

On the Causes and Consequences of Ballot Order Effects

Marc Meredith · Yuval Salant

© Springer Science+Business Media, LLC 2011

Abstract We investigate the effect of ballot order on the outcomes of California city council and school board elections. Candidates listed first win office between four and five percentage points more often than expected absent order effects. This first candidate advantage is larger in races with more candidates and for higher quality candidates. The first candidate advantage is similar across contexts: the magnitude of the effect is not statistically distinguishable in city council and in school board elections, in races with and without an open seat, and in races consolidated and not consolidated with statewide general elections. Standard satisficing models cannot fully explain ballot order effects in our dataset of multi-winner elections.

Keywords Ballot order effects · School boards · City council · Satisficing

Long before there were butterfly ballots or hanging chads, political scientists recognized the potential for ballot design to affect electoral outcomes. Numerous studies dating back to at least the 1920's assessed the relationship between ballot order and candidates' electoral performance. While some early studies on ballot

M. Meredith (✉)
Department of Political Science, University of Pennsylvania, 208 S. 37th Street, Philadelphia,
PA 19104-6215, USA
e-mail: marcmere@sas.upenn.edu

Y. Salant
Department of Managerial Economics and Decision Sciences, Kellogg School of Management,
Northwestern University, 2001 Sheridan Road, Evanston, IL 60208, USA
e-mail: y-salant@kellogg.northwestern.edu

order effects were methodologically flawed, recent studies have used increasingly sophisticated methods to identify ballot order effects.¹

Most of these recent studies find that ballot order influences candidates' *vote shares*. Miller and Krosnick (1998) and Krosnick et al. (2004) uncover statistically significant and sometimes substantively large effects of being listed first on the vote shares of the major party candidates in U.S. state and federal elections. Likewise, King and Leigh (2009) estimate that candidates' vote shares increase by about one percentage point from being listed first in partisan Australian House of Representatives elections. Koppell and Steen (2004) and Ho and Imai (2008) estimate that candidates listed first in primary or non-partisan elections for U.S. state or federal offices gain about two percentage points. Koppell and Steen find that the first candidate advantage is even larger in less salient primary elections. On the other hand, Alvarez et al. (2005) and Ho and Imai (2006, 2008) identify only minimal effects of ballot ordering on vote shares of major party candidates in California general elections.

This paper builds on the literature on ballot order effects in two ways. First, we estimate directly to what extent ballot order affects who *wins office* rather than how it influences *vote shares*. Second, we investigate the mechanisms leading to order effects, and establish that standard models of satisficing cannot fully explain order effects in our dataset.

We begin by estimating to what extent ballot order influences who wins office. Clearly, any effect of ballot order on vote shares implies that ballot order will affect representation from time to time. The estimated first candidate advantage was larger than the winning margin in 4 out of the 118 elections studied in Miller and Krosnick (1998), 7 of the 79 contests studied by Koppell and Steen (2004), 7 of the 59 races examined by Ho and Imai (2008) and 40 of the 595 races studied by King and Leigh (2009). These findings are insufficient, however, to estimate the number of races in which ballot order changed representation because one also has to account for additional factors, such as whether the winner actually comes from the first ballot position, how ballot order influences votes share in other positions and so on. For example, in King and Leigh (2009) the winning candidate was listed first in only five out of the 40 races in which the estimated first candidate advantage was larger than the winning margin.

We estimate *directly* the percentage of California local elections in which ballot order affects the winners' identity and thus quantify the extent to which ballot order affects representation. This estimation is policy-relevant when weighing the costs and benefits of rotation and randomization schemes. On one hand, rotation and randomization schemes are costly to administrate and cause voter confusion. On the other hand, representation should not be decided by the order in which candidates are listed on the ballot. Thus randomization and rotation schemes should be used only when ballot order significantly affects representation.

¹ Miller and Krosnick (1998) criticize most of the 28 papers on ballot order effects that they survey for using non-random variation in candidates' positions that is potentially correlated with pre-existing difference in the support for candidates.

To estimate the effect of ballot order on representation, we use the quasi-random assignment of candidates to ballot positions in California city council and school board elections. After the California local election entry deadline has passed, the California Secretary of State draws a random order of the alphabet according to which candidates are listed on all ballots. Because of this quasi-random assignment of candidates to ballot positions, the distribution of candidates' qualities and other characteristics—and consequently the number of winners—is expected to be similar across ballot positions. We thus calculate the *expected* number of winners from each ballot position absent order effects and compare it to the *actual* number of winners from each ballot position. This comparison informs us how ballot order affects representation.

We find that candidates listed first on the ballot are between four and five percentage points more likely to win office than expected absent order effects. This estimate implies that the first listed candidate wins roughly one out of every ten elections that he or she would otherwise lose. The first candidate advantage is similar across contexts: the magnitude of the effect is similar in city council and in school board elections, in races with and without an open seat, and in races consolidated and not consolidated with statewide general elections. However, the first candidate advantage is larger in races with more candidates and is more beneficial to higher quality candidates. The first candidate advantage comes primarily at the expense of candidates listed in the median ballot position who are between two and three percentage points less likely to win office than expected absent order effects.

After analyzing the effect of ballot order on representation, we investigate what the data can tell us about the mechanisms leading to order effects. A better understanding of these mechanisms may inform us about the effectiveness of different balloting systems.

The leading model used to explain ballot order effects is satisficing (Miller and Krosnick 1998; Brockington 2003; Koppell and Steen 2004). A satisficing voter has in mind a fixed aspiration level that reflects the minimal quality she considers satisfactory. She evaluates the candidates according to the ballot order and selects the first candidate that meets or exceeds her aspiration level without considering candidates listed further down the ballot. One possible rationalization of satisficing behavior is that the cognitive costs required to distinguish between acceptable candidates are smaller than the expected benefits of doing so (Simon 1955).²

We exploit the fact that candidates are randomly assigned to ballot positions in our dataset and that many elections in the dataset are multi-winner elections to examine the satisficing hypothesis. Similar to a single-winner race, in a K -winner race a satisficing voter examines the candidates in order and selects *up to* K candidates who are satisfactory. If such a voter casts fewer than K votes, it means that she examined *all* the candidates on the ballot and found only $M < K$ of them to be satisfactory. Because of the random assignment of qualities to ballot positions in

² For example, Ho and Imai's (2008) model of costly sequential ballot search results in voters that endogenously satisfice because the expected benefit of evaluating the next candidate on the ballot is smaller than the cognitive cost of doing so.

our dataset, such a voter is equally likely to cast a vote for any position. Therefore when aggregating over such voters, their behavior should not generate order effects. If, on the other hand, such a voter casts K votes, then she potentially “runs out of votes” before considering all candidates and is therefore more likely to vote for candidates that appear at the top of the ballot. When aggregating over such voters, their behavior should generate two patterns of order effects.

First, because it is possible that a voter casts all her K votes after having considered only the first K candidates, candidates in position K should perform better than candidates in position $K + 1$ in a K -winner race. Second, because in a multi-winner race voters never “run out of votes” when considering the second candidate on the ballot (unlike in single-winner election), it is expected that candidates listed first and candidates listed second will be equally likely to win.

These two predictions only require that voters are aware of the possibility of casting K votes in a K -winner race. We provide some suggestive evidence that this is indeed the case by examining the number of votes cast by individual voters in a particular race in our dataset.

We find no evidence supporting the two predictions above. Candidates in the second ballot position *do not* perform better than candidates in the third ballot position in two-winner as compared to three-winner elections. Moreover, candidates listed second are about seven percentage points less likely to win office than candidates listed first. Thus, assuming voters are aware of the possibility of casting multiple votes, satisficing with a fixed aspiration level cannot fully explain the patterns of order effects in the data.

We conclude by examining the downstream consequences of ballot order on subsequent election outcomes. Because the vote shares of candidates in the first ballot position increase from a source other than voters’ underlying preferences, a concern is that they are lower quality representatives than representatives elected from other ballot positions. We indirectly test this hypothesis by comparing the performance of candidates elected from the first ballot position and from other ballot positions in the next election. We find that candidates elected from the first ballot position are more likely to lose in the next election, which is consistent with candidates elected from the first ballot position being of lower quality.

