Voting in Low Information Elections:* Bundling and Non-Independence of Voter Choice

Comments Welcome, version 1.0

Elisabeth R. Gerber
Gerald R. Ford School of Public Policy
University of Michigan

Jeffrey B. Lewis
Department of Political Science
University of California, Los Angeles

April 23, 2002

Abstract

Decades of research on voting in American elections have consistently identified political parties as a predominant determinant of vote choice. Parties play a valuable role in helping voters overcome a lack of substantive information by providing them with low-information cues and shortcuts. However, in many elections, party cues are either absent (as in ballot measure elections and non-partisan races) or not informative (as in primary elections when multiple candidates share a party label). This research considers how voters make electoral choices in these settings when party cues are unavailable. We utilize a unique dataset that allows us to study electoral choice across a set of primary election races. We test a number of hypotheses about the determinants of cross-race voting patterns, and find that voters rely heavily on information shortcuts similar to those provided by party cues. As such, the research not only helps us understand how voters navigate the low-information electoral environments they often face, but also allows us to isolate and examine some of the roles played by parties in other elections.

*Prepared for presentation at the Annual Meetings of the Midwest Political Science Association, April 25-28, Chicago, IL. We thank Valenta Kabo for research assistance. Gerber acknowledges the Center for Local, State, and Urban Policy at the University of Michigan.

1 Introduction

Research on voting in American elections has consistently identified party identification as a primary determinant of vote choice. This massive body of research has considered how individuals acquire their party identification;[1] what kinds of information party conveys;[2] and what role party plays relative to other factors in candidate evaluation[3] and vote choice.[4] Contemporary scholars largely concur that by allowing voters to engage in "low-information rationality," parties help regular citizens make "informed" electoral choices despite their low levels of substantive information (Popkin 1991, but see Bartels 1996 for evidence that uninformed voters fail to emulate informed choices).

However, in many elections, party cues are either absent (as in ballot measure elections and non-partisan races) or uninformative (as in primary elections when multiple candidates share a party label). To the extent that voters in these elections seek to make informed decisions, they must rely on other readily available sources of information. Work on presidential primaries shows that citizens rely on candidate qualities (Williams et al. 1976, Norrander 1986, 1993) as well as ideology (Wattier 1983), policy positions and electability (Abramowitz 1989; Stone et al. 1992; Stone et al. 1995). Research on voting in ballot measure races shows that endorsements by interest groups may partially substitute for party endorsements by allowing voters to link their policy preferences to those of the endorsers (see Lupia 1994, Bowler and Donovan 1998 for the role of interest group endorsements in ballot measure elections); work on municipal elections reveals a similar dynamic with newspaper endorsements (see Lieske 1989, Krebs 1998).

In this paper, we study voting in one type of low-information elections where party cues are uninformative: primary elections. Research on voting in primary election provides valuable insight into the dynamics of electoral choice in at least two ways. First, it improves

[1]For contemporary work on political socialization and the acquisition of partisanship, see Beck and Jennings 1991, Cain et al. 1991, Sears and Valentino 1997, Trevor 1999.

[2]Two major strands in this body of literature focus on the role of parties in transmitting policy (see Fiorina 1981, Gerber and Green 1998) and non-policy (Alwin and Krosnick 1991, Rahn 1993, Whiteley 1988) information to voters.

[3]See Kahn and Kenney 1997, Dalton et al. 1998.

[4]See Fiorina 1981, Marra and Ostrom 1989, Brown and Woods 1991, Shanks and Miller 1991.

our understanding of how voters make choices in the low-information electoral settings they commonly face. Most of the electoral contests that US citizens decide lack the relatively rich information available in Presidential, US Senate, or even congressional elections. On a typical ballot, voters are asked to select candidates for state, county, and municipal legislative offices, statewide executive branch offices, judicial offices, school board offices, and party leadership positions. In some states, they decide statewide ballot measures, local ballot measures, bond authorizations, and tax levies. In these down-ballot races, information about candidates and issues is sparse, difficult to access, or entirely non-existent. Most of our theories of voting and information have been developed in the context of high profile elections where information is at least available; research on low-information environments will allow us to extend and/or adapt our theories of voting to account for a wider range of electoral activities.

Second, research on voting in low-information elections may provide analytic leverage to help us better understand some aspects of voting in more information-rich environments. Perhaps most importantly, and somewhat ironically, it allows us to better understand the role of party by unpacking and isolating some of the many roles party plays. Party labels are believed to help voters infer candidate ideology, beliefs, and likely future policy positions; shared interests with organized interest groups or groups of citizens; and personal characteristics (see Popkin 1991 for a summary of this research). Unfortunately, it is extremely difficult to infer from an individual voter's behavior which of these many considerations lead him or her to cast a particular vote. This is because when there are few candidates/electoral options, these many considerations often line up in such a way that they would all lead to the same vote choice (i.e., their many effects are observationally equivalent). For example, a person who uses party to infer the match between her own ideology and those of several candidates would often make the same choice as if she used party to infer the match between specific issue positions, etc. In some low information elections (such as the ones we study below), other organizations besides parties play some but not all of these roles, thereby allowing the researcher to isolate these functions and assess their marginal importance. In particular, we are able to study the effect of candidate "bundling," where a political party (or other organization) puts forward and commonly campaigns for a slate of candidates. This bundling function may be particularly important in a federal system where

citizens elect representatives across levels of government and parties simplify this process by fielding candidates under a common label at all (or some) levels. Our dataset allows us to test hypotheses about the importance of this bundling function by looking at how other organizations besides parties bundle candidates and how voters respond to this bundling.

