARIABLES %</th>
<th>(1) Vote Share</th>
<th>(2) Vote Share</th>
<th>(3) Vote Share</th>
<th>(4) Vote Share</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta \cdot \bar{q}$: Strategic order effects</td>
<td>1.49</td>
<td>1.47</td>
<td>1.47</td>
<td>1.05</td>
</tr>
<tr>
<td></td>
<td>(0.675)</td>
<td>(0.660)</td>
<td>(0.668)</td>
<td>(0.697)</td>
</tr>
<tr>
<td>Strategic share</td>
<td>45.0</td>
<td>45.2</td>
<td>44.4</td>
<td>37.1</td>
</tr>
<tr>
<td></td>
<td>(24.4)</td>
<td>(28.8)</td>
<td>(24.3)</td>
<td>(29.6)</td>
</tr>
<tr>
<td>$\lambda$: Behavioral order effects</td>
<td>1.8</td>
<td>1.78</td>
<td>1.84</td>
<td>1.78</td>
</tr>
<tr>
<td></td>
<td>(1.21)</td>
<td>(1.14)</td>
<td>(1.17)</td>
<td>(1.10)</td>
</tr>
<tr>
<td>$\bar{q}$: Cutoff</td>
<td>44</td>
<td>44</td>
<td>44</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>(13.8)</td>
<td>(14.1)</td>
<td>(28.54)</td>
<td>(20.61)</td>
</tr>
</tbody>
</table>

Table 4.1:

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Voters &gt; 196</th>
<th>Year, month, and county dummies</th>
<th>Incumbent dummies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observations</td>
<td>1,638</td>
<td>1,701</td>
<td>1,638</td>
</tr>
</tbody>
</table>

Standard errors are computed using non-parametric bootstrap, 1000 repetitions, clustered by election.

4.2. Estimation of strategic order effects - winning chances. So far, our analysis was concerned with vote share percentages. An alternative measure of election performance is whether a candidate wins the election (Meredith and Salant (2013)). For example, in elections with 2 candidates, candidates are expected to win on average 50% of all elections. Order effects may then be measured by the difference in the number of elections won by the first-listed candidate and by other candidates.

In the data, in single-vote elections with 3 candidates, the first-listed candidate wins 8.79% (2.46) more elections than other ballot positions. This is a 17.5% increase in the probability to win over the expected winning chance of one third without order effects in three-candidate elections.\(^{43}\) Now, in our model, behavioral voters are ignorant of the election’s outcome, whereas strategic voters only vote for potentially winning candidates. This suggests that strategic voters’ votes have a higher chance of deciding elections than behavioral voters, which would then imply that strategic order effects have a relatively larger impact on who wins.

\(^{42}\)Alternatively, one could use the estimated coefficients of the first stage to estimate strategic order effects, which yields very similar results.

\(^{43}\)To be clear, the first-listed candidate wins 5.9% more elections than the expected winning chances of one third, and thus wins 8.79% (percentage points) more elections than the other candidates.
To test this, I simulate my theoretical model and estimate the share of rational order effects in terms of vote shares and a parameter governing preferences given this model. Then, using these two estimates, I conduct counterfactual simulations to compute the impact of strategic voters on the first-listed candidate’s winning advantage.

Note that the counterfactual simulation are based on my theoretical model, a less parsimonious model might generalize the model in various dimensions. However, despite its narrow focus, its findings suggest that even with fairly restrictive assumptions the main hypothesis that strategic voters have a larger impact on winning percentages than behavioral voters is borne out. This hypothesis simultaneously also implies that the existence of strategic voters substantially reduces the necessary amount of behavioral voters to replicate the observed ballot order effects. This reinforces the empirical importance of strategic voting in creating ballot order effects.

Now let us start with the structural estimation of my model. I first draw mean utilities (or candidate quality) for candidates $U$ from a standard normal distribution $U \sim N(0, 1)$. Given the mean utilities, I use a logit model to determine how many voters prefer a candidate with mean utility $U_j$ over a candidate with mean utility $U_k$ in a direct head-to-head comparison,

$$P(u_{ij} > u_{ik}) = \frac{\exp(U_j/\mu)}{\exp(U_j/\mu) + \exp(U_k/\mu)},$$

where $\mu$ denotes the scale parameter of the logit distribution.

I use the relationship between $P(\cdot)$ and $U$ to compute vote shares. Recall that in single-vote elections, strategic voters only maximize utility between the top-two candidates. It is thus sufficient to define preferences over candidate pairs to determine strategic vote shares. Behavioral voters always vote for the first-listed candidate, their behavior does not depend on the preference distribution in single-vote elections.

As before, let $\lambda$ denote the share of behavioral voters. I use a Method of Moments approach to estimate the two parameters $\mu$ and $\lambda$, i.e., I match moments generated by the model given some values of the parameters to the respective moments generated by the data. I compute standard errors by bootstrapping the procedure 200 times.