Data

Testing whether ballot order affects the results of elections requires that we use a different source of variation in candidates’ ordering than what is used in most other recent papers on ballot order effects. Rather than focusing on how *within* race variation in candidates’ ballot positions affects vote shares, we look at how the likelihood of winning office in different ballot positions varies *across* races. Two necessary properties to implement this identification strategy are a large number of elections for which both the ballot ordering and outcomes are known and the random assignment of candidates to ballot positions. California local elections satisfy both of these properties. California’s 500 incorporated cities and 1,100 school and community college election districts provide a large number of city

council and school board elections from which we can observe outcomes. In addition, unlike many states that list candidates for local office in alphabetical order, ballot ordering in California is determined via a lottery system.³ After the deadline has passed for candidates to enter the race, the California Secretary of State office draws a random ordering of the alphabet according to which candidates names are ordered on the ballot. In municipal and school board elections, candidates are listed according to this ordering across all precincts.⁴

California local election results come from the California Elections Data Archive (CEDA). The CEDA archives election results from 1995 to 2008 for county, city, community college, and school district elections in over 6,000 jurisdictions throughout California.⁵ The CEDA contains information on candidates' names, incumbency status, vote totals, and an indicator about whether a candidate was elected. Information on candidates' names, combined with information about the outcomes of the California alphabet lotteries, allows us to reconstruct the order in which candidates appeared on the ballot without actually seeing the physical ballots.⁶ We focus on 7,846 competitive city council, community college, and school district elections for which we are able to determine the ballot order.⁷

Information on the outcomes of the California alphabet lotteries for the 1996, 1998, 2000, and 2002 elections is taken from Ho and Imai (2008). Lottery outcomes for statewide primary and general elections in the remaining even years and November consolidated elections in odd years are obtained from various county election officials.

We test whether the ballot position assignment mechanism produces balance by examining the percentage of incumbents assigned to different ballot positions. More specifically, we test the null hypothesis that the percentage of incumbents assigned to the first, second, and last ballot positions is equal to the expected 34.2 percentage points implied by random assignment. We find that 35.0% of first-listed candidates ($p = .181$, two-tailed), 34.4% of second-listed candidates ($p = .705$, two-tailed)

³ Alphabetical ordering is problematic for identifying the causal effect of ballot order on electoral outcomes because candidates with certain last names may also be likely to receive more votes than others (Miller and Krosnick 1998).

⁴ In contrast, the ordering of candidates is rotated across assembly districts in elections for federal and state offices.

⁵ These races are non-partisan. Previous work generally finds larger ballot order effects in non-partisan races, although Meredith and Salant (2007) find that the first candidate advantage is similar in partisan and non-partisan Ohio city council elections.

⁶ We verified that candidates' constructed ballot positions matched the actual ballot positions in San Bernardino County's Statement of the Votes 97% of the time. Inconsistencies mostly resulted from errors in the CEDA data about which portion of the name constituted the last name.

⁷ We observe 8,348 competitive city council, community college, and school district races. We partition those races according to "types", where two races are of the same type if the number of candidates competing for office and the number of winners are identical in both (e.g. in both races seven candidates compete for two spots). We restrict attention to election types in which there are 75 or more races for estimation purposes. In the 502 excluded elections, we find that candidates listed first win 8.2 percentage points (s.e. 2.5 percentage points) more races than expected absent order effects.

Table 1 Number of election winners by ballot position in California city council and school board elections from 1995 to 2008

Election type	<i>N</i>	Expected	Ballot position											
			1	2	3	4	5	6	7	8	9			
1 of 2	1,717	858.50	912	805										
1 of 3	609	203.00	221	200	188									
1 of 4	219	54.75	68	49	50	52								
1 of 5	83	16.60	24	16	18	9	16							
2 of 3	1,108	738.67	763	721	732									
2 of 4	816	408.00	430	407	397	398								
2 of 5	523	209.20	247	213	200	174	212							
2 of 6	276	92.00	104	92	96	82	82	96						
2 of 7	164	46.86	58	43	56	37	40	45	49					
2 of 8	78	19.50	26	15	19	19	16	21	23	17				
3 of 4	708	531.00	557	553	511	503								
3 of 5	557	334.20	372	338	319	319	323							
3 of 6	462	231.00	284	218	220	204	221	239						
3 of 7	260	111.43	137	101	108	92	106	110	126					
3 of 8	167	62.63	80	72	53	51	62	59	61	63				
3 of 9	99	33.00	41	46	27	35	30	32	29	28	29			

Note X of Y indicates that the election is selecting X winners out of Y candidates

and 33.8% of last-listed candidates ($p = .464$, two-tailed) are incumbents.⁸ We also test the joint hypothesis that incumbents are distributed independently across all positions using simulation, and find that the actual distribution of incumbents to ballot positions is near the middle of the simulated distribution that would be expected under the null that incumbents are randomly assigned to ballot positions.⁹

Results: Winning Elections and the Effect of Ballot Position

Table 1 presents the number of winners by ballot position across all elections in our dataset. Table 1 indicates for example that there were 523 races in which five candidates competed for two positions. If winning office were independent of ballot order, then we expect that on average $2/5 * 523 = 209.2$ winners would come from

⁸ There is some overlap here because candidate listed second in two candidate elections are also listed last.

⁹ We first calculate the absolute value of the difference between the actual number of incumbents in each ballot position and the expected number of incumbents in each ballot position if incumbents were uniformly distributed across ballot positions. We then compare this *actual value* to 1,000 simulated values generated by randomly assigning incumbents to ballot positions. We find that the actual value falls in the 48th percentile of the simulated distribution, which is consistent with the alphabetical lottery approximating the random assignment of candidates to ballot positions.

each ballot position. We instead observe 247, 213, 200, 174, and 212 winners from the first, second, third, fourth, and fifth ballot positions respectively. We observe a similar pattern of more winners than expected from the first position and fewer winners than expected from middle ballot positions across most of the rows in Table 1, suggesting that ballot order is an important determinant of winning office.

To formally test these observed patterns, we derive the expected distribution of winners by ballot position under the null hypothesis of no order effects. Let $Y_{p,j}$ be a Bernoulli random variable that is equal to one if the candidate listed in ballot position p in election j wins office, and zero otherwise. If K_j of the N_j candidates in election j win office, then under the null hypothesis the probability of any given candidate winning office in election j is $\pi_j = K_j / N_j$. The statistic $\hat{\theta}_p$ defined in Eq. 1 is our estimate in the T elections we observe of the percentage point change in the probability of winning office in ballot position p relative to the expected probability under the null of no order effects.

$$\hat{\theta}_p = \frac{1}{T} \sum_{j=1}^T (Y_{p,j} - \pi_j) \tag{1}$$

Equation 2 gives the distribution of $\hat{\theta}_p$ implied by the Central Limit Theorem under the null hypothesis.

$$\hat{\theta}_p \sim N \left[0, \frac{1}{T^2} \sum_{j=1}^T \pi_j (1 - \pi_j) \right] \tag{2}$$

Candidates listed first are expected to win in 50.3% of the races in our dataset absent any order-effects, i.e. $\frac{1}{T} \sum_{j=1}^T \pi_j = 0.503$. Figure 1 shows that candidates listed first actually win 55.1% of the races, which is 4.8 percentage points ($\sigma = 0.5$ percentage points) more races than expected absent order effects.¹⁰ This implies that the first listed candidate wins about 1 in 10 elections (i.e., $[55.1 - 50.3]/[100 - 50.3] \approx 0.1$) that he or she would lose without the benefit of order. Even in two-candidate one-winner races candidates listed first win 53.1% of the races, which is significantly different from 50.0 at the $p = 0.01$ level, two-tailed. Figure 1 also illustrates that the first candidate advantage comes primarily at the expense of candidates listed in the median ballot position,¹¹ who are 2.5 percentage points less likely to win than expected absent order effects ($\sigma = 0.5$ percentage points).

Results: Moderators of the First Candidate Advantage

Having found evidence of position effects, we next investigate the moderators of the first position advantage. We test whether incumbents and non-incumbents benefit

¹⁰ Because the effect of order on the probability of winning office varies by election type, θ_p is defined conditional on the observed percentages of election types.

¹¹ In an election with an even number (N) of candidates, the median ballot position is defined as $N/2 + 1$.