We take advantage of a unique dataset that reports the choices of individual voters across a number of primary election races. Studying voter choice across primary elections allows us to test different sorts of hypotheses. Rather than studying the determinants of a single vote for one of a few candidates, our dataset allows us to think about and analyze patterns of votes. In other words, we can test hypotheses about how voters cast a set of votes (i.e., are votes independent or correlated across races? If they are independent, what factors motivate individual choices? If they are correlated, what is the common basis for choice across races? Ascriptive characteristics? Ideology? Campaign spending?).

In the next section, we offer a number of hypotheses about voting patterns that derive from different views of voter choice. We then describe the data and methods that allow us to test these hypotheses. Next, we report the results of our empirical analyses, and conclude with a discussion of implications and future research.

2 Hypotheses

Theories of voter choice, designed to understand the relative importance of multiple factors in shaping a voter's decision to support or not support a given candidate, abound. Rather than summarize that voluminous literature here, we simply refer the reader to the introductory section of this paper. Instead, we shift our focus in this paper to thinking about what some of these theories imply for the patterns of votes we would expect to observe citizens casting across a set of races. In particular, we focus on two such theories in this paper.

2.1 Spatial Voting

One of the dominant views of individual choice is the spatial voting model (Riker and Ordeshook 1968, Enelow and Hinich 1984). In the model's starkest form, well-informed voters compare their own policy preferences to the positions of two or more candidates and select

the candidate whose position is closest to their own. The story becomes complicated by uncertainty, multidimensionality, considerations of candidate viability, and the like, but in the end, the basic story remains one of choice based on policy or preference proximity.

How does the spatial voting model extend to understanding voter choice across a set of races? To the extent that voters can identify which candidate in each race most closely reflects his or her own policy preferences, then under the spatial voting model, we would expect conservative voters to consistently vote for conservative candidates in each race, for moderate voters to vote for all moderate candidates, and for liberal voters to vote for liberal candidates. Of course, not all primary election races field a full ideological range of candidates; many are characterized by no real competition and often field only a single (often incumbent) candidate. Still, we expect that when a choice does exist, voters will select the candidate ideologically closest to them. And to some extent, we recognize that the spatial voting model, in its starkest form, is a bit of a straw-man, since we also expect many other factors to impact voter choices for individual candidates including gender, race, incumbency, quality, visibility, endorsements, etc. Still, we interpret existing theoretical and empirical research on voting as indicating a strong and important role for ideology, and so put the simple spatial voting model forward as a baseline.

2.2 Bundling

We compare the simple spatial model with an alternative view of voting that we believe may be extremely important in the elections we study, and possibly in other elections as well. Specifically, we consider how voter behavior would look if voters took cues from groups or organizations that put forward and campaign for a bundle or slate of candidates. Naturally, part of the answer to this question depends on how the bundling organizations behave. Based on observations of many campaigns, we argue that any of the following factors may motivate organizations to advance a slate of candidates:

Ideology: one type of bundling occurs when ideologically motivated organizations endorse (and bundle together) candidates who share the organization's ideological perspective. These ideological organizations may include organized interest groups, unorganized coalitions of

interests, industry representatives, legislative caucuses, etc. Advancing and campaigning for a slate of ideologically-minded candidates may be an effective way for an interest group to make the most of its scarce campaign resources, and being associated with an ideologically identifiable group can help a candidate define and publicize his or her ideological positions. To the extent that organizations bundle candidates based on ideological similarity, and that voters who share the organizations' ideology respond to these cues, then we expect voter behavior to closely resemble the behavior predicted by the spatial voting model.

Issues: other interest groups define themselves more narrowly than the broad-based ideological groups described above. To the extent that they seek to leverage campaign resources and information, they may bundle candidates based on a shared position on a single issue. Examples of groups who engage in single issue bundling might include, for example, gun control organizations, gun rights organizations, environmental organizations, etc. On these slates, we would expect to observe some ideological differences in the endorsed candidates (i.e., thus, an environmental organization may endorse candidates who share a similar position on environmental issues, but who differ on other (say economic or social) issues). To the extent that voters respond to the cues provided by single-issue slates, we expect to observe some deviations from the behavior predicted by the spatial voting model, with a clustering of votes one a single issue dimension.

Candidate Groups: slating or bundling of candidates is often done by professional campaign consultants, who distribute candidate "slates" to registered voters. In recent years, targeting of these slates has become a highly sophisticated, multi-million dollar industry in many places. The consultants who coordinate and implement these slate mailing campaigns work for individual candidates or groups of candidates. One common strategy is for a few well-known candidates to forge a personality-based coalition (often across levels and branches of government), and to subsidize the campaign costs of some of their less well-endowed colleagues in return for future legislative support. These candidate groups are typically (but not always) comprised of members of the same party. They may or may not share an ideology. Voter responding to the cues provided by candidate-based slates will cast votes that reflect the slate's endorsements, but that diverge from the predictions of the spatial voting model.