Let me discuss the identification of the parameters. I use conditional order effects $b(q)$ in single-vote and multi-vote elections for all vote share percentiles to estimate the share of behavioral voters in the electorate.

First, strategic voters do not create order effects for candidates with vote shares in low percentiles in single-vote elections and (almost) none in multi-vote elections, so order effects

---

44 One could generalize the model in two important dimensions: one could allow for sincere voters and relax the stringency placed on strategic voters’ beliefs, similar to Kawai and Watanabe (2013). However, as mentioned before, sincere voters only affect the distribution of vote shares rather than directly create ballot order effects. Similarly, strategic voters’ beliefs already allow for strategic ballot order effects as is, although laxer restrictions could generate positive ballot order effects for lower-ranked candidates and negative ballot order effects for highly-ranked candidates (even though the latter would be of minor magnitude).

45 This models follows from assuming an error $\varepsilon_{ij}$ with an extreme value distribution and $u_{ij} = U_j + \varepsilon_{ij}$. 
for these elections must come from behavioral voters. Thus, these order effects identify the share of behavioral voters \( \mu \).

Next, the variation of order effects for larger vote share percentiles in single-vote elections must come from strategic voter. However, strategic ballot order effects only exist when vote shares are sufficiently close without behavioral voters so that behavioral voters would then change the election outcome. Note that the vote share distribution is determined by preferences, so the extent of strategic order effects identifies the preference parameter \( \mu \). For example, if behavioral voters would never change the election outcome, i.e., the vote share distribution would imply that vote shares are far apart from each other, we would not get any strategic order effects and thus no curvature of total order effects.

I estimate the preference parameter \( \mu \) to be \( \hat{\mu} = 0.503, CI_{90} = [-1.13, 0.7] \). In three-candidate elections, this value implies with the logit specification the following preferences: on average, 76% of voters prefer the candidate with the highest mean utility over the one with the second-highest mean utility (and similarly the 2nd over the 3rd); and 90% prefer the candidate with the highest mean utility over the one with the lowest mean utility.

I estimate the share of behavioral voters to be \( \hat{\lambda} = 1.73\%, CI_{95} = [0.05, 2.42] \), which is similar to my previous estimate of behavioral voters in three-candidate elections as well as to the share of behavioral voters in two-candidate elections as shown below. Given this share of behavioral voters, I estimate that strategic order effects \( \hat{b} - \hat{\lambda} \) account for around

\[
\hat{b} - \hat{\lambda} = 52.7\%, CI_{95} = [38.5\%, 109\%]
\]

of all order effects. Note that this uses the estimated share of total order effects in the model.\(^{46}\)

Having estimated the parameters of the model, I now use them to conduct counterfactual simulations to estimate the impact of strategic voters on the first-listed candidate’s winning chances. The results are listed in table 5. Specifications (1) through (3) use the parameter values previously estimated. I find that when only behavioral voters produce ballot order effects (i.e., strategic voters do not react to behavioral order effects), the first-listed candidate wins 0.38\% more elections than candidates in other ballot positions (1) or 0.66\% more

\(^{46}\)For four-candidate elections: These results are, due to the lower number of elections (572 elections vs 1,638 in three-candidate elections), not as statistically powerful. I estimate the share of behavioral voters to be

\[
\hat{\lambda} = 0.8\%, CI = [-0.68, 1.6],
\]

and strategic order effects account for

\[
\hat{b} - \hat{\lambda} = 52.94\% CI = [5.88, 55.72].
\]
elections when replacing strategic voters with sincere voters (2). In contrast, the first-listed candidate’s winning advantage increases to 3.91% when strategic voters also create order effects, which is around 5 times larger (3).

Next, specifications (4) through (6) present what share of behavioral voters is needed to replicate the winning proportions in the data; these results are, roughly speaking, scaled up versions of specifications (1) to (3). A purely behavioral version of my model requires a large share of behavioral voters to replicate the data of around 13% to 20%, (4) and (5), while allowing for strategic voters substantially reduces the need for behavioral voters to just 3.2% (6).

Next, I evaluate a policy to reduce order effects: suppose instead of giving every voter in the election one ballot with the same ballot ordering, we give out two ballots with two different ballot orderings randomly distributed across all voters. This introduces some ballot order randomization but is less costly and easier to implement than providing all possible ballot orderings, particularly in elections in which electronic voting is not available or expensive, for example in vote-by-mail elections. We would expect that such a policy has little impact on order effects in a model with only behavioral voters, but may significantly affect a model with both behavioral and strategic voters. Results are listed in table 6. For these simulations, I set the share of behavioral voters to 3.2% behavioral voters, which is the share of behavioral voters needed in my model to create the observed winning proportions in the data. In the two-ballot policy case, the values listed in the figure are the sum of the respective advantages for each candidate. There is no difference between these two candidates in my simulation.