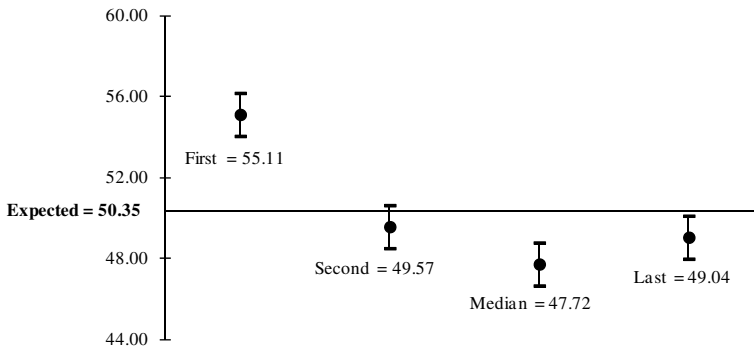


Fig. 1 Frequency of winning office by ballot position in all elections ($N = 7,846$, 95% confidence interval in bars, $\sigma = 0.54$). *Note:* In an election with an even number (N) of candidates, the median ballot position is defined as $N/2 + 1$

differently from being listed first on the ballot. We also examine how the mere presence of incumbents on the ballot interacts with order effects. Incumbency status is generally observable to voters in California because candidates’ occupations appear on the ballot and nearly all incumbents denote themselves as incumbents. If the first candidate advantage results from voters using the order to resolve their indecisiveness among candidates, we expect that observing incumbency status will have a similar role and hence attenuate order effects.

We also compare the magnitude of the first candidate advantage in elections that are concurrent and not concurrent with statewide general elections. If the first candidate advantage results primarily from uninformed voters mobilized to the polls by other races on the ballot, we expect that the first candidate advantage will be larger in higher turnout general elections. Finally, we investigate whether the first candidate advantage differs between city council and school board elections.

To cope with the interdependence in outcomes for candidates in the same race, we use a structural estimation approach to estimate the moderators of the first candidate advantage. We model a candidate’s *attractiveness* as a linear function of the candidate’s ballot position, the candidate’s incumbency status, and his or her unobserved intrinsic quality component. Assuming that the K most attractive candidates win in a given race, we estimate this model using maximum likelihood, and use the estimated parameters and simulation to investigate the interaction between ballot position and the above moderators.

Formally, let $a_{p,j}$ denote the attractiveness of the candidate located in position p in race j . Let $t(j)$ denote the “type” of race j , where two races are of the same type if the number of candidates and the number of winners are identical in both. Equation 3 models $a_{p,j}$ as a linear function of the ballot position effect $\alpha_{p,t(j)}$; a candidate’s incumbency status $\text{Inc}_{p,j}$; and an unobserved candidate quality $\varepsilon_{p,j}$.

$$a_{p,j} = \alpha_{p,t(j)} + \text{Inc}_{p,j} \lambda_{t(j)} + \varepsilon_{p,j} \tag{3}$$

We allow the ballot position effect $\alpha_{p,t(j)}$ to vary with individual- and race-level observable characteristics. Equation 4 models $\alpha_{p,t(j)}$ as a linear function of a position

specific constant $\delta_{p,t(j)}$, a candidate’s incumbency status $\text{Inc}_{p,j}$, and a vector of race-level covariates X_j that includes indicators for whether this is a school board race, whether the election takes place on a statewide election date, and whether there is an open seat (i.e., at least one incumbent is not running for reelection).

$$\alpha_{p,t(j)} = \delta_{p,t(j)} + \text{Inc}_{p,j}\gamma_{p,t(j)} + X_j\beta_{p,t(j)} \tag{4}$$

In Eqs. 3 and 4, the coefficients λ , δ , γ , and β vary with the “race type” $t(j)$. That is, we allow the effect of position and incumbency to vary with the number of winners and candidates in the race. Because we can only identify relative differences in attractiveness, we normalize $\delta_{N,t(j)}$, $\gamma_{N,t(j)}$, and $\beta_{N,t(j)}$ to zero where N denotes the last ballot position. This implies that the coefficients in Eq. 4 are the difference in the effect of the specific observable in position p relative to the last ballot position.

Assuming that the $\varepsilon_{p,j}$ ’s are drawn from a logistic distribution, we use Maximum Likelihood to estimate the parameters λ , δ , γ , and β (See Appendix 1 for details). We then calculate for each candidate i in race j , the difference in the probability of winning implied by our parameter estimates when candidate i is in the first ballot position relative to when candidate i is rotated into every ballot position once, holding fixed all the other characteristics of the candidates in race j (See Appendix 2 for an example). We then pool the implied treatment effect across all candidates sharing a specific characteristic (e.g., school board candidates), and compare it to the implied treatment effect across all candidates not holding that specific characteristic (e.g., city council candidates). The difference between the two groups is our estimated effect of the specific characteristic on the first candidate advantage. We estimate the standard error of this difference by calculating the standard error of the estimated difference in 1,000 bootstraps blocked by race.

Table 2 presents our estimates of the how the first position advantage varies by candidate and race characteristics. Column (1) presents the counterfactual average first position effect for all candidates in a given election type. For example, column (1) indicates that candidates listed first are 5.2 percentage points ($\sigma = 2.0$ percentage points) more likely to win office than expected in five-candidate two-winner races. The bottom row of Table 2 indicates that the sample-weighted average (i.e. weighting the estimate from each election type by the number of elections in the sample) of the first candidate advantage across all election types is 4.4 percentage points ($\sigma = 0.5$ percentage points).

We find little difference in the magnitude of the first candidate advantage across electoral contexts. Column (2) indicates that the difference in the first candidate advantage in statewide general elections relative to other lower turnout elections is not statistically significant: the first candidate advantage is about 0.8 percentage point larger ($\sigma = 1.0$ percentage points, $p = 0.436$ two tailed) in races held in conjunction with statewide general elections than in local-only elections. This finding hints that the first candidate advantage does not result primarily from uninformed voters mobilized to the polls by other races on the ballot. Columns (3) and (4) also indicate no substantively important or statistically significant differences between races with and

Table 2 Counterfactuals implied by MLE parameter estimates by election type

Election type	Main effect	Statewide	Open seat	School board	Incumbent
1 of 2	0.031 (0.012)	-0.024 (0.022)	-0.032 (0.025)	0.027 (0.026)	0.014 (0.012)
1 of 3	0.023 (0.017)	0.054 (0.036)	0.008 (0.037)	0.009 (0.038)	0.040 (0.037)
1 of 4	0.034 (0.028)	-0.014 (0.054)	0.045 (0.053)	-0.036 (0.056)	-0.059 (0.073)
1 of 5	0.085 (0.045)	0.147 (0.093)	-0.095 (0.088)	0.106 (0.084)	0.298 (0.170)
2 of 3	0.025 (0.013)	0.050 (0.025)	-0.017 (0.026)	-0.050 (0.028)	0.006 (0.020)
2 of 4	0.033 (0.016)	-0.057 (0.031)	0.005 (0.033)	-0.018 (0.032)	0.009 (0.028)
2 of 5	0.050 (0.021)	-0.015 (0.039)	-0.024 (0.040)	-0.052 (0.041)	-0.009 (0.036)
2 of 6	0.033 (0.026)	0.019 (0.053)	-0.066 (0.053)	0.031 (0.053)	0.065 (0.057)
2 of 7	0.036 (0.032)	-0.024 (0.068)	0.068 (0.063)	0.077 (0.069)	0.041 (0.085)
2 of 8	0.119 (0.048)	0.040 (0.088)	-0.101 (0.092)	-0.048 (0.093)	0.188 (0.074)
3 of 4	0.040 (0.014)	0.048 (0.029)	0.024 (0.026)	-0.009 (0.031)	0.001 (0.026)
3 of 5	0.067 (0.018)	-0.021 (0.037)	-0.039 (0.037)	0.058 (0.041)	0.027 (0.031)
3 of 6	0.107 (0.021)	0.009 (0.041)	0.036 (0.041)	0.074 (0.043)	0.003 (0.036)
3 of 7	0.089 (0.028)	0.029 (0.057)	-0.026 (0.054)	0.019 (0.056)	0.006 (0.055)
3 of 8	0.068 (0.036)	0.072 (0.068)	-0.059 (0.073)	0.000 (0.068)	0.091 (0.066)
3 of 9	0.053 (0.044)	0.079 (0.087)	-0.068 (0.092)	-0.007 (0.087)	0.150 (0.089)
Sample-weighted average	0.044 (0.005)	0.008 (0.010)	-0.013 (0.010)	0.005 (0.011)	0.020 (0.009)
<i>p</i> -Value (two-tailed)	0.000	0.436	0.203	0.657	0.020

Note Standard errors constructed using 1,000 bootstraps blocked by race

without an open seat, and between school board and city council elections. Column (5) shows 2.0 percentage points larger increase ($\sigma = 0.9$ percentage points, $p = 0.020$ two tailed) in incumbents' probability of winning from being listed first relative to non-incumbents, suggesting that incumbents benefit more from being listed first.