Gender/Racial/Ethnic: bundling or slating candidates may also be used to identify can-

didates who share the same gender or similar racial or ethnic backgrounds. When voters respond to these endorsements, they vote for candidates based on ascriptive, as opposed to ideological or issue-based, characteristics.

Financial: finally, some organizations may construct bundles or slates of candidates simply on the basis of ability to pay. Vote patterns based on these endorsements may show little or no resemblance to the predictions of the spatial voting model.

3 Data and Methods

We use a unique dataset that allows us to study electoral choice across a set of primary election races. The dataset consists of the electronic images of each of the 927,024 ballots cast in the 1992 Democratic Party primary in Los Angeles County, California.[5] Each ballot contains votes on up to 11 nomination races[6] plus three statewide, two countywide, and as many as five local ballot measure races. Given the multiple overlapping political jurisdictions in the county, there are a total of 189 unique ballot combinations (referred to as "ballot groups") which include a common set of candidates. Ballot groups are identified on the individual ballots, as is the precinct in which the ballot was cast.

3.1 Slates

Given our interest in slate mailers, and their role in bundling candidates and coordinating voters, we next consider how those slates are organized and how they relate to the ballots themselves. Each organization that produces a slate mailer is required to disclose two pieces of information: dollars spent in various reporting periods, and the amount of money received

[5]Given California's closed primary system, each ballot contains the nomination races for only a single party. In this paper, we restrict our analysis to Democratic Party ballots only, for several reasons. First, limiting our analysis to a single party greatly reduces its technical complexity, although it still requires us to program the analysis of 189 unique ballot groups (see below). Second, given our interest in testing the bundling function of slate mailer organizations, we find that in Los Angeles in 1992, the most important such organizations worked for Democratic Party candidates. At a future date, we hope to expand our analysis to add Republican Party and perhaps minor party ballots.

[6]Each ballot contains nomination races for President, US Senate (2 year term and 6 year term), US House of Representatives, State Assembly, and District Attorney. Some also contain races for State Senate, County Supervisor, Superior Court Justice, Municipal Court Justice, and City Council.

Los Angeles County Slate Mailer Organizations, January–June 1992

Organizations	Statewide Spending Jan-June 1992
A.N.G.L.E. slate	31,005
California Democratic Alliance	135,688
California Democratic Voter Checklist	349,623
California Black Caucus Voter Guide	85,454
CARAL PAC Official Endorsement Guide	70,000
Citizens for Republican Values	124,776
Democratic Party of San Fernando Valley Slate Mailer Org	27,912
Independent Californians for Affordable Healthcare	100,000
Independent Voters League of California	17,616
LA Vote	111,392
Non-partisan Candidate Evaluation Council, Inc.	350,546
Republic Communications	53,829
Team for the 1990s	254,709
United Democratic Campaign	71,468
Unity Slate	66,850
Voter Education Project	222,992
Voter Guide	1,795,376
Your Ballot Guide	252,1778
Voter's Guide to Pro-Choice Republicans	55,793
Your Pro-Choice Voter Guide	132,856
Total all groups	4,310,062

Table 1: *Table reports statewide spending for the reporting periods Jan-May and May-June 1992 for all slate mailer organizations that reported activity during that period in Los Angeles County.*

from each candidate listed on its slate(s).[7] Table 1 lists each of the slate mailer organizations that reported activity in LA County between January and June of 1992.

For the purposes of this paper, we focus on one particular organization: "Voter Guide." As Table 1 reveals, this is the county slate mailer organization that spent the most money ($1,795,376) during the 1992 primary season. This slate was organized by "B.A.D. Campaigns, Inc.," the firm of Michael Berman and Carl D'Agostino, who specialized in running California (especially West Los Angeles) political campaigns in the 1980s and 1990s. Located

[7]These reporting requirements result from the state's Political Reform Act of 1974.

"Voter Guide" Endorsements, 1992 Democratic Primary Election, Los Angeles County

Office	Races (Total/Contested)	On BAD slate (Total/Contested)	Average payment per voter	Average payment % expend.
US Senate	2/2	2/2	0.31	17
District Attorney	1/1	1/1	0.04	na
US Congress	17/11	10/5	0.10	4
State Assembly	25/15	22/13	0.10	10
State Senate	8/4	5/3	0.05	2
County Supervisor	3/3	2/2	0.10	na
Superior Court	3/1	1/1	0.09	39
Municipal Court	5/5	2/2	0.04	na
Statewide Proposition	3/3	2/2	0.00	na
Local Propositions	10/10	6/6	0.08	na
Total	77/55	53/37		

Table 2: *Table summarizes races of various types on the LA County 1992 Democratic Primary Ballot and BAD slate mailer endorsements in those races. "Contested" races are those where more than a single candidate was listed on the ballot. Average payment "per voter" and as a percent of expenditure show the average amount paid by endorsed candidates in contested races per registered Democrat and the average fraction of all campaign expenditures used to buy a position on the BAD slate in contested races.*

in Beverly Hills, BAD Campaigns was widely viewed as one of the most powerful West LA political machines. Table 2 reports the Voter Guide endorsements.

