American Economic Journal: Economic Policy 2009, 1:1, 229–257
http://www.aeaweb.org/articles.php?doi=10.1257/pol.1.1.229

Who Misvotes? The Effect of Differential Cognition Costs on Election Outcomes†

By Kelly Shue and Erzo F. P. Luttmer*

If voters have negligible cognition costs, ballot layout should not affect election outcomes. We explore deviations from rational voting using quasi-random variation in candidate name placement on ballots from the 2003 California recall election. We find that minor candidates' vote shares almost double when their names are adjacent to the names of major candidates. All else equal, vote share gains are larger in precincts with higher percentages of poorly educated, poor, or third-party voters. A major candidate that disproportionally attracts voters from such precincts faces an electoral disadvantage. We also explore which voting technology platforms and brands mitigate misvoting. (*JEL* D72)

We use the quasi-random variation in layouts of 152 types of ballots used in the 2003 California recall election to study how different ballot designs led to systematic deviations in the vote shares of candidates. We refer to votes for a candidate that are attributable solely to that candidate's favorable position on the ballot as misvotes because differences in ballot layout should not affect the decisions of fully rational voters.[1] Previous literature has focused on position misvotes—when candidates receive more votes because they are listed first on the ballot page or column. In this paper, we consider another type of misvote, adjacency misvotes—when candidates receive more votes because they are adjacent to a popular candidate.

Adjacency misvotes can occur if voters accidentally select candidates adjacent to their preferred choice. Evidence from simulated voting experiments suggests that some voters become confused by the ballot layout and voting technology and, as a result, cast accidental, but valid, votes for candidates adjacent to their intended choice

* Shue: Department of Economics, Harvard University, Littauer Center, 1805 Cambridge St., Cambridge, MA 02138, (e-mail: shue@fas.harvard.edu); Luttmer: Harvard University Kennedy School of Government, 79 John F. Kennedy St., Cambridge, MA 02138, National Bureau of Economic Research (NBER), and Institute for the Study of Labor (IZA) (e-mail: erzo_luttmer@harvard.edu). We would like to thank Attila Ambrus, Gary Chamberlain, Stefano DellaVigna, John Friedman, Edward Glaeser, Caroline Hoxby, Brian Jacob, Lisa Kahn, Rebecca Kalmus, two anonymous referees, as well as seminar participants at the University of California, Berkeley, Harvard University, and the 2007 Public Choice Society meetings for helpful conversations and insightful comments. We are also grateful to the staff at the University of California, Berkeley Institute of Governmental Studies Statewide Database for their generous assistance. We thank Andra Hibbert for excellent research assistance. Kelly Shue gratefully acknowledges support from National Science Foundation Graduate Research Fellowship Grant 2006042992. Erzo Luttmer gratefully acknowledges funding from the National Institute on Aging through Grant Number T32-AG00186 to the NBER. All errors are our own.

† To comment on this article in the online discussion forum visit the articles page at: http://www.aeaweb.org/articles.php?doi=10.1257/pol.1.1.229

[1] We use rational in the traditional sense of the term, someone who has a preference ordering over candidates that is complete and transitive. Implicit in this definition is that choice is not affected by cognition costs.

(e.g., Susan K. Roth 1998). These adjacency misvotes are especially likely if the ballot columns are unclearly marked, the ballot text is small, or the punch card grid is confusing. Appendix A shows some sample ballots that may produce adjacency misvotes. It is also conceivable that adjacency misvotes can occur if voters scan for popular candidates and then give increased consideration to candidates listed near the popular candidate. We will present strong evidence that accidental voting rather than increased consideration is the primary cause of adjacency misvotes.

Misvotes are important for two reasons. First, and at a more fundamental level, misvotes provide an important example of why demographic differences in human cognition matter. While Herbert A. Simon's (1955) seminal article has spurred much research on the limits of rational decision making (see John Conlisk 1996 for an overview), relatively few papers show how variation in cognitive ability across individuals can have important aggregate consequences. We examine how ballot design alters voting behavior and for which types of voters these effects are strongest. Following John A. List and David Lucking-Reiley (2002), we assume that the rationality of voter behavior decreases with the cognitive cost of processing a complicated ballot design. We show how demographic variation in cognitive costs can affect the relative vote shares of major candidates and thereby determine election results. Our paper thus contributes to a relatively new literature that examines the consequences of demographic variation in people's decision-making abilities in real world situations (e.g., Brigitte C. Madrian and Dennis F. Shea 2001, James Choi et al. 2004, Richard H. Thaler and Shlomo Benartzi 2004, and Annamaria Lusardi and Olivia S. Mitchell 2007).

Our paper also adds to the very limited literature concerning behavioral confusion summarized in Stefano DellaVigna (2007). While most of the behavioral literature explores systematic biases in judgment, Michael S. Rashes (2001) shows that investors can be confused, accidentally investing in the ticker *MCI* when they intended to invest in the phone company with ticker *MCIC*. We explore the consequences of pure confusion, when people (possibly) have the right judgment but fail to implement their intention. Although there is no individual-level systematic bias due to confusion, we show that confusion leads to significant aggregate bias because voters favoring certain political candidates may be more prone to confusion. Our methodology also provides a clean estimate of the incidence of pure errors. We estimate the error rate as roughly 1 in 300, which is of the same order of magnitude as the estimate by Rashes.

Second, misvotes can change election outcomes. Previous literature on position misvotes has predominantly focused on who gains from misvotes. Joanne M. Miller and Jon A. Krosnick (1998), Jonathan G. S. Koppell and Jennifer A. Steen (2004), and Amy King and Andrew Leigh (forthcoming) show that candidates gain up to 2 percent in vote share if they are listed near the start of the ballot. However, the gain from being listed near the top of the candidate list tends to be much smaller in well-publicized elections featuring party labels and well-known or incumbent candidates (e.g., Delbert A. Taebel 1975, Miller and Krosnick 1998, Koppell and Steen 2004, and Daniel E. Ho and Kosuke Imai 2006). Further, the bias caused by position misvotes can be minimized if the list of candidate names is randomized and rotated through voting districts. Therefore, what often matters more for the outcome of major elections in the United States is which of the two front-runners systematically *loses* more votes to misvotes. This point is best illustrated by the unexpected large number of votes that Patrick J. Buchanan received on the "butterfly ballot" used in Palm Beach County, FL during the 2000 US Presidential Election. While even Buchanan himself acknowledges that many of the 3,407 votes he received in Palm Beach County are due to voter mistakes,[2] the bigger question is whether these misvotes came sufficiently more from Al Gore than from George W. Bush to overturn Bush's 537 vote margin of victory. Numerous studies using demographic trends and Buchanan's performance in surrounding counties indicate that Al Gore would have won if there had been no misvoting in Palm Beach County (e.g., Jonathan W. Wand et al. 2001 and Richard L. Smith 2002).

Unlike position misvotes, adjacency misvotes can offer insight into differences in vote share lost by major candidates. Vote share gained by a candidate adjacent to a major candidate is vote share lost by the associated major candidate. We assess which of the major candidates disproportionally loses votes to adjacent candidates in two ways. First, we examine whether minor candidates in the vicinity of one major candidate gain as much vote share as minor candidates in the vicinity of another major candidate. Second, we relate the extent of misvoting in each election precinct to demographic characteristics at the precinct level and estimate whether major candidates that appeal to voters with certain demographic characteristics lose more votes to adjacent candidates. Both methods indicate that the leading Democratic candidate, Cruz M. Bustamante, disproportionally lost more votes due to misvoting.

We measure adjacency misvoting by exploiting the rotational structure of ballots used in the 2003 California recall election. Candidates are listed in order according to a randomized alphabetization in District 1, and then rotated one position at a time in subsequent districts. Due to the interaction between ballot rotation and each county's unique page layout, there exists significant exogenous variation in whether a minor (unpopular) candidate's name is adjacent to that of a major (popular) candidate. The increase in vote share experienced by a minor candidate when she is adjacent to a major candidate is an estimate of adjacency misvoting. Figure 1 provides motivation for testing the changes in vote share caused by adjacency. We see that vote share for minor candidate Ronald J. Palmieri jumps significantly in each county-district in which he is adjacent to a major candidate. While this figure shows vote share changes for a single minor candidate, the pattern it represents is more general, and suggests a strong effect of adjacency on minor candidate vote share. In addition to testing for adjacency effects, we use precinct-level voting results and demographic data to test if certain demographic groups are more likely to cast misvotes than others. Finally, we test if, controlling for demographic characteristics, certain kinds of voting technologies correspond to a reduction in the level of misvoting. We study misvoting in the California recall election because the structure of the ballot rotation in this election allows for credible identification of adjacency effects. However, the types of voting errors we uncover are likely to occur more generally.

[2] In an appearance on the "Today" show on November 9, 2000, Buchanan said, "When I took one look at that ballot on Election Night … it's very easy for me to see how someone could have voted for me in the belief they voted for Al Gore."

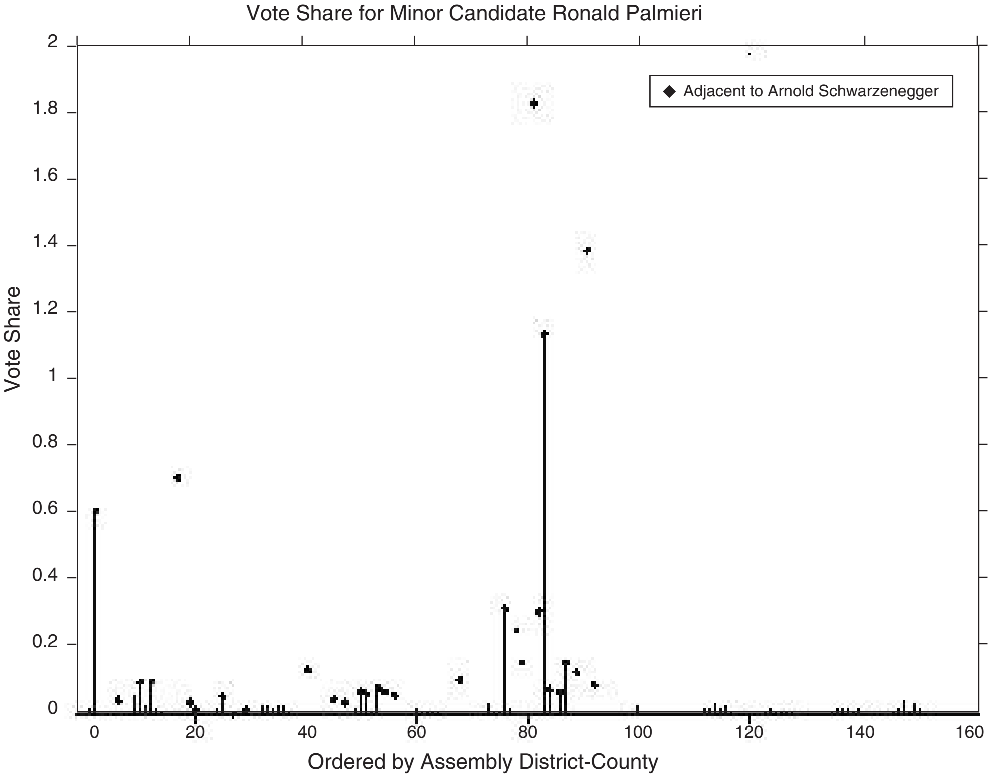


FIGURE 1

Notes: The percentage vote shares for minor candidate Palmieri are ordered from left to right by assembly district (numbered 1 through 80) and alphabetically by county within each district. A diamond marks Palmieri's vote shares in county-districts in which he is adjacent to major candidate Arnold Schwarzenegger on the ballot. We see that Palmieri's vote share never exceeds 0.08 percent when he is not adjacent to Schwarzenegger on a ballot. Palmieri's vote share is significantly higher (by as much as 1.8 percentage points) in county-districts in which he is adjacent to Schwarzenegger, suggesting the effects of misvoting.