Table 3 Vote share by ballot position and percentile of the vote share distribution in five-candidate, two-winner elections

Ballot position	Mean	Percentile				
		10th	25th	50th	75th	90th
First	43.08%	21.53%	30.91%	43.85%	54.05%	65.33%
Second	40.21	18.94	27.76	39.77	51.31	61.10
Third	38.46	16.95	26.24	38.78	50.73	59.25
Fourth	38.36	19.26	27.00	37.50	48.42	59.06
Fifth	39.90	19.04	28.64	40.35	51.04	59.42
First	43.08	21.53	30.91	43.85	54.05	65.33
Not first	39.23	18.67	27.32	39.01	50.35	59.84
Difference	3.85	2.85	3.59	4.85	3.70	5.49
Implied first candidate advantage (e.g. Difference * (N - 1)/N)	3.08 (0.72)	2.28 (0.90)	2.88 (1.04)	3.88 (0.88)	2.96 (0.99)	4.39 (1.36)

Note Standard errors in parentheses constructed using 1,000 bootstraps blocked by race

This last finding may hint that higher quality candidates benefit more than lower quality candidates from being listed first. We examine this hypothesis by looking at the distribution of votes shares in each ballot position and comparing the difference in vote shares between candidates in the same percentile of these distributions across different ballot positions. Given the random assignment of candidates to ballot positions and assuming that vote share is monotone in quality, candidates in the same percentile of these distributions are of similar quality. If higher quality candidates indeed benefit more from being listed first, then we expect the difference between the first and the remaining candidates’ vote shares to be greater in the *p*th percentile than in the *q*th < *p*th percentile.

Table 3 illustrates how we examine the heterogeneity in the effect of ballot order with respect to candidates’ quality. The table provides the mean and 10th, 25th, 50th, 75th, 90th percentile of the distribution of vote shares among candidates listed in different ballot positions in five-candidate two-winner races.¹² The mean vote share of candidates listed first in these elections is 43.08% whereas the mean vote share of candidates not listed first is 39.23% implying that, on-average, candidates’ vote shares increase by more than three percentage points from being listed first.

The remainder of Table 3 indicates that not all candidates benefit equally from being listed first. We see that the vote share in the 10th percentile of the vote share distribution of candidates listed first is 21.53%, as compared to 18.94, 16.95, 19.26, and 19.04% for candidates listed in the second, third, fourth, and fifth ballot positions respectively. The 2.85 percentage point difference between the 10th percentile vote share of candidates listed first and the 10th percentile vote share of candidates not listed first translates into a normalized 2.28 percentage points

¹² Vote share is defined as the number of votes received by a candidate multiplied by the number of winners in the race divided by the total number of votes cast in the race.

($\sigma = 0.90$ percentage points) advantage from being listed first.¹³ As a point of comparison, this effect increases to a 4.39 percentage point ($\sigma = 1.42$ percentage point) advantage for candidates in the 90th percentile of the vote share distribution.

Table 4 aggregates these results across election types. The bottom row of Table 4 shows that the sample-weighted average first-candidate advantage increases from 1.91 percentage points ($\sigma = 0.29$ percentage points) for candidates in the 10th percentile of the vote share distribution to 2.90 percentage points ($\sigma = 0.28$ percentage points) for candidates in 90th percentile.¹⁴ Figure 2a and 2b depict how the first candidate advantage by percentile varies with the number of candidates in the race.¹⁵ Figure 2a shows that in races with three to four candidates, the average first candidate advantage is 1.48 percentage points ($\sigma = 0.25$ percentage points), and it remains relatively constant across percentiles. Figure 2b shows that in races with five to nine candidates the average first candidate advantage increases to 3.37 percentage points ($\sigma = 0.31$ percentage points) and that it increases with the percentile of the vote share distribution. For example, candidates in the 10th percentile of the vote share distribution benefit from an increase of 2.52 percentage points ($\sigma = 0.44$ percentage points) whereas candidates in the 90th percentile of the distribution benefit from a 4.41 percentage points ($\sigma = 0.49$ percentage points) increase.

Results: Satisficing and Ballot Order Effects

Many recent papers on ballot order effects highlight satisficing behavior (see Simon (1955)) as a key mechanism leading to the first candidate advantage (Miller and Krosnick 1998; Brockington 2003; Koppell and Steen 2004). Satisficing has also been used to explain many other political phenomena including positive turnout despite Downs' Paradox of Voting, incrementalism in budgeting, and party dynamics in a multi-dimensional issue space (see Bendor 2003 for a summary).

A satisficing voter has in mind a fixed aspiration level that reflects the minimal quality she considers satisfactory. In a single-winner election, a satisficing voter evaluates the candidates according to the ballot order and selects the first candidate that meets or exceeds her aspiration level, without considering candidates listed further down the ballot. In a K-winner election, a satisficing voter evaluates the candidates in order and chooses the first $M \leq K$ candidates that meet or exceed her aspiration level.

¹³ This normalization accounts for the fact that an additional vote received by the first candidate comes at the expense of a candidate listed in another ballot position. We account for this mechanical relationship in vote totals across positions by multiplying the difference between the votes shares of candidates listed first and candidates not listed first by $(N - 1)/N$ (e.g., $4/5 * 2.85 = 2.28$).

¹⁴ Elections with two candidates and one winner are excluded from this analysis because the effects are symmetric by construction. That is, the difference between the 90th percentile vote shares in the first and second ballot positions is also the difference between the 10th percentile vote shares in the first and second ballot position.

¹⁵ Note that the differences we observe with respect to the number of candidates cannot necessarily be interpreted causally because other features of a race may also both affect the magnitude of order effects and be related to the number of candidates in the race.

Table 4 Percentage point advantage of first candidate by election type and percentile of the vote share distribution

Election type	Mean	Percentile				
		0.1	0.25	0.5	0.75	0.9
1 of 2	1.26% (0.35)	1.53% (0.45)	1.46% (0.39)	1.03% (0.52)	1.46% (0.39)	1.53% (0.45)
1 of 3	1.48 (0.65)	1.46 (0.79)	1.14 (0.73)	1.30 (1.02)	2.26 (0.93)	2.36 (0.92)
1 of 4	2.08 (0.96)	0.83 (0.89)	2.32 (1.53)	3.07 (1.44)	2.63 (1.51)	2.93 (1.92)
1 of 5	3.81 (1.45)	1.96 (1.39)	3.65 (1.55)	4.12 (2.00)	5.88 (2.67)	6.55 (2.81)
2 of 3	1.05 (0.43)	0.65 (0.67)	0.95 (0.61)	1.01 (0.58)	1.16 (0.38)	1.12 (0.48)
2 of 4	1.60 (0.53)	1.54 (0.69)	0.94 (0.98)	1.25 (0.58)	1.96 (0.59)	2.04 (0.76)
2 of 5	3.08 (0.72)	2.28 (0.90)	2.88 (1.04)	3.88 (0.88)	2.96 (0.99)	4.39 (1.36)
2 of 6	1.99 (1.00)	-0.18 (0.78)	0.15 (1.39)	2.89 (1.82)	3.21 (1.16)	3.67 (1.98)
2 of 7	1.15 (1.18)	1.10 (0.87)	-0.39 (0.80)	0.68 (2.36)	2.77 (1.44)	3.10 (2.24)
2 of 8	3.41 (1.84)	2.12 (1.71)	2.31 (1.36)	1.83 (2.33)	4.85 (2.73)	9.81 (4.25)
3 of 4	1.82 (0.54)	2.70 (1.20)	2.30 (0.84)	1.61 (0.41)	1.87 (0.52)	1.42 (0.50)
3 of 5	3.11 (0.69)	2.28 (1.07)	2.40 (1.01)	4.54 (0.66)	3.41 (0.67)	3.71 (0.76)
3 of 6	4.10 (0.70)	2.69 (1.25)	4.02 (1.18)	4.35 (0.90)	3.71 (0.75)	5.07 (0.82)
3 of 7	4.22 (0.93)	5.07 (1.74)	4.78 (1.30)	4.17 (1.31)	4.94 (0.85)	3.36 (1.18)
3 of 8	6.04 (1.23)	4.86 (1.86)	7.62 (2.35)	6.02 (1.54)	6.01 (1.75)	6.42 (2.19)
3 of 9	3.32 (1.42)	4.41 (1.19)	4.90 (2.23)	1.38 (2.74)	5.75 (2.09)	2.92 (1.55)
Sample-weighted average (excludes 1 of 2)	2.30 (0.19)	1.91 (0.29)	2.09 (0.29)	2.44 (0.25)	2.67 (0.23)	2.90 (0.28)