Specifications (1) and (2) evaluate the effects of the policy in a purely behavioral model with sincere voters and behavioral voters. In these models, I find that the policy has indeed no effect on order effects. In contrast, when considering the model with both behavioral and strategic order effects in specifications (3) and (4), I find that the winning advantage of 7.06% without the policy reduces to −0.2% when the policy is implemented. In terms of vote shares, the advantage reduces from 9.07% without the policy to 1.86% with the policy.

47The logit specification can readily be extended to comparisons of multiple candidates as is necessary for sincere voters:

\[ P(u_{ij} = \arg \max_{k \leq m} u_{ik}) = \frac{\exp(U_j/\mu)}{ \sum_{k \leq m} \exp(U_k/\mu) } \]

I assume that preferences are independent of voter type, so sincere voters and strategic voters share a common preference parameter \( \mu \).

48These specifications also take into account possible nonlinear effects introduced by the vote share distribution.

49Costs here could refer to both administrative costs of election officials or costs that voters might occur, such as a lack of a sample ballot, etc.

50The small differences are due to finite data - the first candidate may just have had a slightly higher mean utility on average.

51Interestingly, the size of the advantages with the policy are smaller in the model with strategic order effects than in the model without. In a nutshell, strategic order effects in my model exist because the behavioral advantage eliminates equilibria in which the first candidate would be pushed up in the rankings,
Thus, the policy has substantial effects on the advantages of the first-listed candidate when strategic order effects also exist.

<table>
<thead>
<tr>
<th>Model %</th>
<th>(1) Pure behavioral</th>
<th>(2) Pure behavioral</th>
<th>(3) Behavioral + strategic</th>
<th>(4) Pure behavioral</th>
<th>(5) Pure behavioral</th>
<th>(6) Behavioral strategic</th>
</tr>
</thead>
</table>

**Parameter values:**
- Share of behavioral voters: 1.37, 1.37, 1.37, 20.0, 13.1, 3.2
- Share of strategic voters: 98.63, 0, 98.63, 80, 0, 96.8
- Share of sincere voters: 0, 98.63, 0, 0, 86.9, 0

**Estimated moments:**
- Winning advantage: 0.38, 0.66, 3.91, 7.02, 7.0, 7.06 (Observed: 7.08)
- Vote share advantage: 1.37, 1.37, 4.83, 20.0, 13.1, 9.07 (Observed: 3.35)

Table 5. Simulation of counterfactuals, three-candidate elections. Logit parameter set to 0.49. 6,000,000 simulated elections.

<table>
<thead>
<tr>
<th>Model %</th>
<th>(1) Pure behavioral</th>
<th>(2) Pure behavioral</th>
<th>(3) Behavioral + strategic</th>
<th>(4) Behavioral + strategic</th>
</tr>
</thead>
</table>

Winning advantage (sum): 1.54, 1.56, 7.06, -0.2
Vote share advantage (sum): 3.2, 3.24, 9.07, 1.86

Two-ballot policy: No, Yes, No, Yes

Table 6. Counterfactual evaluation of two-ballot policy, three-candidate elections. Logit parameter set to 0.49. 6,000,000 simulated elections. Share of behavioral voters set to 3.2%, purely behavioral models includes no strategic voters.

4.3. Discussion of results. My results show that, in my model, strategic voters amplify the initially behavioral vote share advantage and account for roughly half of all ballot order effects in terms of vote shares. Moreover, simulations suggest that behavioral voters alone can generate only a portion of the increased winning chances of the first candidate observed in the data. This points to the potential empirical importance of strategic voting which are no longer feasible equilibria, and the equilibrium selection function I presume in my model says that every equilibrium has equal probability. However, with two candidates gaining such an advantage, the only other candidate that can be overtaken is the candidate not listed first in any election. But since that only happens when that other candidate has a relatively low vote share (as the first candidates need to be able to overtake her in terms of vote shares), we eliminate only equilibria where the non-first candidate has a low vote share. This ultimately reduces her disadvantage.
when considering order effects and shows how strategic complementarity in elections can naturally lead to a rational amplification of behavioral tendencies.

These findings also supply indirect evidence of strategic voting. In fact, the take on strategic voting here is different from the literature as I focus on the interaction between strategic (rational) voting and behavioral voting.