As table 2 shows, the BAD slate endorsed a total of 45 candidates (plus 8 ballot measures), about half of whom ran in contested races. Twenty-two of the endorsed candidates paid for their position on the slate, in amounts ranging from $1,500 (James Blatt, State Assembly and Brian Finander, State Senate) to $636,000 (Mel Levine, US Senate). Most of the candidates who where supported by the Voter Guide, and who did not pay, were either running unopposed and/or were popular and well-known incumbents such as members of Congress Maxine Waters and Henry Waxman.[8] Among those who did pay for their posi-

[8]Because we only collected data on payments made to the Voter Guide group between January 1 and the primary election, it may be the case that some the candidates who appear to have been unpaid actually paid for their positions after the election or before January 1. Causal inspection of additional filings reveals

tions on the slate, the average cost per voter in their district was generally between 5 and 10 cents. The very large cost per voter reported for US Senate candidates probably reflects the fact that these candidates were paying for Voter Guide mailers that were sent outside of LA county (in other words, the effective denominator should be the number of Democratic voters statewide, not just within the county). Many candidates dedicated a large fraction of their campaign resources to buying a position on the Voter Guide slate. For example, the $500,000 that Gray Davis paid was nearly one-quarter of the $2.08 million he spent overall. Similarly, the $4,200 paid by Superior Court candidate Bob Henry was nearly 40 percent of his total spending of $10,801. At the other end of the spectrum, State Senate candidate Teresa Hughes spent only about 1 percent of her total campaign funds on the Voter Guide mailer.

Despite the large resource commitments, however, a place on the Voter Guide slate was no guarantee of electoral success. Overall, 19 of the 29 endorsed candidates running in contested races won. However, many of the candidates that paid the most to be included on slate lost on election day, including "top-of-the-ballot" U.S. Senate candidates Mel Levine and Gray Davis. A number of "down-ballot" endorsed candidates, such as Superior Court candidate Bob Henry, lost as well. Still, we note that with many races fielding multiple candidates, the 65 percent victory rate for Voter Guide endorse candidates is much better than chance would dictate.

Given how much campaigns pay to be included on the slate, it seems reasonable to suppose that candidates believe that the Voter Guide slate mailer has influence. Many dedicate a considerable share of their campaign resources to the getting on the Voter Guide slate. Indeed, in terms of costs per registered voter, gaining a position on Voter guide slate is similar in cost to sending a piece of direct mail.

3.2 Voting Patterns

We now begin to consider the influence that the Voter Guide had on voting in the 1992 Democratic Primary. In an optimal research design, we would randomly assign some voters

no evidence of such out-of-period payments.

to receive the slate mailer and other voters to not receive the mailer. We could then note the difference between support for each endorsed candidate among the two sets of voters and attribute systematic differences to the influence of the slate mail. Unfortunately, we do not know exactly which voters in the county received the "Voter Guide" slate and which did not. However, as described above, we do know the exact pattern of votes cast by each of the nearly one million voters who cast ballots in the Democratic primary. Our approach, then, is to seek to identify the influence of the mailer by looking for evidence that a block of voters followed the mailers recommendations across races. We attempt to identify a subset of the voters that are B.A.D. "straight-ticket voters." Of course, no such voter need exist for the slate to have influence. It could be that being on the slate benefits each candidate, even if it does not change the tendency for voters who support one endorsed candidate to support other endorsed candidates. Voters may simply weigh B.A.D. endorsement as one of many considerations in casting a voting each race. If each voter weighed the endorsement in a similar way, then the slate would have influence, yet we would observe no systematic cross-race correlation in vote choices. It is, however, our claim that the slate does act as a coordinating mechanism tying the fortunes of the endorsed candidates together, much in same way as a party label does in a general election. For the lesser known candidates, this might cut both ways. On one hand, lesser-known candidates gain support from those voters favoring the well-known incumbents and other high-profile candidates on the slate. On other hand, lesser-known candidates lose support among those voters who favor other candidates (not their slate-mates) in high profile races.

Because candidates were endorsed across many offices and districts, the number of endorsements made to each voter varied greatly. Voters living in some areas of LA could support B.A.D. endorsed candidates in as many as 13 contested races; others could support as few as five. Table 3.2 describes the distribution of the number "Voter Guide" endorsements in contested races that were followed as a function of the total number of opportunities to follow those recommendations. It is immediately clear from the table that relatively few voters followed the recommendations across the entire slate. One important caveat here is that this table treats abstentions (i.e., roll-off or under-votes) as failures to follow the slate. Many of the zero matches, for example, arise from voters who only voted for President or a

Distribution of the number of Voter Guide recommendations in Contested Races followed by LA County Voters

	BAD endorsements in contested races on ballot:								
Matches	5	6	7	8	9	10	11	12	13
0	15	10	8	4	4	3	2	2	1
1	28	21	17	9	6	6	4	3	2
2	31	29	25	16	10	9	6	5	4
3	18	23	24	20	14	11	9	7	6
4	6	12	15	21	19	14	11	9	8
5	2	4	7	16	22	16	14	12	10
6		1	3	10	15	17	17	14	13
7			1	4	7	13	17	17	14
8				1	3	8	11	15	15
9					1	3	6	12	12
10						1	2	4	9
11							0	1	4
12								0	1
13									0
Total	99,741	151,875	92,449	155,500	22,507	32,740	221,063	47,366	103,783

few offices and left the remainder of the ballot blank. Hence, our estimates of slate behavior are highly conservative and likely understate the importance of the slate. It should also be noted that many of the elections involve more than two candidates, further lower the chances of observing matches by chance. Thus, while only a small fraction of voters exactly reproduced the endorsed slate on their ballots, the distribution of matches strongly rejects the notion that support for endorsed candidates is independent across races.