Note Standard errors in parentheses constructed using 1,000 bootstraps blocked by race

Satisficing behavior may emerge when assessing the precise quality of a candidate is costly while figuring out whether the candidate is simply “good enough” is less so. Satisficing behavior may also emerge when a rational voter

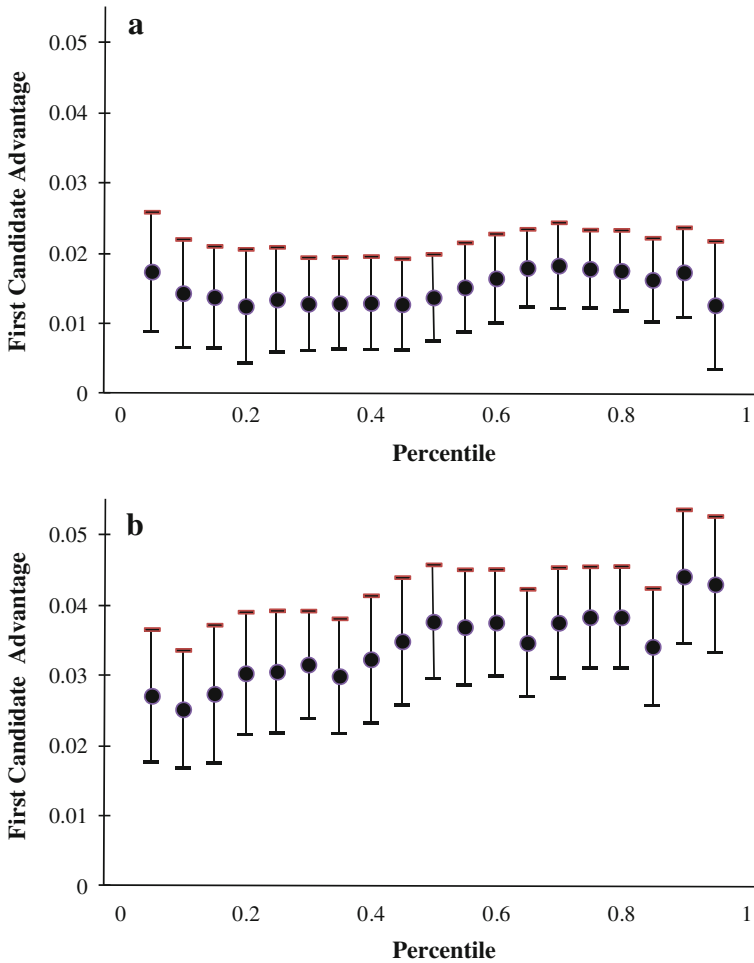


Fig. 2 **a** Sample-weighted average of first position advantage by percentile in elections with 3 to 4 candidates ($N = 3,460$). **b** Sample-weighted average of first position advantage by percentile in elections with 5 to 9 candidates ($N = 2,669$)

wishes to economize on the cognitive costs involved in evaluating the precise quality of the next candidate on the ballot. The voter then chooses the candidate he just evaluated if the expected benefit of evaluating the next candidate on the ballot is smaller than the cognitive cost of doing so, as in Ho and Imai (2008).

The first candidate advantage in single-winner elections is consistent with satisficing as well as many other mechanisms. For example, the first candidate advantage may result from uninformed voters casting a vote for the candidate in the first position and from informed voters considering all candidates and voting for the best one. Such a mechanism, as well as any other mechanism that predicts no interaction between candidates' quality and the first candidate advantage, is not consistent with our findings in the previous section. However, a plethora of

mechanisms, in addition to satisficing, predict an interaction between candidates' quality and the first candidate advantage. Our goal in this section is to derive additional testable predictions of satisficing in multi-winner elections and explore whether they hold in our dataset.

Consider a K -winner race and a satisficing voter who is aware of the fact that he can cast *up to* K votes. If such a voter casts only $M < K$ votes, it means that he examined *all* the candidates on the ballot and found only M of them to be satisfactory. Because of the random assignment of candidates to ballot positions in our dataset, such a voter is equally likely to cast a vote for any position. When aggregating over such voters, their behavior should therefore not generate order effects. If, on the other hand, such a voter casts K votes, then she potentially “run out of votes” before considering all candidates and is therefore more likely to vote for candidates that appear at the top of the ballot. In particular, it is possible that such a voter cast all her K votes after having considered only the first K candidates.

When aggregating over such voters, their behavior should generate order effects. Assuming voters are aware that they can cast K votes and some of them indeed cast K votes, we should observe that candidates in position K perform better than candidates in position $K + 1$. In single-winner races, this leads to the standard prediction that candidates listed first perform better than the candidates listed second. In multi-winner races, this leads to a more nuanced prediction. Because a satisficing voter may have exhausted all her votes before considering the candidate listed third in a two-winner race but not in a three-winner race, we expect that:

(P1) The difference in the likelihood of winning office between the second and third candidates should be larger in two-winner elections than in three-winner elections.

Moreover, if all voters are aware of the fact that they can cast K votes in a K -winner race, then they never “run out of votes” in a multi-winner race when considering the second candidate on the ballot. Because of the random assignment of candidates to ballot positions in our dataset, this implies that voters are equally likely to vote for the candidates in the first and second position, i.e.:

(P2) In multi-winner races, there should be no difference in the likelihood of winning office between candidates listed first and second.

As mentioned above, predictions (P1) and (P2) require some assumptions about voters' awareness of the possibility of casting multiple votes. For (P1) to hold, it is enough that a fixed proportion of voters are aware of the possibility of casting K votes:

(A1) A fixed proportion of voters are aware of the possibility of casting K votes in a K -winner election. Other voters cast only one vote.

If (A1) holds, then a satisficing voter who casts only one vote (because he is unaware of the possibility of casting K votes) behaves similarly in a two- and a three-winner race and hence does not influence the gap in the difference between

Table 5 Distribution of votes by number of votes cast in 2006 Fullerton Joint Union school board election

# Votes cast (% of obs.)	0 (46.15%)	1 (13.48%)	2 (4.72%)	3 (35.57%)	4 (0.08%)
% of Ballots					
1. Hathaway (incumbent)	–	24.29	77.42	91.22	100.00
2. Day	–	44.63	30.65	30.62	100.00
3. Buchi (incumbent)	–	17.51	32.26	89.08	100.00
4. Singer (incumbent)	–	13.56	59.68	89.08	100.00

two- and three-winner races. As explained above, satiscing voters who are aware of the possibility of casting K votes generate the pattern described in (P1).

For (P2) to hold, a stronger assumption is needed:

(A2) All voters are aware of the possibility of casting K votes in a K -winner election.

Indeed, if some voters think they can only cast one vote, then they may “run out of votes” after considering the first candidate and this will result in the first listed candidate having a higher likelihood of winning office than the second listed candidate.