We therefore believe our results on misvoting also apply to elections that do not share the specifics of the recall election (such as name rotation or a large number of candidates).

We find that the vote shares of minor candidates almost double when their names are adjacent to the names of major candidates. Misvotes account for at least 0.30 percent of all votes cast.[3] We find that the amount of misvoting depends on the voting technology used. Punch card technologies yield about twice the rate of misvoting compared to optical scan and touch screen platforms, which confirms findings by Michael R. Alvarez et al. (2004) and Thomas S. Dee (2007). Controlling for technology, the amount of misvoting also varies with voter characteristics. Using precinct-level voter registration and census data, we find that, all else equal, vote share gains are larger in precincts with higher percentages of poorly educated, poor, or third-party voters.

[3] This is a lower bound because we only consider votes gained by candidates immediately adjacent to the top three frontrunners in our calculation of misvotes.

After controlling for the relative popularity of the two major candidates, we find that the vote share lost by the Democratic frontrunner, Bustamante, exceeds the vote share lost by his Republican rival, Arnold A. Schwarzenegger, by 62 percent. While Schwarzenegger beat Bustamante by a comfortable 17 percentage points, the disproportionate misvoting would have given Schwarzenegger a 0.07 percent advantage if both had been tied at 47 percent of the vote. This margin is larger than the margins of victory in the 2000 US presidential, 2004 Washington gubernatorial, 1974 New Hampshire senatorial, and 1985 Indiana representative elections. Further, our estimate of the difference in vote share lost by the major candidates underestimates the difference that might occur in elections in which two major candidates are adjacent to each other, as is the case in many US presidential elections. In such situations, not only does one major candidate lose more votes to adjacent candidates, these losses are accrued by her opponent, thereby exacerbating differences in vote share.

Adjacency misvotes pose a unique challenge to policy makers. While candidate list rotation by district can eliminate the bias posed by positional misvoting, rotation does nothing to reduce the bias caused by adjacency misvoting. No matter how a candidate list is randomized across districts, major candidates will continue to lose votes to adjacent candidates. Our results suggest two strategies that may reduce the level of adjacency misvotes. First, some voting technologies (electronic voting and certain brands of optical scan ballots) lead to lower levels of misvoting. Second, we find that adjacency misvotes are most prevalent when it is difficult to discern which candidate selection box is associated with a given candidate name. For example, on multi-column ballots, a single candidate name might be bordered by two selection bubbles. Thus, clearer separation of the selection boxes corresponding to each candidate's name may reduce the level of adjacency misvotes.

I. The California Recall Election Dataset

A statewide special election was held in California on October 7, 2003 concerning the recall of Governor Gray Davis. Voters were presented with two recall questions:

(a) *"Should Gray Davis be recalled (removed) from the office of Governor?" Yes/ No*

(b) "*Candidates to succeed Gray Davis as Governor if he is recalled: Vote for one"* [*List of 135 names*]

Voters were asked to respond to the second question regardless of their response to the first question. If more than 50 percent answered "Yes" to the first question, the candidate with the plurality of the vote on the second question would become governor. Perhaps due to the low entry requirements (65 signatures and a $3,500 registration fee), a total of 135 candidates ran to replace Davis. The resulting candidate list was extremely diverse and included such personalities as actor Gary W. Coleman, erotic actress Mary "Mary Carey" Cook, and sumo wrestler Kurt E. "Tachikaze" Rightmyer.

Table 1—2003 California Recall Election Results

Rank	Candidate	Vote share	Number of votes
1	Arnold Schwarzenegger	48.60	3,747,446
2	Cruz M. Bustamante	31.63	2,439,133
3	Tom McClintock	13.33	1,027,926
4	Peter Miguel Camejo	2.77	213,547
5	Arianna Huffington	0.55	42,654
6	Peter V. Ueberroth	0.29	22,267
7	Larry Flynt	0.20	15,489
8	Gary Coleman	0.16	12,712
9	George B. Schwartzman	0.14	10,960
10	Mary Cook	0.13	10,129
⋮			
25	Edward Thomas Kennedy	0.0335	2,586
50	Michael J. Wozniak	0.0179	1,384
75	Scott A. Mednick	0.0103	791
100	Dennis Duggan McMahon	0.0067	517
135	Todd Richard Lewis	0.0022	172

Ballot format was also extremely diverse. The ballot layout and technology was determined at the county level (58 counties). The ballot ordering of candidates varied by assembly district (80 districts). Because multiple counties can be in the same district and vice versa, the recall election used 157 different ballots (county-district combinations). Candidates were ordered according to a randomized alphabet, "RWQOJMVAHBSGZXNTCIEKUPDYFL," in the first assembly district and then rotated one candidate at a time by district.

Davis was recalled as governor of California (55 percent of voters voted to recall him). Table 1 shows the distribution of vote shares among the candidates running to replace him. The top three candidates captured a total of 93.56 percent of the vote share, leaving an average vote share of 0.05 percent for each of the remaining 132 "minor" candidates. For the remainder of this paper, we consider the three frontrunners: Schwarzenegger (R), Bustamante (D), and Tom McClintock (R) as **major candidates**. Inclusion of the next most popular candidate, Peter M. Camejo, as a major candidate does not significantly change the baseline results. All candidates other than the top three are included in the analysis as **minor candidates**.[4]

California uses three types of voting technology platforms: punch card, optical scan, and touch screen. Punch card technology requires that voters use a stylus to punch out a prescored hole corresponding to the desired candidate. Some forms of punch card technology require that voters search for the candidate's number on a separate punch card grid while others place the prescored hole next to the candidates name. Optical scan technology requires that voters fill in an oval or other marking to the left or right of each candidate's name. Finally, touch screen technology requires that voters touch a box on the computer screen in order to select their desired candidate. After selection, the screen asks the voter to confirm her choice. Technology is determined at the county level. In our sample, 45 percent of voters (19 counties) use

[4] Results do not change significantly if (a) Camejo is dropped from the dataset, (b) all observations representing minor candidates earning more than 5 percent of the vote share in a precinct are dropped from the dataset, or (c) the definition of minor candidates is restricted to include only those candidates in the lower half of the vote share distribution.

punch card, 43 percent (33 counties) use optical scan, and 11 percent (4 counties) use touch screen technology. Each technology platform has several technology brands. The distribution of brands is presented in Appendix B.

We utilize sample ballots mailed to voters for 152 county-district combinations to code which candidates had which type of advantaged position in each county-district.[5] We obtain precinct-level statements of vote and registration data from the Institute of Governmental Studies (IGS) at the University of California, Berkeley. The statement of vote data has precinct-level vote tallies for every candidate.[6] The registration dataset contains precinct-level registration data such as voter party registration, race, and age group. Because registration data does not include other important demographic variables such as income and education, we merge precinct-level data with US Census data from 2000 at the blockgroup level using IGS conversion files. Summary statistics of these demographic variables are available in Appendix B.

II. Empirical Strategy

In general, it is difficult to measure the level of adjacency misvotes because misvotes are recorded as valid votes. This may be the reason why previous literature (e.g., Alvarez et al. 2001 and Henry E. Brady et al. 2001) has focused on the residual vote—votes that are discarded because of hanging chads or other mismarkings. To our knowledge, the papers written by Alvarez et al. (2004) and Dee (2007) are the only papers that analyze adjacency misvotes. Using county-level data from the 2003 California recall election, Dee argues that the positive correlation between the vote shares of a major candidate and the vote shares of "bookend candidates"—minor candidates listed immediately before or after the major candidate—is evidence of voting mistakes. Dee shows that the vote share of a bookend candidate increases by roughly 0.15 percent of the vote share increase of the adjacent major candidate. Alvarez et al. (2004) focus on electoral complexity in general. However, they note a possible "vertical proximity" effect, i.e., that the gain in votes to a vertically adjacent minor candidate is equal to 0.4 percent of the votes received by the corresponding major candidate. Consistent with our findings, both papers find that punch card technologies lead to significantly higher levels of misvoting.

Both Alvarez et al. (2004) and Dee (2007) restrict their analysis to the limited number of candidates that are almost always vertically adjacent to major candidates and therefore cannot test vote share gains caused by variation in the identity of adjacent candidates. Thus, the critical assumption behind their results is that the vote shares of the four bookend or vertically proximate candidates are not correlated with the vote shares of major candidates due to reasons other than adjacency misvotes. This assumption may be invalid given that the names of bookend or vertically proximate

[5] The actual ballots used on election day are identical to the sample ballots. We have been unable to collect sample ballots from Stanislaus and Imperial counties (which account for 1.4 percent of the vote). In the case of touch screen technology, voters viewed electronic screens that displayed candidate names in the same layout as was presented on the sample paper ballots.

[6] We are grateful to IGS for updating their files over the course of our conversations. However, precinct-level data from IGS remains incomplete due to incomplete geography matching files. Full statement of vote and registration data along with matching files account for 85 percent of all votes cast in the recall election and cover all California counties. The missing IGS data is not concentrated in any county.

minor candidates closely resemble the names of major candidates. For example, Dee's results are strongest for a bookend candidate named George B. Schwartzman. Since "Schwartzman" looks similar to "Schwarzenegger" and Schwartzman's name is almost always adjacent to Schwarzenegger's name on the ballot, it is difficult to separate the effects of name confusion from adjacency-related mistakes. Thus, Dee and Alvarez's results may be specific to elections with a large number of candidates where there are likely to be minor candidates with names resembling those of major candidates. In contrast, we use random variation in the identity of adjacent candidates, drawn from a pool of more than 50 minor candidates.