4 Model and Estimation

In this section, we describe and estimate two models of voter behavior in the 1992 LA Democratic Primary election. In each of these models, the electorate is assumed to be comprised of two types of voters. In the first model, one type of voter closely follows the Voter Guide recommendations while the other type votes independently across races. In the second model, one type again follows the Voter Guide recommendations while the other type follows a simple one-dimensional spatial model. These mixture models allow us to address two

central problems of inference that arise in estimating the effect of bundling/ endorsements. First, inspection of Table 3.2 reveals that few voters supported every candidate endorsed by the Voter Guide. Yet, it seems overly restrictive to suppose that a voter who "follows" the Voter Guide must support the Guide's endorsed candidate in every race. Race-specific idiosyncratic factors and errors may also play a role. Our model does not require that slate voters support every endorsed candidate. Thus avoiding the downward bias on the estimated influence of the mailer that would result from a method that simply counted the number of ballots that were perfectly congruent with the Voter Guide. Second, our method controls for the possibility that voting patterns that are congruent with the Voter Guide simply because the Guide endorses either popular or ideologically similar candidates. By building plausible null models of primary voting into the estimation, we avoid the attribution of influence of the slate mailer to patterns of voting that would commonly arise even if no voters followed the slate mailer. Thus, the modeling strategy avoids the underestimation of slate mailer influence that would result from requiring voter agreement with every recommendation, and the overestimation of influence that would result from attributing to the effect of the slate mailer every vote that was consistent with its endorsements.

The two models are estimated by maximum likelihood. The likelihood functions are each constructed out of three conditional components. The first component captures the likelihood of casting a particular pattern of votes, given that a voter is of the slate-following type. The second component captures the probability of casting a particular vote if votes are cast independently across races. The third component captures the probability of casting a particular pattern of votes if voting is governed by a simple one-dimensional spatial voting model similar to Poole & Rosenthal's NOMINATE (1985). The components are then combined in mixture likelihoods and estimated using a simple EM algorithm.

4.1 The slate following likelihood

Consider a vector of votes $\mathbf{V} = (v_1, v_2, \ldots, v_K)$ cast by a voter and a vector of slate recommendations $\mathbf{S} = (s_1, s_2, \ldots, s_K)$ across K offices. If $v_k = s_k$, then the voter followed the recommendation of the slate in race k; otherwise, she did not. If the voter is of the

slate-following type, we assume that she follows each slate recommendation with probability λ; $\text{Prob}(v_k = s_k) = \lambda$ for $k = 1, 2, \ldots, K$. When a voter deviates from the slate recommendation, we assume (for the moment) that she chooses among the remaining candidates at random. Thus, we define the probability that $v_k = v'_k$ as $(1-\lambda)/(n_k - 1)$ for all $v'_k \neq s_k$ where n_k is the number of candidates contesting race k. The likelihood of the full vector of votes is

$$L_S(\mathbf{V}_i|\mathbf{S}) = \prod_{v_{ik}=s_k} \lambda \prod_{v_{ik}\neq s_k} (1-\lambda)/(n_k - 1).$$

The maximum likelihood estimate of λ is simply the fraction of times that $v_{ik} = s_k$ across all voters $i = 1, 2, \ldots, N$ and all K races. Thus, if we knew which voters were slate-following, the probability of following each slate recommendation could be easily estimated. The ML estimate of λ is simply the fraction of times that the slate mailer recommendation is followed across all races and voters.

4.2 Independent voting across races

A second type of voter is assumed to vote in each race on an idiosyncratic basis. Electoral outcomes among this population are driven by factors such as advertising, name recognition, door-to-door canvassing, and any other factors that are not coordinated across candidates running for different offices (such as slate endorsements). We make no attempt to model which candidates will be popular and which will not. We simply assume that voting is independent across races. That is, information about which candidate a voter selected in one race is assumed to reveal no information about a voter's choice in another. For each of the K races, we estimate the parameters of a multinomial probability distribution function P_k over the votes for each of the n_k candidates where $P_k(v)$ yields the probability of voting for candidate $v = 1, \ldots, n_k$ in race k. The likelihood of a given vector of votes, conditional on belonging to this population, is

$$L_I(V_i, P_1, \ldots, P_K) = \prod_k P_k(v_{ik}).$$

Because voting is assumed to be independent across races (and voters), maximizing this

likelihood is also straightforward. The parameters of each of the K multinomial distributions are estimated by the frequencies with which each candidate is supported in each race. For example, the probability of an individual voting for candidate v in race k is estimated by the fraction of voters who supported candidate v in race k.