We use the data collected in Hill (2010) to provide some indication that a substantial number of voters are aware of the possibility of casting multiple votes in line with assumption (A1). Table 5 outlines the number of votes cast by a non-random subset of 1,313 voters in four precincts in the 2006 Fullerton Joint Union School Board Election in which voters could cast up to three votes. Table 5 shows that about 46% of the voters in this election did not cast any votes in this particular race. Among the voters who did cast votes, 66% cast three votes implying that they were aware of the possibility of casting three votes in line with Assumption (A1). About 25% of the active voters in this race cast a single vote, providing an upper bound on the number of voters that are unaware of the possibility of casting multiple votes. This finding does not necessarily invalidate assumption (A2) as a plurality of these single votes were for the one non-incumbent in the race, which is more consistent with strategic voting than with being unaware of the ability to cast multiple votes.¹⁶

We find no evidence supporting (P1). We define $W_{p,k,n}$ to be the percentage of winners from ballot position p in elections with k winners and n candidates, and evaluate the difference-in-difference DD_n in the advantage of the second over the third candidate in two- versus three-winner elections:

$$DD_n = (W_{2,2,n} - W_{3,2,n}) - (W_{2,3,n} - W_{3,3,n}) \tag{5}$$

Table 6 shows that in contrast to (P1), the second candidate’s relative advantage over the third candidate does not increase in two-winner elections as compared to

¹⁶ This example also highlights the difficulty of inferring the distribution of votes cast without observing individual ballot data. If we observe that 100 ballots were not cast in a 2-winner election, it could result from 50 voters not casting either of their votes or 100 voters not casting one of their votes.

Table 6 Percentage of winners from second and third ballot position by number of winners and candidates

Number of candidates	2 Winner elections			3 Winner elections			Diff-in-Diff
	2nd	3rd	Diff.	2nd	3rd	Diff.	
4 Candidate elections	0.499	0.487	0.012	0.781	0.722	0.059	-0.047 (0.051)
5 Candidate elections	0.407	0.382	0.025	0.607	0.573	0.034	-0.009 (0.047)
6 Candidate elections	0.333	0.348	-0.014	0.472	0.476	-0.004	-0.010 (0.057)
7 Candidate elections	0.262	0.341	-0.079	0.388	0.415	-0.027	-0.052 (0.071)
8 Candidate elections	0.192	0.244	-0.051	0.431	0.317	0.114	-0.165 (0.657)
Precision-weighted average							-0.027 (0.027)

three winner elections. For example, in four-winner elections $DD_4 = -0.047$ ($\sigma = 5.1$ percentage points) reflecting the fact that candidates listed second won 1.2 and 5.9 percentage points more often than candidates listed third in two- and three-winner elections respectively. A precision weighted-average of the DD_n 's indicates that the difference in the probability of winning between the second and the third candidates is 2.7 percentage points ($\sigma = 2.7$ percentage points) *smaller* in two-winner elections than in three-winner elections in contrary to (P1).

We also find no evidence supporting (P2). We define the statistic $\hat{\theta}_1 - \hat{\theta}_2$ which is our estimate of the percentage point difference in the probability of winning office from the first and the second positions relative to the expected probability under the satisficing hypothesis. The distribution of $\hat{\theta}_p$ implied by the Central Limit Theorem under (A2) is:

$$\hat{\theta}_1 - \hat{\theta}_2 \sim N \left[0, \frac{1}{T} \sum_{j=1}^T 2\pi_j \left((1 - \pi_j) - \left(\frac{K_j - 1}{N_j - 1} - \pi_j \right) \right) \right] \tag{6}$$

Figure 3 shows that candidates listed first win 6.9 percentage points (s.e. 1.3 percentage points) more races than candidates listed second, and thus provides clear evidence against (P2). As indicated above, the data in Table 5 suggests that there is potentially a group of voters who are unaware that they can cast multiple votes in contrast to Assumption (A2). In Appendix 3, we design a test that addresses this possibility. We develop a simple model of satisficing with a fixed aspiration level in which voters cast a single vote. For parameterizations of the model consistent with the observed first candidate advantage, the model predicts that the difference in the likelihood of winning office between the first and second candidates should be only slightly larger than the difference in the likelihood of winning between the second and third candidates. In contrast, Fig. 3 demonstrates that the difference in the

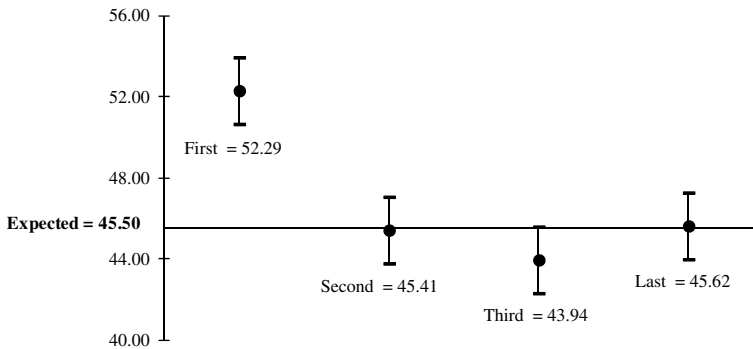


Fig. 3 Percentage of candidates winning office by ballot position in elections with 2 or more winning and losing candidates ($N = 3,402$, 95% confidence interval in bars, $\sigma = 0.84$)

likelihood of winning is substantially larger when comparing the first and second positions than when comparing the second and third positions. Whereas candidates listed first are 6.9 percentage points ($\sigma = 1.3$ percentage points) more likely to win office than candidates listed second, candidates listed second are only 1.4 percentage points ($\sigma = 1.3$ percentage points) more likely to win office than candidates listed third. This suggests that satisficing with a single vote also does not fully explain the patterns observed in the dataset.

The above findings suggest that satisficing behavior in which voters have a fixed aspiration level cannot fully explain the pattern of order effects in our dataset. It is possible, however, that more elaborate forms of satisficing in which the aspiration level changes across positions in certain ways could predict better the pattern of order effects that we observe. It is also possible that there are other cognitive mechanisms at play. The fact that the first candidate receives an advantage that is not shared by other candidates, that the second and last candidates perform as expected and that the remaining middle positions perform worse than expected may be related to how much attention voters devote to different candidates as a function of their position. Krosnick and Alwin (1987) suggest that alternatives considered earlier receive deeper cognitive processing, and thus the attributes of candidates listed near the top of the ballot may be more salient when voting. The tendency of decision makers, when presented with several pieces of information, to weigh more heavily information presented to them in the beginning and in the end (see Thompson (1993) and Hogarth and Einhorn (1992) for surveys) also hints that voters may pay more attention to candidates at the top and at the bottom of the ballot, which may in turn bias voting behavior in their favor. A deeper investigation of such a limited attention mechanism likely will require a combination of laboratory experiments and data on individual ballots.

Results: Downstream Effects

The results in previous sections show that candidates listed in the first ballot position win elections that they would lose otherwise. A concern is that this will cause the

average quality of representatives elected from the first ballot position to be worse than the average quality of representatives elected from other ballot positions. While it is generally hard to measure the quality of representatives, we can test this hypothesis indirectly by looking at how representatives elected from different ballot positions perform in the next election. If representatives elected from the first ballot position are of lower quality, we expect them to perform worse in the next election than candidates elected from other ballot positions. Such a finding would suggest that it is preferable to use rotation schemes rather than randomization to set ballot order in local elections.

Because the final year we observe is 2008, we restrict our analysis to representatives elected to full four-year terms from 1995 to 2004. For each representative we observe one of four possible outcomes in our data four years after they won office. First, we can observe that the representative ran for reelection and won. Second, we can observe that the representative ran for reelection and lost. Third, we can observe an election for the representative’s office that the representative did not compete in. Fourth, we can fail to observe an election for the representative’s office. Because many counties do not report election results for non-competitive races, this last outcome often represents cases where the incumbents ran unopposed for reelection.

We observe that representatives who were elected from the first ballot position are more likely to lose in the next election. Table 7 reports the coefficients and standard errors from four separate regressions where we regress an indicator for each of the four possible outcomes in year $t + 4$ on a dummy variable for the election type that the representative was elected in year t and an indicator for whether the representative was listed in the first ballot position in year t . We see that winners from the first ballot position are 1.4 percentage points ($\sigma = 0.6$ percentage points) more likely to lose in the next election than winners from other ballot positions. If we restrict our analysis to elections with two or more winning and losing candidates, which are the elections in which the first listed candidates received the greatest benefits from ballot positioning, the increase in the probability of losing grows to 2.9 percentage points ($\sigma = 1.0$ percentage points). One caveat to this finding is that Table 7 also indicates that a substantial portion of the increased probability of losing is offset by a decreased probability of observing winners from the first ballot position not running in an election in their district four years later.