To better capture variation in adjacency, we define an adjacent candidate as any minor candidate that vertically or horizontally borders a major candidate on a particular ballot (see sample ballots in Appendix A). We call these adjacent candidates north, south, east, and west adjacent. Identification of adjacency misvoting results from random variation in when a minor candidate is an adjacent candidate. North and south adjacent candidates are identified due to candidate name rotation through page and column breaks. Because north and south adjacent candidates are located in the same column as the major candidate, they are always present and the same unless column and page breaks cause the major candidate to be listed at the top or bottom of a column.

Identification of east and west adjacent candidates occurs because column length varies by county. In the 60 percent of counties (corresponding to 51 percent of voters) that use multi-column ballots, the candidates horizontally adjacent to the major candidates depend on the county's column length. For example, if Schwarzenegger is listed in the middle column on a three-column ballot, he has east and west adjacent candidates. If Schwarzenegger is listed in the left-most column on a two-column ballot, he only has an east adjacent candidate. Further, this east adjacent candidate may not be the same east adjacent candidate as in the previous ballot if the two ballots have columns of differing lengths. Statistics describing the identification of each type of adjacent candidate are available in Appendix B.[7] Because the data contains greater variation in horizontal adjacency than in vertical adjacency, we retest all specifications using only horizontal adjacency. For brevity, we do not report those results, but they are very similar to those derived from the original specifications.

Combining district-level candidate name rotation and county-level variation in ballot layout, we identify 57 minor candidates that are adjacent to a major candidate in some but not all county-districts. However, since even minor candidates who are never adjacent help to pin down coefficients on control variables, the vote shares of all minor candidates are used in our baseline specification:

$$Voteshare_{pdc} = \beta_0 + Adjacent_{pdcm}\,\beta_1 + controls + \varepsilon_{pdc}, \quad (1)$$

[7] We also identify a special class of adjacent candidates with numbers located in squares adjacent to the number corresponding to the major candidate on a punch card (see ballot 2 in Appendix A). These punch card adjacent candidates only exist in the 13 percent of counties (corresponding to 39 percent of voters) that use the Votomatic or Pollstar brands of voting machines. Because column length in the punch card grid differs from column length in the ballot text, a candidate can be punch card adjacent without being adjacent on the regular ballot and vice versa.

where

$$Adjacent_{pdcm} \equiv I_{dc}^{adjacent} \times Major\ Candidate\ Voteshare_{pdcm}.$$

An observation is a minor candidate c in a precinct p, with voting technology and ballot layout varying by county-district d. Each observation is weighted by the number of votes cast in precinct p. $Voteshare_{pdc}$ is defined as the share of votes candidate c receives in precinct p in county-district d. We do not use the log of vote share as the dependent variable because the vote share gained by an adjacent candidate should be a level effect and should not depend on the average vote share of the minor candidate. $I_{dc}^{adjacent}$ is a dummy variable equaling one if candidate c is adjacent to a major candidate in county-district d. In modifications of the baseline regression, $I_{dc}^{adjacent}$ may represent specific types of candidate adjacency (e.g., the minor candidate listed directly above major candidate Schwarzenegger) or a vector representing several types of candidate adjacency. If minor candidate c is adjacent to a major candidate in precinct p in county-district d, then $Major\ Candidate\ Voteshare_{pdcm}$ equals the vote share in precinct p of the major candidate m to whom the minor candidate is adjacent. $Adjacent_{pdcm}$ is defined as the product of $I_{dc}^{adjacent}$ and $Major\ Candidate\ Voteshare_{pdcm}$. Finally ε_{pdc} is an error term that we allow to be clustered by county-district because ballot layout varies at that level. Our sample consists of observations at the candidate-precinct level. There are 132 data points per precinct, one for each minor candidate. This results in a total of 1,839,024 observations covering 13,932 precincts and corresponding to 80 districts and 56 counties.

In the baseline specification, β_1 represents the gain in vote share experienced by an adjacent minor candidate when the corresponding major candidate's vote share increases by one. In other words, β_1 measures the percentage of voters originally intending to vote for a major candidate that in actuality vote for *each* adjacent minor candidate. (The total percentage of voters that misvote is not β_1 but rather the product of β_1 and the average number of minor candidates adjacent to major candidates per ballot.) Implicit in this interpretation of β_1 is that adjacency misvoting is a positive function of the popularity of the associated major candidate; i.e., if Schwarzenegger is more popular in precinct A than in precinct B and Schwarzenegger's voters are equally likely to misvote in precincts A and B, then candidates adjacent to Schwarzenegger should gain more vote share in precinct A than in precinct B.

We also introduce the following set of controls to the baseline regression and all extensions of the baseline regression: candidate ballot location controls (dummy variables for first and last on the ballot overall, ballot page, and ballot column), county-district fixed effects, candidate fixed effects, candidate fixed effects interacted with demographic controls, and candidate fixed effects interacted with the vote shares of the major candidates. Altogether, 2,402 control variables are included in our baseline specification to ensure correct identification of adjacency misvoting.

The ballot location controls are county-district-candidate level dummy variables that equal one if the minor candidate is located first or last overall, on a page, or in a column on the ballot. County-district fixed effects control for the relative popularity of all minor candidates compared to major candidates in each county-district. The candidate fixed effects control for differences in the popularity of minor candidates that could lead to bias if relatively popular minor candidates happen to be adjacent to major candidates on a disproportionate share of the ballots. We also control for candidate fixed effects interacted with precinct-level demographic characteristics. This ensures that adjacency misvoting is not incorrectly identified from precinct demographic characteristics, which, by chance, might favor candidates that are adjacent in that precinct. Similarly, we control for candidate fixed effects interacted with the precinct-level vote share of each major candidate to control for the correlation between the popularity of each minor candidate and the popularity of the major candidates. In the Section III, we introduce several modifications of the baseline specification to explore adjacency misvoting as it relates to vote share lost by the major candidates, precinct-level demographic characteristics, and voting technology.

III. Results

A. *The Baseline Result*

Regression (1) in Table 2 shows the results of the baseline specification. It is a least squares regression of precinct-level vote shares of each minor candidate on *Adjacent*, a variable equal to the vote share of the associated major candidate if the minor candidate is adjacent to a major candidate and zero otherwise. As noted previously, standard errors are clustered by county-district. Because vote shares of major candidates are expressed as fractions and vote shares of minor candidates as percentages, the coefficient on *Adjacent* represents the percentage of voters that misvote for each minor candidate adjacent to their preferred major candidate. Regression (1) reveals that each minor candidate adjacent to a major candidate will experience an increase in votes equal to 0.105 percent of the intended votes for that major candidate. If we suppose that there is an average of three adjacent candidates per major candidate, this implies that 0.315 percent of voters cast misvotes. This result is significant at the 1 percent level and is clear evidence of the bounded rationality of some voters.[8] A more precise calculation, which takes into account the magnitude of adjacency misvoting for each type of adjacency (i.e., north, south, east, and west) along with the average number of adjacent candidates per ballot, reveals that misvotes accounted for at least 0.30 percent (s.e. = 0.04 percent) of all votes cast, corresponding to an error rate of roughly 1 in 300 (see Appendix C).

Regression (2) in Table 2 decomposes adjacency misvoting by major candidate. It shows a regression of minor candidate vote share on the interaction between *Adjacent* and a vector of dummy variables representing each major candidate. We find that Bustamante loses 0.144 percent of his vote to each of his adjacent minor candidates, but that the corresponding figure is only 0.089 percent for Schwarzenegger and 0.109 percent for McClintock. This result is striking because it implies that Bustamante

[8] We also test that the results are insensitive to weighting. If we do not weight observations by the number of votes in each precinct, the coefficient on *Adjacent* becomes 0.116 (s.e. = 0.027). We repeat the complete set of analyses using unweighted observations and find very similar results as long as we exclude the very smallest precincts (precincts with less than 50 votes, accounting for 0.27 percent of total votes).

Table 2—Baseline Results

Dependent variable *Voteshare* = (votes/total votes) × 100	(1)	(2)	(3)
Adjacent	0.105*** (0.024)		
Adjacent × Schwarzenegger		0.089*** (0.025)	
Adjacent × Bustamante		0.144*** (0.028)	
Adjacent × McClintock		0.109** (0.046)	
Adjacent dummy			0.037** (0.009)
Observations	1,839,024	1,839,024	1,839,024
R^2	0.8686	0.8687	0.8686

Notes: Robust standard errors, adjusted for clustering on county × district (152 clusters), are reported in parentheses. *Adjacent* equals the precinct-level vote share of the associated major candidate if a minor candidate is adjacent to a major candidate and zero otherwise. *Schwarzenegger*, *Bustamante*, and *McClintock* are dummy variables that equal one if a minor candidate is adjacent to major candidates Schwarzenegger, Bustamante, or McClintock, respectively. Observations are weighted by the total number of votes cast in a precinct. Controls are included for candidate fixed effects, candidate fixed effects interacted with demographic controls (percent black, percent Hispanic, percent Asian, percent age 18–24, percent age 65 plus, percent in poverty, median household income in $1,000 per year, percent high school graduates, percent college graduates, percent lacking English fluency, percent Republican, percent other party affiliation, and percent independent—percent Democrat is the omitted category), candidate fixed effects interacted with the precinct-level vote shares of major candidates Schwarzenegger, Bustamante, and McClintock, and ballot position controls (dummy variables for first and last overall, on a ballot page, and in a ballot column) representing the ballot position of the minor candidate in each county-district.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

voters are 62 percent more likely to misvote than are Schwarzenegger voters (the difference is significant at the 5 percent level). This result has important consequences because vote share gained by an adjacent candidate is vote share lost by the associated major candidate; if Schwarzenegger and Bustamante had received equal shares of the votes received by the major candidates, misvotes would have given Schwarzenegger an edge of 0.07 percent. In later regressions, we present evidence that Bustamante lost more votes than the other major candidates because he attracted voters from demographic groups associated with higher levels of misvoting.

Regression (3) in Table 2 shows a variation of the baseline regression in which minor candidate vote share is regressed on *Adjacent Dummy*, a dummy variable equaling one if a candidate is adjacent to a major candidate. The coefficient on *Adjacent Dummy* represents the average gain in vote share a minor candidate experiences when she is adjacent to a major candidate. Results show that minor candidates gain 0.037 percentage points in vote share if they are adjacent to a major candidate. Given that the average vote share of minor candidates is 0.049 percent, being adjacent to a major candidate nearly doubles a minor candidate's vote share.