4.3 A simple spatial model of voting

A great many analyses of U.S. politics — both scholarly and popular — assume a simple left-right ideological continuum. Indeed, empirical work has found that a single policy dimension describes much of voting that takes place in U.S. Congress (Poole & Rosenthal 1997) and the California Legislature (Lewis 1998, Gerber & Lewis 2001). While much of this seeming unidimensionality may result from the existence of a two party system (Hinich & Munger 1994), it is still common for journalists and pundits to describe primary contests as battles between the centrists and extremists of a party [XXX ADD CAL JOURNAL QUOTES]. Indeed, primary election outcomes in California are frequently described in terms of the ideological positions taken by the candidates. To say that electoral outcomes can be understood in terms of candidates' positions on a single ideological continuum suggests that the primary electorate has preferences over the dimension, that candidates stake out positions on that continuum, and that voters observe (at least partially) those positions. Taken together, these assertions further imply that votes for candidates across races will be correlated, because voters that support centrist candidates in one race, for example, are likely to support centrist candidates in other races as well. Thus, if voters observe the ideological positions of candidates and have preferences over those positions, then we would expect to see dependencies in their choice of candidates across races. Finding such dependencies would be evidence that ideology is an important determinant of vote choice and that voters have at least some information about candidates' positions.

Beyond our intrinsic interest in the existence of ideological voting in primaries, ideological voting (as captured by the standard spatial model) also functions as a useful control in our estimation of the effect of slate mailers. If the Voter Guide slate presented a collection of candidates who shared similar spatial locations, then under spatial voting, we would expect

to see voters casting ballots that are consistent with the slate, even if no voter was directly influenced by the slate's endorsements.

A full description of the spatial model and likelihood function that it implies is given in the appendix. The model follows the basic setup presented in Lewis 2001 and Gerber & Lewis 2001. Each voter is assumed to have an ideal point θ along the single ideological dimension. Support for each candidate is estimated by a multinominal logistic regression on those ideal points. The parameters of these multinominal regressions reveal information about both the location of each candidate in the space and the overall valence of the candidate. Unfortunately, these two components cannot be fully separated, so each candidate cannot be fully located in the space, though the left-right order of the candidates for each office can be established [XXX NEED TO PROVE THIS]. Estimation is complicated because the ideal points of the voters θ are not observed. However, by assuming the distribution of voter ideal points within the county, the likelihood of observing each vote pattern can be obtained by marginalizing the logistic regression likelihood with respect to θ;

$$L_M(\mathbf{V}_i|\boldsymbol{\beta}) = \int \prod_k \mathrm{MNL}(v_{ik}|\theta, \boldsymbol{\beta}_k) f(\theta) d\theta. \tag{1}$$

where MNL is the multinomial regression likelihood of the choice v_{ik} given parameters $\boldsymbol{\beta}_k = (\beta_{0k}, \beta_{1k})$, the MNL regression "intercept" and "slope," and f is the probability density function for the ideal points θ. The MNL parameters for each race, $\boldsymbol{\beta}_k$, can be estimated using an EM algorithm first presented in Bock & Aitken (1981).

4.4 Mixture of voter types

The preceding paragraphs posited three voter types and described the likelihood of observing each pattern of votes $\mathbf{V}$ for each of these types. However, we do not know the type of any given voter. The full likelihood for any given vector of votes (unconditional on voter type) is thus a mixture of the likelihoods for each of the types in the population. We consider two populations. As described above, one population is comprised of slate followers and of voters who vote independently across races. The other population is comprised of slate followers and of voters who follow the spatial model. Construction of the two mixture populations is

similar. We present the derivation of the slate–spatial population. While not presented, the derivation of the slate–independent population is parallel.

The likelihood of observing a pattern of votes $\mathbf{V}_i$ in the slate–spatial population is

$$L(\mathbf{V}_i|p, \lambda, \boldsymbol{\beta}) = pL_S(\mathbf{V}|\lambda) + (1-p)L_M(\mathbf{V}_i, \boldsymbol{\beta})$$

where $p \in [0, 1]$ is fraction of the population that is of the slate-following type. While the full likelihood could be maximized directly, it is well known that such finite mixture models are amenable to estimation by a simple EM algorithm (see McLachlan & Peel 2000). In the EM approach, the voter's type is treated as missing data in the log likelihood

$$\ln L(\mathbf{V}_1, \ldots, V_N|\lambda, \boldsymbol{\beta}, t) = \sum_i t_i \ln L_S(\mathbf{V}_i|\lambda) + \sum_i (1-t_i) \ln L_M(\mathbf{V}_i, \boldsymbol{\beta})$$

where $t_i = 1$ indicates that voter i is of the slate-following type. Because the slate and spatial components of the complete data log likelihood are additively separable, L_A and L_M can be maximized independently using the estimators described above. Similarly, if t where observed, the ML estimator p is $\hat{p} = \sum t_i/N$. However, t is not observed. In the EM algorithm, the approximate expectation of t_i as a function of $\mathbf{V}_i$, p, λ, and $\boldsymbol{\beta}$ is calculated using Bayes' rule,

$$\mathrm{E}t_i = \frac{pL_S(\mathbf{V}_i|\lambda)}{pL_S(\mathbf{V}_i|\lambda) + (1-p)L_M(\mathbf{V}_i, \boldsymbol{\beta})}$$

substituting an initial set of guesses about p, λ, and $\boldsymbol{\beta}$. Then the expected value of the likelihood (with respect to t),

$$\mathrm{E_t} \ln \mathrm{L}(\mathbf{V}_1, \ldots, \mathrm{V_N}|\lambda, \boldsymbol{\beta}) = \sum_{\mathrm{i}} \mathrm{Et_i} \ln \mathrm{L_S}(\mathbf{V}_\mathrm{i}|\lambda) + \sum_{\mathrm{i}} (1 - \mathrm{Et_i}) \ln \mathrm{L_M}(\mathbf{V}_\mathrm{i}, \boldsymbol{\beta}),$$

is maximized with respect λ, and $\boldsymbol{\beta}$ and p is estimated as $\sum Et_i/N$. The new parameter values are then used to obtain new estimates of Et_i for $i = 1, 2. \ldots, N$, and these values used to obtain new estimates of λ, and $\boldsymbol{\beta}$ and p, and so forth, until the estimated quantities converge. Under fairly general conditions, the solution arrived at by this algorithm will be the solution to the original maximization problem presented in (XX). However, the