Table 7 Outcomes in next election among winners listed first in previous election relative to all other winners

	Candidate won	Candidate lost	Candidate not matched	Race not listed
All elections ($N = 10,595$ candidates in 5,208 elections)	-0.001 (0.010)	0.014 (0.006)	-0.010 (0.009)	-0.002 (0.008)
Election with two or more winning and losing candidates ($N = 6,041$ candidates 2,465 elections)	-0.001 (0.015)	0.029 (0.010)	-0.021 (0.014)	-0.007 (0.010)

As not running for reelection may also indicate being of lower quality, we see this result as only being suggestive that winners from the first ballot position are of lower quality than winners from other ballot positions.

Conclusion

This paper contributes to the study of ballot order effects in a number of ways. From a policy perspective, randomization and rotation procedures are costly to administer and may cause voters' confusion.¹⁷ Thus, they should only be undertaken if ballot order alters the results of elections. We demonstrated that in city council and school board elections in California this is indeed the case. In between four and five percent of the elections we examined, the candidate listed first won office as a result of her or his ballot position. Given that states are the least likely to use rotated or randomized ballots in local elections, our results imply that a non-negligible portion of local governmental policies is likely being set by individuals elected because of their ballot position. In particular, the current use of alphabetical ordering for local elections in many states not only provides the same candidates with the advantage of ballot position in election after election, but also gives the beneficiaries of ballot positioning the subsequent advantage of incumbency.

The substantial increase in the likelihood of winning office when being listed first also suggests that strong incentives exist to strategically manipulate ballot order. The results of the supposedly random assignment of parties to ballot positions in the March 2007 Russian regional parliamentary elections emphasize this point. While the Unified Russia party would be expected to be listed first in 2.1 of the fourteen regions under a random allocation of parties to ballot positions, it was actually listed first in eight of the fourteen regions. The probability of Unified Russia being listed first in eight or more regions by random chance is less than 1 in 2,000. This finding underscores the importance of ensuring that any policies done to mitigate order effects, like randomization or rotation, are done so in a fair manner.

In the second part of the paper, we examined some of the mechanisms leading to ballot order effects. We found that candidates listed second perform worse than candidates listed first and that candidates listed third perform relatively similar to candidates listed second. We also found that the relative performance of the second and third candidates does not differ significantly between two-winner and three-winner races. Finally, we found that the first position advantage is comparable across percentiles of the vote share distributions in races with small number of candidates, but is greater for candidates receiving larger vote share in races with five or more candidates. These findings suggest that voting behavior cannot be explained fully by simple forms of satisficing. Exploring further the mechanisms leading to order effects in order to evaluate the effectiveness of current randomization and rotation schemes is a task for future research.

¹⁷ It is estimated that the cost of ballot rotation in the 1994 Alaska primary and general elections was \$64,024 (see Sonneman v. State of Alaska, 969 P 2d 632), which is about \$137 per precinct. In California, for example, there are more than 24,000 precincts.

Acknowledgments We thank Jon Bendor, Jonah Berger, Michael Hanmer, Daniel Kessler, Peter Reiss, and seminar participants at MIT, Stanford, and the 2008 American Political Science Association Annual Meeting for comments and suggestions. We also thank Seth Hill for providing us with the individual-level ballot data. Salant acknowledges the financial support of the Stanford Institute for Economic Policy Research, the Leonard W. and Shirley R. Ely Fellowship. Replication data is available at <http://www.sas.upenn.edu/~marcmere/replicationdata/BallotOrderReplicationData.zip>.

Appendix 1: Maximum Likelihood Estimation

We wish to estimate the coefficients λ , δ , γ , and β using maximum likelihood in the model $a_{p,j} = \alpha_{p,t(j)} + \text{Inc}_{p,j}\lambda_{t(j)} + \varepsilon_{p,j}$ where $\alpha_{p,t(j)} = \delta_{p,t(j)} + \text{Inc}_{p,j}\gamma_{p,t(j)} + X_j\beta_{p,t(j)}$.

Assume that the $\varepsilon_{p,j}$'s are drawn from a logistic distribution. Then the probability that the candidate in position p is the most attractive candidate in race j is $\Pr_j(p = 1) = \frac{\exp(\alpha_{p,t(j)} + \text{Inc}_{p,j}\lambda_{t(j)})}{\sum_i \exp(\alpha_{i,t(j)} + \text{Inc}_{i,j}\lambda_{t(j)})}$. Similarly, the probability that the candidate in position p is the second most attractive candidate given that the candidate in position p' is the most attractive candidate is $\Pr_j(p = 2|p' = 1) = \frac{\exp(\alpha_{p,t(j)} + \text{Inc}_{p,j}\lambda_{t(j)})}{\sum_{i \neq p'} \exp(\alpha_{i,t(j)} + \text{Inc}_{i,j}\lambda_{t(j)})}$. Finally, the probability that the candidate in position p is the third most attractive candidate given that the candidates in positions p' and p'' are the two most attractive candidates is $\Pr_j(p = 3|p', p'' = 1, 2) = \frac{\exp(\alpha_{p,t(j)} + \text{Inc}_{p,j}\lambda_{t(j)})}{\sum_{i \neq p', p''} \exp(\alpha_{i,t(j)} + \text{Inc}_{i,j}\lambda_{t(j)})}$.

Let $Y_j = [Y_{1,j}, \dots, Y_{N_j,j}]$ denote the vector of observed outcomes in race j where $Y_{i,j} = 1$ if the candidate in position i wins office and $Y_{i,j} = 0$ otherwise. We construct the likelihood function $L(\delta, \lambda, \gamma, \beta|Y_j)$ as follows.

In a one winner race, $L(\delta, \lambda, \gamma, \beta|Y_j) = \prod_{i=1}^{N_j} \Pr_j(i = 1)^{Y_{i,j}}$ is the likelihood that the winning candidate is the most attractive candidate. In a two-winner race, $L(\delta, \lambda, \gamma, \beta|Y_j)$ is the likelihood that the two winning candidates are the two most attractive candidates. By the definition of conditional probability, the probability that the candidates in position p and p' are the most attractive and second most attractive candidates respectively in race j , $\Pr_j(p = 1, p' = 2)$, is equal to $\Pr_j(p' = 2|p = 1)\Pr_j(p = 1)$. Thus, $L(\delta, \lambda, \gamma, \beta|Y_j) = \prod_{i=1}^{N_j-1} \prod_{i'=i+1}^{N_j} (\Pr_j(i = 1) \Pr_j(i' = 2|i = 1) + \Pr_j(i' = 2) \Pr_j(i = 1|i' = 2))^{Y_{i,j}Y_{i',j}}$. Similarly, in a three-winner race, $L(\delta, \lambda, \gamma, \beta|Y_j)$ is the likelihood that the three winning candidates are the three most attractive candidates. Again by the definition of conditional probability $\Pr_j(p = 1, p' = 2, p'' = 3)$ is equal to $\Pr_j(p'' = 3|p' = 2, p = 1)\Pr_j(p' = 2|p = 1)\Pr_j(p = 1)$. Thus,

$$\begin{aligned} L(\delta, \lambda, \gamma, \beta|Y_j) &= \prod_{i=1}^{N_j-2} \prod_{i'=i+1}^{N_j-1} \prod_{i''=i'+1}^{N_j} (\Pr_j(i = 1) \Pr_j(i' = 2|i = 1) \Pr_j(i'' = 3|i = 1, i' = 2) \\ &+ \Pr_j(i = 1) \Pr_j(i'' = 2|i = 1) \Pr_j(i' = 3|i = 1, i'' = 2) \\ &+ \Pr_j(i' = 1) \Pr_j(i = 2|i' = 1) \Pr_j(i'' = 3|i' = 1, i = 2) + \Pr_j(i' = 1) \Pr_j(i'' = 2|i' = 1) \\ &\Pr_j(i = 3|i' = 1, i'' = 2) + \Pr_j(i'' = 1) \Pr_j(i = 2|i'' = 1) \Pr_j(i' = 3|i'' = 1, i = 2) \\ &+ \Pr_j(i'' = 1) \Pr_j(i' = 2|i'' = 1) \Pr_j(i = 3|i'' = 1, i' = 2))^{Y_{i,j}Y_{i',j}Y_{i'',j}}. \end{aligned}$$

We solve for our parameter estimates $\hat{\delta}$, $\hat{\lambda}$, $\hat{\gamma}$, and $\hat{\beta}$ by finding the values that maximize $\prod_{j=1}^T L(\delta, \lambda, \gamma, \beta | Y_j)$.