Table 3 presents a series of regressions that support the robustness of our primary findings and that explore the mechanisms underlying misvoting. Regression (1) in Table 3 shows a regression of minor candidate vote share on *Adjacent Dummy*, a dummy variable equaling one if a minor candidate is adjacent to one of the major candidates and *Adjacent*, equal to the product of *Adjacent Dummy* and the precinct-level vote share of an adjacent candidate's associated major candidate. Regression (1) tests our assumption that adjacent candidates gain vote share as a positive function of the popularity of the major candidate. We find that this assumption is valid, adjacent candidates gain an additional 0.00081 percentage points for each additional

Table 3—Robustness Checks and Mechanism

Dependent variable: *Voteshare* = (votes/total votes) × 100	(1)	(2)	(3)	(4)	(5)	(6)
Adjacent	0.081***			0.105***	0.105***	
	(0.019)			(0.024)	(0.024)	
Adjacent dummy	0.011**					
	(0.006)					
Adjacent dummy × CA voteshare		0.113***				
		(0.026)				
North adjacent			0.081***			0.081***
			(0.022)			(0.022)
South adjacent			0.121***			0.121***
			(0.034)			(0.034)
East adjacent			0.143***			
			(0.046)			
West adjacent			0.038***			
			(0.013)			
Diagonally adjacent				0.004		
				(0.003)		
Punchcard adjacent					0.032**	
					(0.015)	
Horizontally adjacent						0.031***
						(0.009)
Horizontally adjacent × confusing side						0.123**
						(0.049)
Observations	1,839,024	1,839,024	1,839,024	1,839,024	1,839,024	1,839,024
R^2	0.8687	0.8786	0.8687	0.8686	0.8687	0.8687

Notes: Robust standard errors, adjusted for clustering on *county* × *district*, are reported in parentheses. If a minor candidate is adjacent to a major candidate, *CA Voteshare* equals the average statewide vote share of the associated major candidate. *North, South, East*, and *West Adjacent* equal the precinct-level vote share of the associated major candidate if a minor candidate is north, south, east, or west adjacent, respectively, and zero otherwise. *Diagonally Adjacent* and *Punchcard Adjacent* equal the precinct-level vote share of the associated major candidate if a minor candidate diagonally borders a major candidate (i.e., is listed to the northeast, northwest, southeast, or southwest of a major candidate) or is adjacent to a major candidate on a punch card grid, respectively. *Horizontally Adjacent* equals the precinct-level vote share of the associated major candidate if a minor candidate is east or west adjacent and zero otherwise. *Confusing Side* is a dummy variable equaling one if the minor candidate is horizontally adjacent to a major candidate and listed in the column with selection boxes closest to the name of the major candidate. All other controls and variables are as described in Table 2.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

percentage point gained by a major candidate. This result is significant at the 1 percent level.

Regression (2) in Table 3 tests that our baseline results are driven by variation in when a candidate is adjacent to a major candidate rather than by variation in the vote shares of associated major candidates. Instead of regressing minor candidate vote share on the product of an adjacency dummy and the *precinct-level* vote shares of associated major candidates, regression (2) uses the product of an adjacency dummy and the *statewide* vote shares of the associated major candidates as the independent variable. The coefficient on *Adjacent Dummy* × *CA Voteshare* is 0.113 and is significant at the 1 percent level. More importantly, it is very similar in magnitude to the coefficient of 0.105 on *Adjacent* in regression (1) of Table 2, confirming that our results are driven by variation in the adjacency status of minor candidates rather than variation in the popularity of a given major candidate.

Regression (3) in Table 3 decomposes adjacency misvoting by type of adjacency (north, south, east or west). This regression shows that misvoting is strongest when a minor candidate is east adjacent and weakest when a minor candidate is west adjacent. 0.143 percent of voters misvote for east adjacent candidates, only 0.038 percent of voters misvote for west adjacent candidates, while 0.081 percent and 0.121 percent of voters misvote for north and south adjacent candidates, respectively. We explore why east adjacent candidates gain more in vote share than west adjacent candidates in our discussion of regression (6). For now, we note that regression (3) shows that adjacency misvoting is not driven merely by name confusion. One might think that voters vote for north and south adjacent candidates because candidates are listed in order (according to a randomized alphabet) down a column. For example, on the majority of ballots, the south adjacent candidate for Schwarzenegger is Schwartzman, whose last name resembles "Schwarzenegger." We use candidate fixed effects to control for the direct effect of name confusion, but one might believe that it is the interaction between name confusion and adjacency that drives adjacency misvotes. However, east and west adjacent candidates are one column off from a major candidate and therefore have last names that are dissimilar to the names of their associated major candidates. The fact that these adjacent candidates also gain vote share of comparable magnitude suggests that it is adjacency rather than name confusion that drives adjacency misvotes.

Regression (4) in Table 3 explores the vote share gained by candidates that are diagonally adjacent to major candidates. Diagonally adjacent candidates are one row and one column off from their associated major candidate. Regression (4) shows that diagonally adjacent candidates earn much less in vote share than candidates that are horizontally or vertically adjacent. Only 0.004 percent of voters misvote for each diagonally adjacent candidate compared to the 0.105 percent of voters that misvote for each vertically or horizontally adjacent candidate. Low vote share gains for diagonally adjacent candidates likely occur because voters are less likely to misvote for candidates one row and one column removed from their intended choices. Because diagonally adjacent candidates gain very little in vote share, we restrict our analysis to vertically or horizontally adjacent candidates.

We now explore the possible mechanisms underlying misvoting. Misvotes can derive from unintentional mistakes when voters accidentally vote for candidates adjacent to their preferred choices. Misvotes can also result if voters first search for major candidates and then give increased consideration to adjacent minor candidates. Regression (5) in Table 3 explores the cause of misvoting by comparing adjacency effects on text ballots to adjacency effects on punch card grids. Thirty-nine percent of voters (13 percent of counties) used punch card voting machines that require each voter to search for her desired candidate and that candidate's associated number from a text ballot of candidate names. Each voter then punched out the desired candidate's number on a separate punch card grid. A sample punch card grid is shown in ballot 2 of Appendix A. Because the column length on the punch card grid differs from the column length in the text ballot, a minor candidate can be adjacent to a major candidate on a punch card grid without being adjacent on the text ballot and vice versa. The punch card grid lists only numbers, so vote share gained by punch card grid adjacent candidates should represent gains due to voting mistakes. Voters are unlikely to give increased consideration to numbers adjacent to the major candidate's number if they do not see any names associated with those numbers. Regression (5) shows that, controlling for the adjacency of minor candidates on the text ballot, 0.032 percent of voters misvote for each punch card grid adjacent candidate. This result is significant at the 5 percent level and supports our hypothesis that adjacency misvotes do not derive only from increased consideration of adjacent candidates.

Regression (6) in Table 3 presents further evidence that adjacency misvotes are driven by accidental votes rather than by increased consideration. In a single row of ballot text, we might read from left to right: bubble, minor candidate, bubble, major candidate, bubble, minor candidate. See ballot 1 in Appendix A for an example. In this case, the bubble for the east adjacent candidate is very close to the name of the major candidate and a voter may accidentally fill in the east adjacent candidate's bubble believing that she has cast a vote for the major candidate. In this example, we would label the east adjacent candidate as the "confusing" candidate and the west adjacent candidate as the nonconfusing candidate. Note that the actual names of the east and west adjacent candidates are equally distant from the name of the major candidate. Accidental mistakes should lead voters to disproportionally misvote for the more confusing adjacent minor candidate. Increased consideration should lead voters to give equal attention to the minor candidates listed to the left and to the right of a major candidate. We find that voters overwhelmingly misvote for the more confusing of the two horizontally adjacent minor candidates.[9]

We regress minor candidate vote share on *Horizontally Adjacent* and *Horizontally Adjacent* × *Confusing Side*. *Confusing Side* is a dummy variable equaling one if the minor candidate is horizontally adjacent to a major candidate and listed in the column with selection boxes closest to the name of the major candidate. We find that the coefficient on *Horizontally Adjacent* is 0.031 while the coefficient on *Horizontally Adjacent* × *Confusing side* is 0.123, which implies that voters are five times more likely to misvote for candidates listed on the confusing side. It follows that adjacency misvotes are, in large part, driven by accidental voting mistakes, because increased consideration should not depend on how each candidate's bubble is located relative to the major candidates. Regression (6) may also explain why the coefficient on *East Adjacent* is larger than the coefficient on *West Adjacent* in regression (3). For 64 percent of the multi-column ballots used in the election, the selection boxes of east

[9] Increased consideration should be much more likely than accidental mistakes to lead voters to misvote for candidates two positions north or south of a major candidate. We find that 0.011 percent of voters misvote for each minor candidate two positions removed from a major candidate. While this result is significant at the 5 percent level, it is 10 times smaller than the effect measured in the baseline specification, suggesting that the impact of increased consideration is small. Increased consideration would also predict the adjacency effect to be stronger for better-known minor candidates because voters presumably only give increased consideration to candidates they know something about. We test this prediction by adding to our baseline regression an interaction between *Adjacent* and the average vote share of the minor candidate in county-districts where she is not adjacent to any major candidate (the direct effect of the average minor candidate vote share is absorbed by the candidate fixed effects). This interaction effect is not statistically significant (p-value: 0.40) indicating that increased consideration is not likely an important driver of misvoting since the adjacency gain is not significantly higher for better-known minor candidates.

adjacent candidates were closest to the names of the major candidates. Finally, the best policy to reduce misvoting depends on the underlying cause of the misvoting. Our finding that misvoting is driven by accidental votes (or pure confusion) rather than increased consideration or name confusion delivers a clear policy implication. Adjacency misvoting can be reduced if ballot layout clearly specifies which selection box corresponds to each candidate's name.

B. *Demographic Characteristics*

The cognitive costs of voting, given complex ballot designs and voting technology, may increase with certain demographic characteristics. This leads to systematic disadvantages for major candidates that attract voters with characteristics associated with misvoting. Below, we show how the amount of misvoting varies with precinct-level demographic characteristics. We find that voters from precincts with high misvote rates disproportionally favored Democratic candidate Bustamante and that voter demographics may explain why Bustamante lost a higher proportion of his votes to adjacent candidates than did Schwarzenegger or McClintock.

A growing literature suggests that the level of residual votes, i.e., votes that are discarded because voters vote for more than one candidate, leave hanging chads, etc., depends on the racial and socioeconomic composition of each precinct. Michael Tomz and Robert P. Van Houweling (2003) find that in areas with punch card or optically scanned ballots, the black-white gap in residual vote share ranged from 4 to 6 percentage points while lever and touch screen machines cut the gap in residual vote shares by a factor of 10. Controlling for income and education, Stephen Knack and Martha Kropf (2003) find similar results.