EM calculations are simpler and more stable than those required for direct maximization. Standard errors for the model parameters can be obtained by evaluating the inverse Hessian of the full likelihood or by a parametric or non-parametric bootstrap [XXX STILL NEED TO IMPLEMENT THIS].

4.5 Results

In order to apply the models developed above to the LA County primary voting data, we first must identify subsets of contested races in which the Voter Guide made endorsements, and the set of voters who were eligible to vote in that set of races. We have identified four subsets of races and voters with which to estimate the size of the Voter Guide slate-following population. The first subset identifies the five contested races in which the Voter Guide made endorsements (three candidate races and two propositions) and for which all voters in the county were eligible to vote. The second focus on voters in the Western part of Los Angeles city (West LA) who were eligible to vote for Voter Guide-endorsed alternatives in 13 contested races (8 candidate races and 5 propositions). The third subset considers only the 8 candidate races in West LA. The forth considers only the 5 ballot measures in West LA.

Before describing the results, one more issue needs to be addressed. Abstentions (roll-off or under-votes) are quite common in the data. Approximately 1/3 of all voters in the four subsets of races described above abstained in at least one of the races considered. Yet, we do not have a model that explicitly incorporates abstention. Instead, we assume that abstentions are random with respect to the features of the data that we are modelling. If abstentions are random, then list-wise deletion of cases with one or more abstentions will lead to a reduction in the efficiency with which we estimate the models, but will not result in bias.

The results of the estimations are shown in Table 3. The first thing to note is that the number of voters who exactly follow the slate is generally small and is a function of the number of possible voting patterns. On the other hand, these slate patterns where among the most frequent in the data. For example, the slate pattern was the 3rd most likely out of

Estimating the fraction of voters following the Voter Guide Slate in the 1992 Los Angeles County Democratic Primary

		Slate voting		With Spatial Null			With Indep. Null			
Sample	No. of Races	Percent Match	Pattern Rank	P	λ	Posterior Prob.	P	λ	Posterior Prob	N
County-wide races	5	4.2	3/360	5.2	92	83	4.2	95	79	677,015
West LA, All	13	0.3	14/10,461	1.2	85	78	2.9	85	99	23,517
West LA, Candidate races	8	0.7	12/3,558	1.5	85	64	3.4	86	97	27,107
West LA, Propositions	5	24.3	1/32	7.2	86	13	17.6	97	64	27,148

Table 3: Estimating the fraction of voters who follow the voter guide. *Table reports estimates of number of voters in selected samples that "follow" voter guide recommendations. As described in the text, the voter population is split in two parts. P percent follow the slate recommendations voting for each slate candidate with probability λ. The remaining $(100 - P)$ percent either follow a one dimensional spatial voting model or cast votes that are independent across races. The "posterior probability" columns give probability that a voter casting the voter guide's recommended slate is a member of the population who follows the recommendations as a voting strategy.*

360 observed patterns in the 5 countywide races, the 14th most likely among 10,461 observed patterns across the 13 West LA races, and the most likely among 32 observed patterns on the West LA ballot proposition races.

Estimates of the fraction of the electorate that is slate-following range from 1.2 to 7.2 percent when the slate-following population is mixed with a spatial voting population, and from 2.9 to 17.6 percent when the slate-following population is mixed with an independent voting population. Our model constrains the probability that each slate recommendation is followed by a slate following voter, λ, to be at least 85 percent. This constraint is imposed to insure the underlying consistency of the notion of a "slate-follower." The cut-off of 85 percent is arbitrary, but it seems reasonable to require that slate voters support slate candidates with at least an 85 percent probability. In several of the specifications, this lower bound is hit and the constraint binds. The fact that this constraint binds suggests the presence of other types in the data. Perhaps there are some voters who simply flip coins. We note that there are numerous other slate mailers and newspaper endorsements which may also have adherents in the population. Future analyses will incorporate some of these additional types. [XXX STANDARD ERRORS FOR THE ESTIMATED PARAMETERS ARE ABSENT AND NEED TO BE PROVIDED]

Next, we consider the extent to which the slate patterns are consistent with the independent or spatial voting models. The "posterior probability" columns in table 3 gives the estimated probability that a voter who cast a ballot that exactly matches the slate recommendations is drawn from the slate follower population. This quantity is calculated in the E-step of the EM algorithm defined above. Notice that in some cases this probability is very high, indicating that the slate pattern was very unlikely to have been generated by the other voting models. By contrast, one case in which the slate recommendations closely resemble the spatial model predictions is the West LA ballot propositions. For this subset of races and voters, the posterior probability of being of slate-follower type, conditional on casting the slate-recommended votes on all five races, is only 13 percent. This is nearly twice the unconditional probability of being a slate-follower type, but is still far less than for the other subsets. These posterior probabilities highlight the potential dangers that would arise if simpler methods of measuring slate mailer influence, based simply on counting the