Appendix 2: Constructing Counterfactuals

We illustrate how the Maximum Likelihood estimates are used to construct implied treatment effects in races with one winner and three candidates. Let $C_{p,p',j}$ be the counterfactual probability that the candidate who is actually listed in position p wins office in race j were he listed in position p' . Given the observed incumbency status of the three candidates, $Inc_{1,j}$, $Inc_{2,j}$, and $Inc_{3,j}$ and race-level covariates, X_j , the counterfactual probability of winning of the candidate actually listed in position 1 in race j where he listed in position 1 is $C_{1,1,j} =$

$$\frac{\exp(\hat{\delta}_{1,t(j)} + Inc_{1,j}(\hat{\lambda}_{t(j)} + \hat{\gamma}_{1,t(j)}) + X_j \hat{\beta}_{1,t(j)})}{\exp(\hat{\delta}_{1,t(j)} + Inc_{1,j}(\hat{\lambda}_{t(j)} + \hat{\gamma}_{1,t(j)}) + X_j \hat{\beta}_{1,t(j)}) + \exp(\hat{\delta}_{2,t(j)} + Inc_{2,j}(\hat{\lambda}_{t(j)} + \hat{\gamma}_{2,t(j)}) + X_j \hat{\beta}_{2,t(j)}) + \exp(Inc_{3,j} \hat{\lambda}_{t(j)})}$$

If we rotate the candidate actually listed first to the second position, the candidate actually listed second to the third position, and the candidate actually listed third to the first position, then $C_{1,2,j} =$

$$\frac{\exp(\hat{\delta}_{2,t(j)} + Inc_{1,j}(\hat{\lambda}_{t(j)} + \hat{\gamma}_{2,t(j)}) + X_j \hat{\beta}_{2,t(j)})}{\exp(\hat{\delta}_{1,t(j)} + Inc_{3,j}(\hat{\lambda}_{t(j)} + \hat{\gamma}_{1,t(j)}) + X_j \hat{\beta}_{1,t(j)}) + \exp(\hat{\delta}_{2,t(j)} + Inc_{1,j}(\hat{\lambda}_{t(j)} + \hat{\gamma}_{2,t(j)}) + X_j \hat{\beta}_{2,t(j)}) + \exp(Inc_{2,j} \hat{\lambda}_{t(j)})}$$

Finally, if we rotate the candidates so that the candidate actually listed first is listed third, the candidate actually listed second is listed first, and the candidate actually listed third is listed second, then $C_{1,3,j} =$

$$\frac{\exp(Inc_{1,j} \hat{\lambda}_{t(j)})}{\exp(\hat{\delta}_{1,t(j)} + Inc_{2,j}(\hat{\lambda}_{t(j)} + \hat{\gamma}_{1,t(j)}) + X_j \hat{\beta}_{1,t(j)}) + \exp(\hat{\delta}_{2,t(j)} + Inc_{3,j}(\hat{\lambda}_{t(j)} + \hat{\gamma}_{2,t(j)}) + X_j \hat{\beta}_{2,t(j)}) + \exp(Inc_{1,j} \hat{\lambda}_{t(j)})}$$

Similar counterfactual probabilities can be constructed for the candidates actually listed in position 2 and position 3 in race j . The estimated treatment effect of being listed first for the candidate actually listed in position p in race j is defined as $C_{p,1,j} - (C_{p,1,j} + C_{p,2,j} + C_{p,3,j})/3$.

Appendix 3: Satisficing with a Single Vote

Consider elections with N candidates in which the qualities X_1, \dots, X_N of the candidates located in positions 1, ..., N are random variables distributed i.i.d. with c.d.f. $F()$. A representative satisficing voter evaluates the candidates according to the ballot order and chooses the first candidate with quality above his aspiration threshold C . If no candidate is above C , then the voter selects the highest-quality candidate. The probability of choosing the first candidate in this model is:

$$P(1st) = P(X_1 \geq C) + P(X_1 < C, X_1 > X_2, \dots, X_N) = (1 - F(C)) + (1/N)F(C)^N.$$

Similarly, the probabilities of selecting the second and third candidates are:

$$\begin{aligned}
 P(2\text{nd}) &= P(X_2 \geq C, X_1 < C) + P(X_2 < C, X_2 > X_1, \dots, X_N) \\
 &= F(C)(1 - F(C)) + (1/N)F(C)^N \\
 P(3\text{rd}) &= P(X_3 \geq C, X_1 < C, X_2 < C) + P(X_3 < C, X_3 > X_1, \dots, X_n) \\
 &= F(C)^2(1 - F(C)) + (1/N)F(C)^N.
 \end{aligned}$$

This implies that the expected difference in the number of winners from the first and second ballot positions is $P(1\text{st}) - P(2\text{nd}) = (1 - F(C))^2$, while the expected difference in the number of winners from the second and third ballot positions is $P(2\text{nd}) - P(3\text{rd}) = F(C)(1 - F(C))^2$.

Figure 3 indicates that $P(1\text{st}) - P(2\text{nd}) = 0.069$ in elections with two or more winners and losers, and thus $F(C)$ is about 0.74. This implies that the expected difference in the number of winners between the second and third ballot positions should be about three quarters of the difference between the first and second positions. As Fig. 3 indicates, however, the actual difference between the second and third is only one-fifth of the difference between the first and second candidates. Thus, the difference between the first and second candidates is larger than expected even if voters are satisficing with a single vote.

References

- Alvarez, R. M., Sinclair, B., & Hasen, R. L. (2005). How much is enough? The “ballot order effect” and the use of social science research in election law disputes. *Election Law Journal*, 5(1), 40–56.
- Bendor, J. (2003). Herbert A. Simon: Political scientist. *Annual Review of Political Science*, 6, 433–471.
- Brockington, D. (2003). A low information theory of ballot position effect. *Political Behavior*, 25(1), 1–27.
- Hill, S. (2010). Changing composition and changing allegiance in American elections. Ph.D. Thesis, University of California, Los Angeles.
- Ho, D. E., & Imai, K. (2006). Randomization inference with natural experiments: An analysis of ballot effects in the 2003 California recall election. *Journal of the American Statistical Association*, 101(475), 888–900.
- Ho, D. E., & Imai, K. (2008). Estimating casual effects of ballot order from a randomized natural experiment: California alphabet lottery, 1978–2002. *Public Opinion Quarterly*, 72(2), 216–240.
- Hogarth, R. M., & Einhorn, H. J. (1992). Order effects in belief updating: The belief adjustment model. *Cognitive Psychology*, 24(1), 1–55.
- King, A., & Leigh, A. (2009). Are ballot order effects heterogeneous? *Social Science Quarterly*, 90(1), 71–87.
- Koppell, J. G., & Steen, J. A. (2004). The effects of ballot position on election outcomes. *Journal of Politics*, 66(1), 267–281.
- Krosnick, J. A., & Alwin, D. F. (1987). An evaluation of a cognitive theory of response-order effects in survey measurement. *Public Opinion Quarterly*, 51(2), 201–219.
- Krosnick, J. A., Miller, J. M., & Tichy, M. P. (2004). An unrecognized need for ballot reform: The effects of candidate name order on election outcomes. In A. N. Crigler, M. R. Just, & E. J. McCaffery (Eds.), *Rethinking the vote: The politics and prospects of American election reform* (pp. 51–73). New York: Oxford University Press.
- Meredith, M. & Salant Y. (2007). Causes and consequences of ballot order effects. *SIEPR Discussion Paper No.* 0629.
- Miller, J. M., & Krosnick, J. A. (1998). The impact of candidate name order on election outcomes. *Public Opinion Quarterly*, 62(3), 291–330.
- Simon, H. (1955). A behavior model of rational choice. *Quarterly Journal of Economics*, 69(1), 99–118.
- Thompson, A. M. (1993). Appendix: Serial position effects in the psychological literature. In Dunne, B. J., Dobyys, Y. H., & Jahn, R. G. Series position effects in random event generator experiments. *Journal of Scientific Exploration*, 8(2), 197–215.