To our knowledge, no research has yet explored how the prevalence of misvoting varies by demographic characteristics. To test how misvoting varies by precinct-level demographic characteristics, we interact the variable *Adjacent* with a vector of precinct-level demographic characteristics, $\mathbf{Char}_{pd}$, and add this interaction to the baseline specification.[10] If the voting technology chosen by a county varies with its demographic characteristics, our estimates of the effect of demographics on misvoting could be confounded by the effect of voting technology on misvoting. To prevent this, we add controls for $\mathbf{Tech}_d$ and *Adjacent* $\times$ $\mathbf{Tech}_d$, where $\mathbf{Tech}_d$ is a vector of technology dummy variables representing the brand of technology used in each county.

There are two caveats to the demographic results. First, it is difficult to establish causation. For example, it is conceivable that education is correlated positively with experience taking standardized exams, and experience with standardized exams is what truly reduces levels of misvoting, not more education. Second, the ecological fallacy summarized in Christopher H. Achen and W. Phillips Shively (1995) implies that analysis using precinct-level data does not necessarily imply direct links between individual demographics characteristics and the magnitude of adjacency

[10] The direct effect of $\mathbf{Char}_{pd}$ is absorbed by the $\mathbf{Char}_{pd}$ $\times$ *candidate fixed effects* interactions that are already included in the baseline specification.

misvoting.[11] Nevertheless, we believe that the effect of precinct-level demographics on misvoting likely offers a guide to the effects of individual demographics.

Table 4 shows the results of a regression of minor candidate vote share on *Adjacent* and the interaction between *Adjacent* and a vector of precinct-level demographics. We find that a 1 percentage point increase in the number of voters with a high school degree corresponds to a 0.0091 percentage point fall in the percentage of voters that misvote for each adjacent candidate. This result is significant at the 1 percent level. The coefficient on *Adjacent* $\times$ *Percent College Graduates* is similar with a value of -0.0072, suggesting that education beyond the high school level has no additional impact on misvoting. The calculations in the last three columns of Table 4 show the difference in the percentage of voters that misvote for each adjacent candidate for precincts at the fifth percentile of a demographic variable compared to precincts at the ninety-fifth percentile. Holding the other demographic variables constant, the percentage of voters that misvote for each adjacent candidate is 0.328 percentage points lower in precincts at the ninety-fifth percentile of high school education than in precincts at the fifth percentile. This difference is very large—it is triple the average percentage of voters that misvote for each adjacent candidate. This result is consistent with evidence from Daniel J. Benjamin, Sebastian A. Brown, and Jesse M. Shapiro (2006) and Lusardi and Mitchell (2007). Lusardi and Mitchell also find that lack of education is a strong predictor of making mistakes but in the context of simple questions about interest, inflation, and risk diversification. Benjamin, Brown, and Shapiro (2006) present experimental evidence that lower cognitive ability, rather than lack of schooling per se, is a powerful predictor of behavioral anomalies.

We also find that, holding other demographics constant, a 1 percentage point increase in Asian residents corresponds to a 0.0013 percentage point increase in the percentage of voters that misvote for each adjacent candidate. This result is significant at the 10 percent level. A one percentage point increase in Hispanic residents corresponds to a 0.0021 percentage point increase, but this result is insignificant. Meanwhile, the coefficient on *Adjacent* $\times$ *Percent Black* is small with a standard error near zero, implying that precincts with high percentages of black residents do not behave significantly differently from precincts with high percentages of white residents (the omitted category). Holding the percentage of black and Hispanic residents constant, the percentage of voters that misvote for each adjacent candidate is 0.046 percentage points higher in precincts at the ninety-fifth percentile of percent Asian than in precincts at the fifth percentile. This difference is equal to half the average percentage of voters that misvote for each adjacent candidate.

Controlling for other demographic factors, we find that a 1 percentage point increase in the number of voters registered with third parties corresponds to a 0.0042 percentage point increase in the level of misvoting and is significant at the 1 percent level. Precincts with higher percentages of registered Republicans or Independents do not misvote at significantly different rates relative to precincts with higher percentages of

[11] For example, a regression might show that a precinct with a 60 percent Asian and 40 percent white population casts misvotes at a higher rate than does a precinct with a 10 percent Asian population and a 90 percent white population. This does not necessarily imply that Asian voters are more likely to cast misvotes. It could instead be the case that the white voters living in predominantly Asian precincts cast misvotes at a higher rate.

Table 4—Effect of Precinct Demographic Characteristics on Misvoting

Dependent variable: *Voteshare* = (votes/total votes) × 100			5th percentile	95th percentile	Difference in adjacency effect
Adjacent	0.8224***	(0.2195)			
Adjacent × percent black	−0.0003	(0.0008)	0.00	0.18	−0.0054
Adjacent × percent Hispanic	0.0021	(0.0015)	0.04	0.64	0.1260
Adjacent × percent Asian	0.0013*	(0.0007)	0.01	0.36	0.0455
Adjacent × percent age 18–24	−0.0014	(0.0011)	0.03	0.14	−0.0154
Adjacent × percent age 65 plus	0.0011*	(0.0006)	0.07	0.36	0.0319
Adjacent × percent in poverty	0.0009**	(0.0004)	0.01	0.32	0.0279
Adjacent × income (K)	−0.0003	(0.0006)	12K	55K	−0.0129
Adjacent × percent HS graduates	−0.0091***	(0.0035)	0.24	0.60	−0.3276
Adjacent × percent college graduates	−0.0072***	(0.0025)	0.12	0.71	−0.4248
Adjacent × percent Republican	0.0010	(0.0013)	0.09	0.60	0.0510
Adjacent × percent other party	0.0042***	(0.0016)	0.03	0.21	0.0756
Adjacent × percent Independent	−0.0005	(0.0018)	0.00	0.24	−0.0120
Adjacent × percent lacking English fluency	−0.0071*	(0.0043)	0.01	0.26	−0.1775
Observations	1,839,024				
R^2	0.8690				

Notes: Robust standard errors, adjusted for clustering on county × district, are reported in parentheses. *Adjacent* equals the precinct-level vote share of the associated major candidate if a minor candidate is adjacent to a major candidate and zero otherwise. Omitted categories: *Percent high school dropouts, Percent other ethnicity, Percent Democrat. Percent HS graduates* and *Percent college graduates* represent the percentage of precinct residents with high school degrees and college degrees (*Percent HS graduates* does not include those who have also graduated from college). *Percent black, Percent Hispanic*, and *Percent Asian* represent the percentage of precinct residents that identify themselves as black, Hispanic, or Asian, respectively (the omitted group includes white and other ethnicities). *Percent Democrat, Percent Republican*, and *Percent other party* represent the percentage of registered voters that are officially registered as a Democrat, Republican, or other party member; all remaining voters are included in *Percent Independent*. Income represents median household income of precinct residents in $1,000 per year. *Percent in poverty* represents the percentage of households in a precinct below the census poverty line. *Percent lacking English fluency* represents the percentage of precinct residents lacking fluency in English. *Percent age 65 plus* and *Percent age 18 to 24* represents the percentage of registered voters in a precinct above the age of 65 and between the ages of 18 and 24, respectively. Controls are included for $\mathbf{Tech}_d$ and $Adjacent_{pdcm} \times \mathbf{Tech}_d$, where $\mathbf{Tech}_d$ is a vector of dummy variables representing the 11 brands of voting technology used in the recall election. All other controls and variables are as described in Table 2.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

registered Democrats (the omitted category). However, these results represent partial effects of party registration. If we do not control for other demographic variables, precincts with more registered Republicans or Independents misvote at significantly lower rates relative to precincts with more registered Democrats (not reported here, but see Shue and Luttmer 2006). Further, this effect is large. The difference in misvote rates for precincts at the niney-fifth and fifth percentiles of registered Republicans or Independents is greater than the average rate of misvoting. This discrepancy between partial and total effects occurs because the percentage of voters registered with the Democratic Party is negatively correlated with education, which is associated with lower levels of misvoting. This explains why the percentage of registered Democratic voters has a positive total effect on misvoting but a nonsignificant partial effect.

We also explore how adjacency misvotes vary with five other precinct-level demographic variables: median household income in thousands of US dollars, percent living below the census poverty line, percent lacking English fluency, percent over the age of 65, and percent between the ages of 18 and 24. We find that precincts with high percentages of residents in poverty exhibit significantly higher levels of misvoting. We find a marginally significant increase in the misvoting rate associated with a higher fraction of residents above the age of 65 and marginally significant decrease associated with the fraction lacking English fluency. We do not find significant effects for the other two demographic characteristics.

Since Table 4 presents partial demographic effects, it is important to note that if we do not control for other demographic factors, higher median household income and a lower percentage of young voters both correspond to significantly lower levels of misvoting (not reported here, but see Shue and Luttmer 2006). The total effects for low-income precincts are consistent with Madrian and Shea (2001), Choi et al. (2004), and Thaler and Benartzi (2004), who in different settings, find that low-income employees are more likely to conform to the default savings options set by their employers, which is likely because these employees have higher cognition costs of making optimal savings decisions. In addition, the total effects for very young and very old voters are consistent with Sumit Agarwal et al. (2007), who show that sophistication in financial decisions peaks in middle age.

This variation in misvoting by demographic characteristics sheds light on why, relative to Schwarzenegger and McClintock, Bustamante lost more votes to adjacent candidates. If the three candidates had been equally popular, Bustamante would have lost 62 percent more votes than Schwarzenegger and 32 percent more votes than McClintock. Using our estimates of the effect of demographics on misvoting, we calculate how much of the difference in vote share lost by each major candidate can be accounted for by differences in the demographic characteristics of each major candidate's voting base.[12] We find that, compared to Schwarzenegger and McClintock, Bustamante attracted voters most likely to misvote. Differences in the observed demographic characteristics of each major candidate's voting base can explain at least one-third of the difference in vote share lost.[13]

C. *Voting Technology*

Numerous studies and policy initiatives have explored the effect of voting technology on the number of discarded ballots, also known as residual votes. These studies consistently find that, relative to optical scan and electronic technology platforms, the punch card technology platform is associated with the greatest number of residual votes. Brady et al. (2001) summarizes the current research relating to voting

[12] We first estimate the number of misvotes cast in each precinct using each precinct's demographic characteristics. The estimated number of misvotes in each precinct multiplied by each major candidate's performance in a precinct (as a fraction of the total votes for the three major candidates) is an estimate of the number of misvotes lost by each major candidate. After adjusting for the relative popularities of the major candidates, we find that demographic characteristics predict that Bustamante lost 24 percent more votes than Schwarzenegger and 14 percent more votes than McClintock. Thus, demographic characteristics can account for at least 38 percent of the difference in votes lost by Bustamante and Schwarzenegger and 42 percent of the difference in votes lost by Bustamante and McClintock.