As evidenced by the frequent requirement of ballot order randomization (Miller (2010)), ballot order effects are typically perceived as detrimental to society because they introduce a systematic advantage for a particular candidate based on characteristics determined by the election design rather than the candidate itself; in addition, this endogeneity may also allow external manipulation of elections. While my study presents further evidence on the existence of ballot order effects, at the same time it also provides a new perspective on the normative desirability of a ballot ordering common to all voters. Namely, a common ballot ordering may serve as a coordination device and can thus improve the election outcome for parts or even all of the electorate. For example, a majority that is split between two alternatives can coordinate on the first-listed candidate to overcome the coordination barrier and thus beat out a less-preferred alternative. In this case, the overall welfare may have increased when evaluated with, for example, a utilitarian welfare function. In fact, there are cases in which ballot order effects even allow for a Pareto improvement. Therefore, ballot order effects may indeed be beneficial to society - a view that is, to the best of my knowledge, new to the literature.

Nevertheless, if one wishes to eliminate ballot order effects, my study provides new and cheaper ways to do so. I show that a randomization scheme that only requires two different ballot orderings instead of all permutations already significantly reduces order effects if strategic voters are partially responsible for ballot order effects. For example, numerical simulations suggest that just by using two different ballot orderings, order effects in elections with few candidates may be substantially reduced (by around 79%) or even removed as strategic order effects decrease due to the spreading of behavioral advantages across candidates.

Adding to the literature, my results also suggest that some election systems are more prone to ballot order effects: plurality elections may provide more incentives to vote strategically and therefore increase ballot order effects relative to runoff elections. I find empirical evidence to see this, consider an election with candidates A, B, and C where voters give out a single vote and one winner is elected. Suppose the majority of size $\alpha > 0.5$ prefers $A \succ B \succ C$ and the minority of size $1 - \alpha$ prefers $C \succ A \succ B$. Without ballot order effects, the only two equilibria possible are either A or B winning (in the model introduced above). Now suppose that a $\lambda$ share of voters always votes for the first-listed candidate but still enjoys the same preferences as other voters (i.e., they fail to recognize their preferences or are unaware of it, etc.). If $\lambda > 1 - \alpha$, an election in which A is listed first leads to the unique equilibrium of candidate A winning. Thus, since every voter indeed prefers A over B, removing the equilibrium in which candidate B wins from the set of feasible equilibria is a Pareto-improvement (given that the equilibrium B has indeed positive probability).

Note that these values are for a particular choice of parameters in the model. A different set of parameters would change the quantitative results, but the qualitative intuition would still hold.
evidence supporting this by using data from the same source on runoff elections: candidates listed first benefit significantly less in runoff elections compared to plurality elections, reducing ballot order effects by roughly 45%. Moreover, my model provides a rationale for observing significant ballot order effects even in salient elections where one would expect behavioral ballot order effects to be fairly small as what matters for strategic voters is their perception of the first candidate’s advantage.

5. Conclusion

This paper studies ballot order effects in elections both theoretically and empirically. I use data from Californian local elections where ballot positions are assigned according to candidates’ names and a randomly drawn alphabet before the election. I first present new empirical regularities of ballot order effects. I find that these patterns are difficult to reconcile with a purely behavioral model in which voters gain extra utility when voting for candidates listed first. Similarly, a model without any behavioral voters cannot explain ballot order effects in elections without strategic incentives.

In contrast, I offer a novel theory combining both rational and behavioral aspects of voting that can explain these empirical regularities and allows for ballot order effects even in elections without strategic incentives. In this model, behavioral voters vote for the first candidate on the list, while strategic voters vote for candidates that are both liked by them and likely to win the election. Thus, strategic voters may respond to the advantage of candidates listed first on the ballot and as such amplify the initially purely behavioral ballot order effects, giving rise to the possibility of strategic ballot order effects.

I then turn to the estimation of the new model and show how to separate strategic order effects and behavioral ones in the data, as both favor the first-listed candidate. I consider a specification that allows for any proportion of behavioral voters and let the data select the best estimate of this proportion. As such, I find that strategic voters account for half of the vote share advantage of the first candidate in the data. Moreover, counterfactual simulations suggest that strategic order effects have substantially larger effect on the winning chances of the first-listed candidate than behavioral voters in my model due to strategic voters’ focus on candidates with winning chances.

Order effects are commonly attributed to behavioral or non-standard behavior. In contrast, my study provides evidence that some part of observed order effects could generally be rational if there exists some strategic interaction between individuals. This finding can be extended to other settings. For example, Jacobs and Hillert (2014) finds order effects in financial markets by observing that stocks with initial letters early in the alphabet enjoy higher trading activity. Since investment decisions are likely to exhibit strategic complementarity, individuals could benefit from accounting for these (behavioral) order effects.

54This is not to say that the estimated effect cannot be due to other differences between plurality and runoff elections.
References


Laurent Bouton and Micael Castanheira. One Person, Many Votes: Divided Majority and Information Aggregation, 2012.


Nate Cohn. Why the surprise over 'brexit'? don’t blame the polls. *New York Times*, 2016.