[13] Our calculation is actually a lower bound for the explanatory power of demographic characteristics. Since our data is at the precinct level, we need to assume that, within a precinct, Schwarzenegger, Bustamante, and McClintock voters are equally likely to misvote for adjacent candidates. This is likely to be false—within a precinct, Bustamante voters are probably more likely to cast misvotes. Therefore, we underestimate the extent to which demographic characteristics can account for differences in vote shares lost by the major candidates.

technology. On average, residual votes account for 2 percent of all votes in presidential races. Using county-level data from the 2000 US Presidential Election, Brady et al. find that touch screen and optical scan voting technology yield half the level of residual votes found in elections using punch card technology. Alvarez et al. (2001) employ similar methods to find results generally consistent with those of Brady et al.

Much less is known about the relationship between technology and adjacency misvotes. Dee (2007) finds that punch card technology increases the vote share gained by "bookend" candidates, i.e., candidates vertically adjacent to major candidates. Dee shows that, relative to touch screen and optical scan technology, punch card technology increases the vote share of bookends by 0.04 percentage points on average, an increase of more than one-third of the mean vote share of the bookends. We extend Dee's analysis to a precinct-level dataset that measures adjacency misvotes based upon variation in when candidates are vertically and horizontally adjacent to major candidates. We also extend Dee's analysis by exploring whether technology differences are driven by platforms (i.e., punch card) or by individual technology brands. Further, we control for precinct-level demographic characteristics. This will better control for the possibility that technology is correlated with demographic characteristics that influence misvoting.

To study the relationship between technology and adjacency misvoting, we interact the variable *Adjacent* with a vector of technology dummies $\mathbf{Tech}_d$ representing technology platforms or technology brands. We add this interaction and $\mathbf{Tech}_d$ itself to the baseline specification. By platform of technology, we refer to broad categories of technology such as touch screen, punch card, or optical scan technology. By brands of technology, we refer to specific brands of voting machines, e.g., the Diebold AccuVote-TS that uses the touch screen technology platform. For a summary of the availability of each type of technology, see Appendix B. In addition to the large set of controls included in the baseline regression, we also include controls for *Adjacent* $\times$ $\mathbf{Char}_{pd}$, where $\mathbf{Char}_{pd}$ is a vector of precinct-level demographic variables. This specification allows for the possibility that voting technology is correlated with precinct-level demographic variables and separates technology effects from demographic effects.

Regression (1) in Table 5 shows the relationship between voting technology platform and adjacency misvoting. Our results indicate that 0.202 percent of voters using punch card technology misvote for each adjacent candidate compared to only 0.109 percent and 0.068 percent of voters using optical scan and touch screen technology, respectively. Further, the coefficient on *Adjacent* $\times$ *Punch Card* is significantly different from the coefficients on *Adjacent* $\times$ *Optical Scan* and *Adjacent* $\times$ *Touch Screen* at the 1 percent level. These results suggest that touch screen and optical scan technology cut the prevalence of misvoting for each adjacent candidate roughly in half relative to punch card technology.[14] This result is consistent with the large

[14] The total misvoting rate for each technology depends on the misvoting rate per adjacent candidate for that technology, as presented in Table 5, and the average number of adjacent candidates for the types of ballot layouts used by that technology. Since the average number of adjacent candidates tends to be higher for touch screen than for optical scan and tends to be higher for optical scan than for punch card technology, the difference in the total amount of misvoting is less pronounced across the three technology platforms. The total misvoting rate is 0.325 (s.e. = 0.036) for punch card, 0.273 (s.e. = 0.049) for optical scan, and 0.211 (s.e. = 0.056) for touch screen. Thus, the ordering of technologies remains the same, but the differences in total misvoting rates are not statistically significant.

Table 5—Interactions with Voting Technology

	Dependent variable: *Voteshare* = (votes/total votes) × 100	(1)	(2)	(3)	(4)
	Adjacent × punch card	0.202***	0.198***	0.202***	
		(0.024)	(0.027)	(0.026)	
	Adjacent × optical scan	0.109***	0.101***	0.114***	
		(0.024)	(0.027)	(0.026)	
	Adjacent × touch screen	0.068***	0.065***	0.079***	
		(0.019)	(0.020)	(0.021)	
Optical	Adjacent × Diebold Accu-Vote-OS				0.171***
					(0.041)
	Adjacent × ES&S 550 Optech				0.083***
					(0.023)
	Adjacent × ES&S Optech Eagle				0.110***
					(0.033)
	Adjacent × Hart Ballot Now				0.067***
					(0.017)
	Adjacent × Mark-A-Vote				0.079**
					(0.032)
	Adjacent × Sequoia Optech				0.095***
					(0.024)
Touch	Adjacent × Diebold AccuVote-TS				0.007
					(0.040)
	Adjacent × Sequoia Edge				0.096***
					(0.018)
Punch	Adjacent × Datavote				0.096***
					(0.033)
	Adjacent × Pollstar				0.242***
					(0.032)
	Adjacent × Votomatic				0.213***
					(0.022)
	Demographic controls	Y	N	Y	Y
	Adjacent = Adjacent dummy × CA voteshare	N	N	Y	N
	Observations	1,839,024	1,839,024	1,839,024	1,839,024
	R^2	0.8689	0.8688	0.8689	0.8690

Notes: Robust standard errors, adjusted for clustering on county × district, are reported in parentheses. In regressions (1), (2), and (4) *Adjacent* equals the precinct-level vote share of the associated major candidate if a minor candidate is adjacent to a major candidate and zero otherwise. In regression (3), *Adjacent* equals the state-level vote share of the associated major candidate if a minor candidate is adjacent to a major candidate and zero otherwise. Controls are included for the direct effect of each technology variable. Demographic controls include $Adjacent_{pdcm} \times \mathbf{Char}_{pd}$, where $\mathbf{Char}_{pd}$ is a vector of precinct-level variables representing percent black, percent Hispanic, percent Asian, percent age 18–24, percent age 65 plus, percent in poverty, median household income, percent high school graduates, percent college graduates, percent lacking English fluency, percent Republican, percent other party affiliation, and percent independent (percent Democrat is the omitted category). All demographic variables are demeaned. All other variables and controls are as described in Table 2.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

literature (e.g., Alvarez et al. 2001 and Brady et al. 2001) arguing that punch card technology causes the greatest number of residual votes.

Regressions (2) and (3) in Table 5 use alternative specifications to confirm the robustness of our results. In regression (2), we do not control for precinct demographic characteristics. Regression (3) uses the state-level major candidate vote share rather than the precinct-level vote shares to determine the variable *Adjacent* = $I^{Adjacent} \times$ *Major Candidate Vote Share*. This ensures that results are driven by variation in adjacency status rather than by variation in precinct-level major candidate vote shares, which may be correlated with other unobservables. In regressions (2) and (3), we obtain results very similar to those from regression (1).

Regression (4) in Table 5 shows the relationship between the brand of voting technology and the magnitude of adjacency misvoting. Eleven brands of voting technology were used during the recall election. Each brand corresponded to one of three technology platforms: optical scan, touch screen, or punch card. Results from regression (4) reveal that, controlling for the impact of demographic characteristics, brands tend to have the same rate of misvoting within each technology class, with three notable exceptions: the optical scan Diebold Accu-Vote-OS has a level of misvoting that is comparable to typical punch card voting technologies, the punch card Datavote brand has a level of misvoting that is comparable to most optical scan and touch screen technologies,[15] and the touch screen Diebold AccuVote-TS is associated with an insignificantly low level of misvoting. In the interest of space, we do not present results for robustness checks similar to regressions (2) and (3) at the technology brand level. However, these alternative specifications for the brand-level regressions yield results very similar to those in regression (4).

In general, our results support evidence by Knack and Kropf (2002), Phillip Garner and Enrico Spolaore (2005), and David Card and Enrico Moretti (2007) showing that areas with low levels of education and income and high levels of black or Hispanic voters are not more likely to have "inferior" technologies such as punch card voting. Moreover, the technology brand analysis suggests that a simple switch away from the punch card technology platform may not automatically reduce the level of adjacency misvoting. Careful consideration of inter-brand differences is necessary.

D. *Applications and Extensions*

While the small but significant loss of votes to adjacent candidates did not change the outcome of the recall election, it may have important implications for elections in general. Numerous presidential, senatorial, and gubernatorial elections have been determined by very slim margins of victory. The margins of victory of the popular vote in the presidential elections of 1880, 1884, 1960, and 2000 were all less than one-quarter of a percent.[16] Recently, the 2004 Washington gubernatorial, 2000 Washington senatorial, 1998 Nevada senatorial, and 2002 South Dakota senatorial elections were determined by vote share margins of less than one-tenth of a percent.

[15] The Datavote punch card brand distinguishes itself from the other two punch card brands in that it allows voters to punch out their selections on the regular ballots (rather than requiring the use of punch card ballots in addition to regular text ballots). Thus, the low level of misvotes associated with the Datavote machine suggests that the act of "punching" out prescored holes may not be the cause of misvoting. Instead, higher levels of misvoting associated with punch card technology may be driven by the use of separate punch card ballots.

[16] Outcomes of presidential elections are officially determined by the Electoral College. For a majority of US states, the popular vote by state determines how state electoral votes are distributed in the Electoral College. In each of the presidential elections listed above, the margin of victory was less than a quarter of a percent in the popular vote of at least one state with enough electoral votes to determine the outcome of the election.

Major candidates that attract voters that are more likely to cast misvotes would suffer key disadvantages in these close elections.

However, one may wonder if results from California's gubernatorial recall election can be applied to other major elections. The recall election featured 135 candidates competing in a single race. The advantage is that the large number of candidates combined with variation in ballot layout and candidate order provides the best available identification of adjacency misvoting. The disadvantage is that most major elections feature fewer candidates competing in a single race. For example, while a total of 72 candidates ran for president in at least one state during the 2004 election, the majority of states listed fewer than 10 candidates. Even primary elections, known for their large candidate pools, typically feature fewer than 30 candidates per race.[17] Yet, this does not imply that the California recall election ballots were unusually confusing. Ballots used in major elections typically contain numerous races and propositions, and thus do not significantly differ in length or complexity from recall election ballots which contained only one race. For example, sample ballots offered by Diebold and ES&S, two companies that together serve the majority of US states, show extensive use of multicolumn ballots with a different race for each column. This creates the same problems of east and west adjacency as in the California recall election.

Future nationwide studies of adjacency misvoting may be particularly beneficial because adjacency misvoting could have a large impact on major elections outside of California. Unlike California, many states do not rotate candidate name order on ballots. Rather, incumbents and major candidates are listed first, followed by a long list of minor candidates. This implies that the candidate adjacent to a major candidate could be another major candidate. If this is the case, votes lost by one major candidate are systematically accrued by another major candidate, exacerbating the differential in vote shares lost by the major candidates.

IV. Conclusion

This paper tests the hypothesis that minor candidates experience gains in vote share when their names are listed near the name of a major candidate on a ballot. We refer to votes for minor candidates resulting from adjacency as "misvotes" because ballot layout should not impact the voting choices of fully rational voters with negligible cognition costs.

We test adjacency misvoting using a dataset combining precinct-level voting results from the 2003 California recall election with precinct-level census and voter registration data. Because ballot layout and candidate order are determined at the county-district level, there exists random variation in when a minor candidate is adjacent to a major candidate. The gain in the total vote shares experienced by minor candidates that are adjacent is an estimate of the number of misvotes. We find that the vote shares of minor candidates almost double when their names are adjacent to the names of major candidates and that misvotes accounted for at least 0.30 percent

[17] Data for presidential elections and primaries are drawn from state election Web sites.

of all votes cast. Misvotes become more prevalent as the fraction of poorly educated, poor, or third-party voters in a precinct increases, suggesting that these characteristics are associated with higher cognitive costs of voting. We also find that, relative to the punch card technology platform, the use of the optical scan or touch screen technology platforms generally corresponds to a reduction in the prevalence of misvoting. Even within technology platforms, however, there are some notable performance differences across voting machine brands.

Adjacency misvoting has important electoral implications because votes gained by adjacent candidates are votes lost by major candidates. A major candidate that attracts voters from demographic groups that cast relatively high levels of misvotes suffers systematic electoral disadvantages. In the case of the California recall election, precincts that were most affected by adjacency misvoting also showed the strongest support for Democratic candidate Bustamante. This helps to explain why, relative to Republican candidate Schwarzenegger, Bustamante lost 62 percent more of the votes intended for him to his adjacent candidates. Although this difference in lost votes represents a small fraction of the total vote share, the loss exceeds the margins of victory in several recent elections, including the presidential election of 2000 and the Washington gubernatorial election of 2004. Thus, adjacency misvoting is powerful enough to determine outcomes in important, highly contested close elections.

Appendix A: Sample Ballots

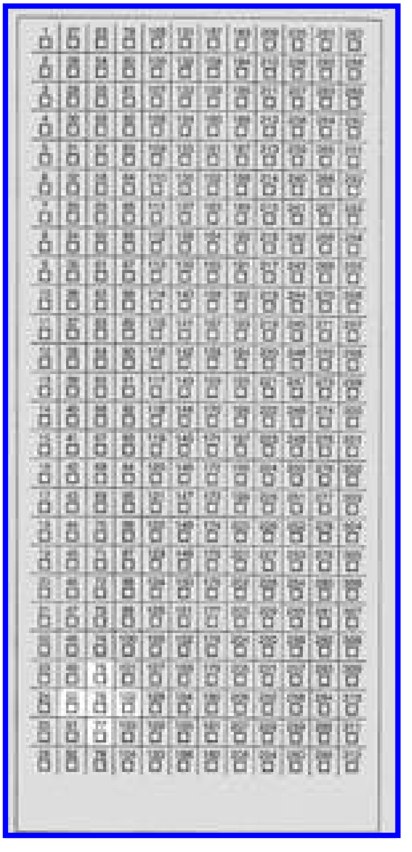
Ballot 1

The image below shows the upper portion of a multi-column optical scan ballot used in Alameda County. The major candidates and their adjacent minor candidates are highlighted in white. Consider major candidate Schwarzenegger, located

in the middle column, third from the bottom. Because the columns are not clearly separated by boundary lines, voters intending to vote for Schwarzenegger may accidentally fill in the bubble corresponding to Ronald Palmieri, Schwarzenegger's east adjacent candidate, instead of the correct bubble to the left of Schwarzenegger's name.

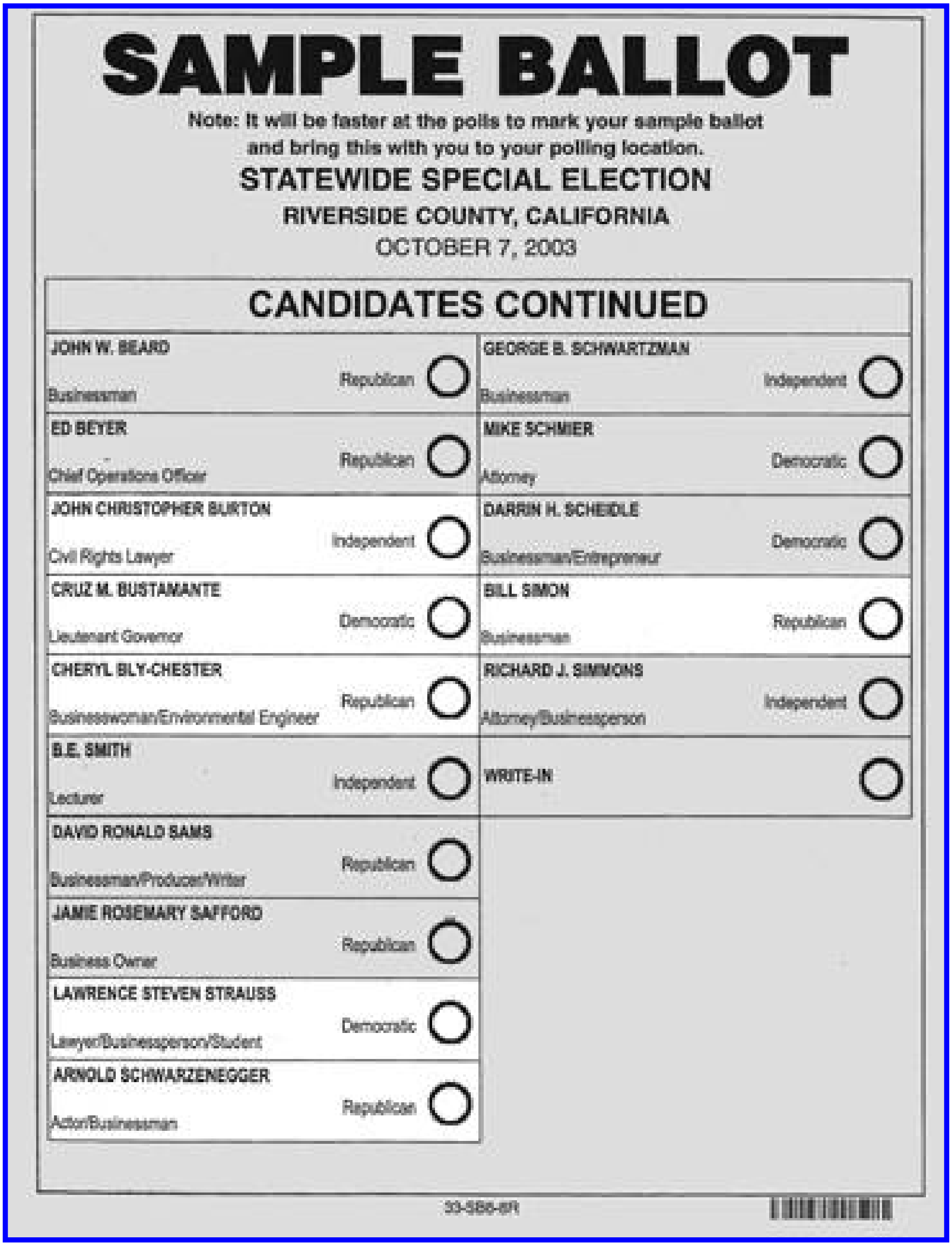
Ballot 2

The images below show a page from a single-column text ballot and a punch card grid used in Sacramento County. Major candidate Bustamante and his adjacent minor candidates are highlighted in white. Voters searched for the number corresponding to the desired candidate's name using standard text ballots. Each voter then used a stylus to punch out the desired candidate's number on a separate punch card grid. The small print suggests voters may accidentally punch out numbers corresponding to candidates adjacent to major candidates on the punch card grid.



Ballot 3

The following image shows a page from a multi-column touch screen ballot used in Riverside County. The major candidates and their adjacent minor candidates are highlighted in white. Voters selected their desired candidates by touching the



SAMPLE BALLOT

Note: It will be faster at the polls to mark your sample ballot and bring this with you to your polling location.

STATEWIDE SPECIAL ELECTION

RIVERSIDE COUNTY, CALIFORNIA

OCTOBER 7, 2003

CANDIDATES CONTINUED

Candidate	Party	Candidate	Party
JOHN W. BEARD Businessman	Republican	GEORGE B. SCHWARTZMAN Businessman	Independent
ED BEYER Chief Operations Officer	Republican	MIKE SCHMIER Attorney	Democratic
JOHN CHRISTOPHER BURTON Civil Rights Lawyer	Independent	DARRIN H. SCHEIDLE Businessman/Entrepreneur	Democratic
CRUZ M. BUSTAMANTE Lieutenant Governor	Democratic	BILL SIMON Businessman	Republican
CHERYL BLY-CHESTER Businesswoman/Environmental Engineer	Republican	RICHARD J. SIMMONS Attorney/Businessperson	Independent
B.E. SMITH Lecturer	Independent	WRITE-IN	
DAVID RONALD SAMS Businessman/Producer/Writer	Republican		
JAMIE ROSEMARY SAFFORD Business Owner	Republican		
LAWRENCE STEVEN STRAUSS Lawyer/Businessperson/Student	Democratic		
ARNOLD SCHWARZENEGGER Actor/Businessman	Republican		

33-SB8-6R

desired candidates' corresponding circles. A second screen then asked voters to confirm their choices. This ballot also shows how column breaks allow identification of vertically adjacent candidates. Minor candidate George Schwartzman normally follows major candidate Schwarzenegger in the candidate list. However, Schwartzman is not south adjacent on this ballot because he is separated from Schwarzenegger by a column break.

Appendix B: Summary Statistics

Table B1—Precinct Level Data

Variable	Mean	Standard deviation
Statement of vote and voter registration data		
Total votes cast in precinct	850.10	580.50
Percent voteshare of a minor candidate	0.05	0.35
Percent voteshare for Arnold Schwarzenegger	48.41	16.97
Percent voteshare for Cruz Bustamante	31.64	17.61
Percent voteshare for Tom McClintock	13.49	4.74
Percent registered with the Democratic party	37.12	48.31
Percent registered with the Republican party	30.30	45.96
Percent registered with other political parties	8.33	27.64
Percent not registered with any political party	24.24	42.85
Percent age 18 to 24	8.08	4.93
Percent age 65 plus	19.14	10.96
Census demographic data		
Percent white	68.58	19.44
Percent black	4.84	8.83
Percent Asian	10.36	11.47
Percent Hispanic	21.92	19.08
Percent lacking English fluency	7.64	8.46
Median household income ($1,000 per year)	27.97	14.03
Percent below the Census poverty line	10.98	10.55
Percent high school graduates (not including college graduates)	43.61	10.89
Percent college graduates	39.40	18.04
Observations	1,839,024	
Precincts	13,932	

Notes: Observations are at the candidate-precinct level. Observations do not include the top three vote share earners. Means and standard deviations are weighted by the total number of votes cast in each precinct.

Table B2—Identification of Adjacent Candidates

	Frequency[a]			Variation
Type	Schwarzenegger adjacent	Bustamante adjacent	McClintock adjacent	Number of minor candidates that occupy the adjacent position
North adjacent	93.7	96.8	94.2	3
South adjacent	91.3	96.1	92.7	3
East adjacent	35.6	39.0	35.5	25
West adjacent	23.6	26.6	15.4	20
Punchcard adjacent	9.0	9.0	9.0	18

Notes: When calculating frequency, observations are weighted by the total number of votes cast in each precinct. This is done because observations are weighted by precinct size in all regression specifications.

[a] Percent of precincts with ballots that contain this type of adjacency.

Table B3—Voting Technology

Technology brand	Technology platform	Prevalence percentage	Used again in 2005 election
Diebold Accu-Vote-OS	optical scan	12.3	Y
ES&S 550 Optech	optical scan	5.4	N
ES&S Optech Eagle	optical scan	5.1	Y
Hart Ballot Now	optical scan	10.6	Y
Mark-A-Vote	optical scan	4.3	Y
Sequoia Optech	optical scan	5.7	Y
Diebold AccuVote-TS	touch screen	5.3	Y
Sequoia Edge	touch screen	6.1	Y
Datavote	punch card	5.9	Y
Pollstar	punch card	9.0	N
Votomatic	punch card	30.3	N

Notes: Prevalence is defined as the percentage of observations in which each brand of technology is used. When calculating prevalence, observations are weighted by the total number of votes cast in each precinct. Brands listed with a "Yes" in the last column were used again by at least one county in the statewide elections in 2005.

Appendix C: Calculation of the Number of Misvotes

Results from the regression, $Voteshare_{pdc} = \beta_0 + \boldsymbol{\beta}_1 \times \mathbf{I}_{dc}^{adjacent} + controls + \varepsilon_{pdc}$, can be used to estimate the total number of misvotes cast by voters in the recall election. Let $\mathbf{I}_{dc}^{adjacent}$ represent a vector of adjacency dummy variables:

$$\left(I^{north\ AS\ adjacent}, I^{south\ AS\ adjacent}, I^{east\ AS\ adjacent}, I^{west\ AS\ adjacent}, \right.$$
$$I^{north\ CB\ adjacent}, I^{south\ CB\ adjacent}, I^{east\ CB\ adjacent}, I^{west\ CB\ adjacent},$$
$$\left. I^{north\ TM\ adjacent}, I^{south\ TM\ adjacent}, I^{east\ TM\ adjacent}, I^{west\ TM\ adjacent}\right).$$

By separately measuring adjacency misvoting for each major candidate and adjacency type, we identify the unique gain in vote share when, for example, a candidate is north adjacent to Schwarzenegger. This distinction is necessary because not all types of adjacency exist with equal frequency. Let $\boldsymbol{F}$ be a vector equal to

$$\left(F^{north\ AS\ adjacent}, F^{south\ AS\ adjacent}, F^{east\ AS\ adjacent}, F^{west\ AS\ adjacent}, \right.$$
$$F^{north\ CB\ adjacent}, F^{south\ CB\ adjacent}, F^{east\ CB\ adjacent}, F^{west\ CB\ adjacent},$$
$$\left. F^{north\ TM\ adjacent}, F^{south\ TM\ adjacent}, F^{east\ TM\ adjacent}, F^{west\ TM\ adjacent}\right).$$

For example, $F^{north\ AS\ adjacent}$ represents the average number of minor candidates that are located directly north of Schwarzenegger, weighted by the total votes cast in a precinct. In a given precinct, Schwarzenegger either has no north adjacent candidates (if Schwarzenegger is at the top of a ballot column) or one north adjacent (if Schwarzenegger is not at the top of a ballot column). Since $\boldsymbol{F}$ is an average across all precincts, each component of $\boldsymbol{F}$ ranges between zero and one. One can think of $\boldsymbol{F}$ as correcting for the fact that adjacent candidates exist with unequal frequency. See Appendix B for detailed frequency statistics.

If $\boldsymbol{\beta}_1$ is the coefficient vector derived from the regression specified above, misvotes as a percentage of total votes cast is equal to $\boldsymbol{\beta}_1 \boldsymbol{F}$. A simple calculation reveals that $\boldsymbol{\beta}_1 \boldsymbol{F}$ equals 0.30 with a standard error of 0.04. Thus, misvotes accounted for 0.30 percent of all votes cast in the 2003 California recall election.

REFERENCES

Achen, Christopher H., and W. Phillips Shively. 1995. *Cross-Level Inference*. Chicago: University of Chicago Press.

Agarwal, Sumit, John C. Driscoll, Xavier Gabaix, and David Laibson. 2007. "The Age of Reason: Financial Decisions over the Lifecycle." National Bureau of Economic Research Working Paper 13191.

Alvarez, R. Michael, Melanie Goodrich, Thad E. Hall, D. Roderick Kiewiet, and Sarah M. Sled. 2004. "The Complexity of the California Recall Election." *PS: Political Science and Politics,* 37(1): 23–26.

Alvarez, Michael R., Stephen Ansolabehere, Erik Antonsson, Jehoshua Bruck, Stephen Graves, Thomas Palfrey, Ron Rivest, et al. 2001. *Voting: What is, What Could Be*. Caltech/MIT Voting Technology Project. http://www.vote.caltech.edu/2001report.htm. (Accessed March 23, 2008).

Benjamin, Daniel J., Sebastian A. Brown, and Jesse M. Shapiro. 2006. "Who is 'Behavioral'? Cognitive Ability and Anomalous Preferences." Unpublished.

Brady, Henry E., Justin Buchler, Matt Jarvis, and John McNulty. 2001. *Counting All the Votes: The Performance of Voting Technology in the United States*. Department of Political Science, Survey Research Center, and Institute of Governmental Studies, University of California, Berkeley. http://ucdata.berkeley.edu:7101/new_web/countingallthevotes.pdf.

Card, David, and Enrico Moretti. 2007. "Does Voting Technology Affect Election Outcomes? Touch-Screen Voting and the 2004 Presidential Election." *Review of Economics and Statistics*, 89(4): 660–73.

Choi, James J., David Laibson, Brigitte C. Madrian, and Andrew Metrick. 2004. "For Better or For Worse: Default Effects and 401(k) Savings Behavior" In *Perspectives in the Economics of Aging*, ed. David A. Wise, 81–121. Chicago, IL: University of Chicago Press.

Conlisk, John. 1996. "Why Bounded Rationality?" *Journal of Economic Literature*, 34(2): 669–700.

Dee, Thomas S. 2007. "Technology and Voter Intent: Evidence from the California Recall Election." *Review of Economics and Statistics*, 89(4): 674–83.

DellaVigna, Stefano. 2007. "Psychology and Economics: Evidence from the Field." National Bureau of Economic Research Working Paper 13420.

Garner, Phillip, and Enrico Spolaore. 2005. "Why Chads? Determinants of Voting Equipment Use in the United States." *Public Choice*, 123(3–4): 363–92.

Ho, Daniel E., and Kosuke Imai. 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.

King, Amy, and Andrew Leigh. Forthcoming. "Are Ballot Order Effects Heterogeneous?" *Social Science Quarterly.*

Knack, Stephen, and Martha Kropf. 2002. "Who Uses Inferior Voting Technology?" *PS: Political Science and Politics*, 35(3): 541–48.

Knack, Stephen, and Martha Kropf. 2003. "Voided Ballots in the 1996 Presidential Election: A County-Level Analysis." *Journal of Politics*, 65(3): 881–97.

Koppell, Jonathan G.S., and Jennifer A. Steen. 2004. "The Effects of Ballot Position on Election Outcomes." *Journal of Politics*, 66(1): 267–81.

List, John A., and David Lucking-Reiley. 2002. "Bidding Behavior and Decision Costs in Field Experiments." *Economic Inquiry*, 40(4): 611–19.

Lusardi, Annamaria, and Olivia S. Mitchell. 2007. "Financial Literacy and Retirement Preparedness: Evidence and Implications for Financial Education." *Business Economics*, 42(1): 35–44.

Madrian, Brigitte C., and Dennis F. Shea. 2001. "The Power of Suggestion: Inertia in 401(K) Participation and Savings Behavior." *Quarterly Journal of Economics*, 116(4): 1149–87.

Miller, Joanne M., and Jon A. Krosnick. 1998. "The Impact of Candidate Name Order on Election Outcomes." *The Public Opinion Quarterly*, 62(3): 291–330.

Rashes, Michael S. 2001. "Massively Confused Investors Making Conspicuously Ignorant Choices (MCI-MCIC)." *Journal of Finance*, 56(5): 1911–27.

Roth, Susan K. 1998. "Disenfranchised by Design: Voting Systems and the Election Process." *Information Design Journal*, 9(1): 1–8.

Shue, Kelly, and Erzo F. P. Luttmer. 2006. "Who Misvotes? The Effect of Differential Cognition Costs on Election Outcomes." National Bureau of Economic Research Working Paper 12709.

Simon, Herbert A. 1955. "A Behavioral Model of Rational Choice." *Quarterly Journal of Economics*, 69(1): 99–118.

Smith, Richard L. 2002. "A Statistical Assessment of Buchanan's Vote in Palm Beach County." *Statistical Science*, 17(4): 441–57.

Taebel, Delbert A. 1975. "The Effect of Ballot Position on Electoral Success." *American Journal of Political Science*, 19(3): 519–26.

Thaler, Richard H., and Shlomo Benartzi. 2004. "Save More Tomorrow: Using Behavioral Economics to Increase Employee Saving." *Journal of Political Economy*, 112(1): S164–87.

Tomz, Michael, and Robert P. Van Houweling. 2003. "How Does Voting Equipment Affect the Racial Gap in Voided Ballots?" *American Journal of Political Science*, 47(1): 46–60.

Wand, Jonathan W., Kenneth W. Shotts, Jasjeet S. Sekhon, Walter R. Mebane, Jr., Michael C. Herron, and Henry E. Brady. 2001. "The Butterfly Did It: The Aberrant Vote for Buchanan in Palm Beach." *American Political Science Review*, 95(4): 793–810.