Fit of Models of Individual Voter Choice, 5 Countywide Races

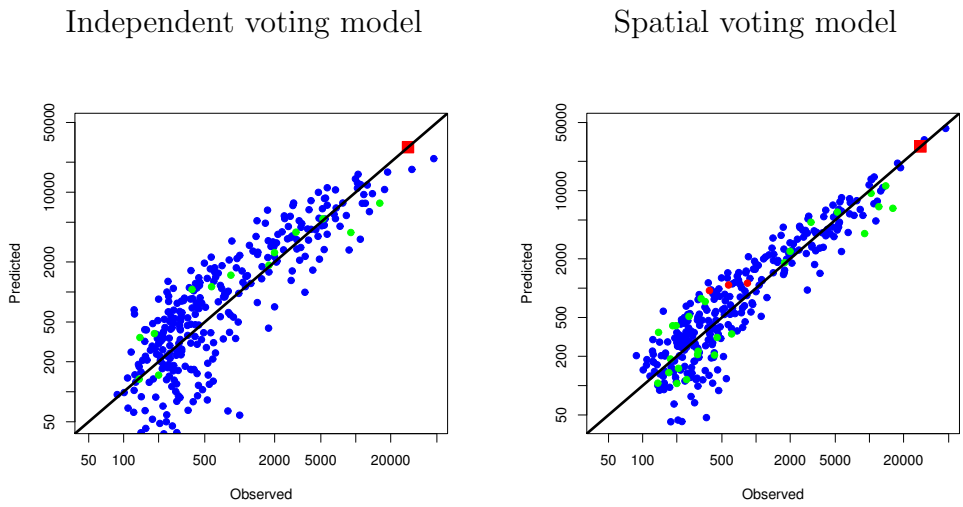

Figure 1: *Each panel plots the observed frequency with which each pattern of votes across the 5 races considered occurs versus the frequency with which those patterns are predicted to occur by each of the two voting models developed in the text. The larger square symbol indicates the voting pattern which matches all of the Voter Guide endorsements.*

frequency with which a slate is followed, are used.

While it is clear that alternatives to the slate voting model must be included in the analysis, it would be useful to know if the models that we have analyzed here seem to describe the patterns of voting that we observe overall. In particular, because the independent voting model is nested within (i.e., is a special case of) the spatial voting model, the increase in fit that is gained when we move from the independent voting model to the spatial model can be used as a measure of the extent to which spatial considerations drove voting in the LA Democratic primary. Figure 1 presents the fit of the two models applied to the countywide races. Each panel plots the observed frequency with which each voting pattern occurs against the number of times it is predicted to occur under the given voting model. The axes scales are logged to allow more detail to be revealed. If the model perfectly fit the data, all of the points would fall along the forty-five degree line through each of the two panels. Because of the log-log scale, deviations above the line are "larger" than equidistant deviations below the line, and deviations in the lower left-hand corners are smaller, in absolute terms, than similar

looking deviations in the upper right-hand corners. Notice that the points fall considerably closer to the forty-five degree line under the spatial model. Of course, this model involves the estimation of twice as many parameters, but even accounting for reduced degrees of freedom, the increase in fit is large and cannot be attributed to chance using standard likelihood ratio tests.

Figure 2 provides information about the size of the effect of spatial considerations on vote choices in the LA county-wide races. Each panel plots the estimated probabilities of supporting each candidate and proposition alternative as a function of the spatial position of a given voter. The solid lines indicate alternatives supported by the Voter Guide. The histogram bars at the bottom of each panel summarize the estimated *a posteriori* distribution of voter positions (ideal points).

The polarity of these graphs has been reversed from the usual left-right meaning (the direction of the underlying scale is arbitrary). Moving from the left to the right of each panel indicates more liberal preferences. Consider for example support for the two statewide Ballot measures. Proposition 152 authorized the state to issue $1.9 billion in education bonds. Proposition 154 would have delayed to reassessment of a home's value after purchase by a low-income family lowering the property tax burden on low-income home buyers. Proposition 152 narrowly passed statewide, while Proposition 154 garnered only 40 percent support. Given the nature of both proposals, we would expect support to be greatest among liberals and least among conservatives. "Yes" votes on each of these proposals were recommended by the Voter Guide. Moving from left to right across the Proposition 152 and 154 panels, support for each proposition increases markedly. Moving from the 25 percentile to the 75 percentile in the distribution of voter locations increases the probability of supporting Proposition 154 from 31 to 82 percent, a difference of 51 percentage points. The same interquartile difference for Proposition 152 is 63 percentage points (from 30 percent to 93 percent).

Similar degrees of spatial voting appear in the two U.S. Senate races. Both the Voter Guide endorsed candidates, Mel Levine for the full term seat and Gray Davis for the two-year seat, were most strongly supported by the moderate/conservative end of the Democratic party. Support for Statewide winners Barbara Boxer and Diane Feinstein was heavily drawn from the liberal end of the Democratic party. Boxer, represented by the steeply rising dashed

Spatial Determinants of Voting for Countywide Candidates in the 1992 Los Angeles County Democratic Primary

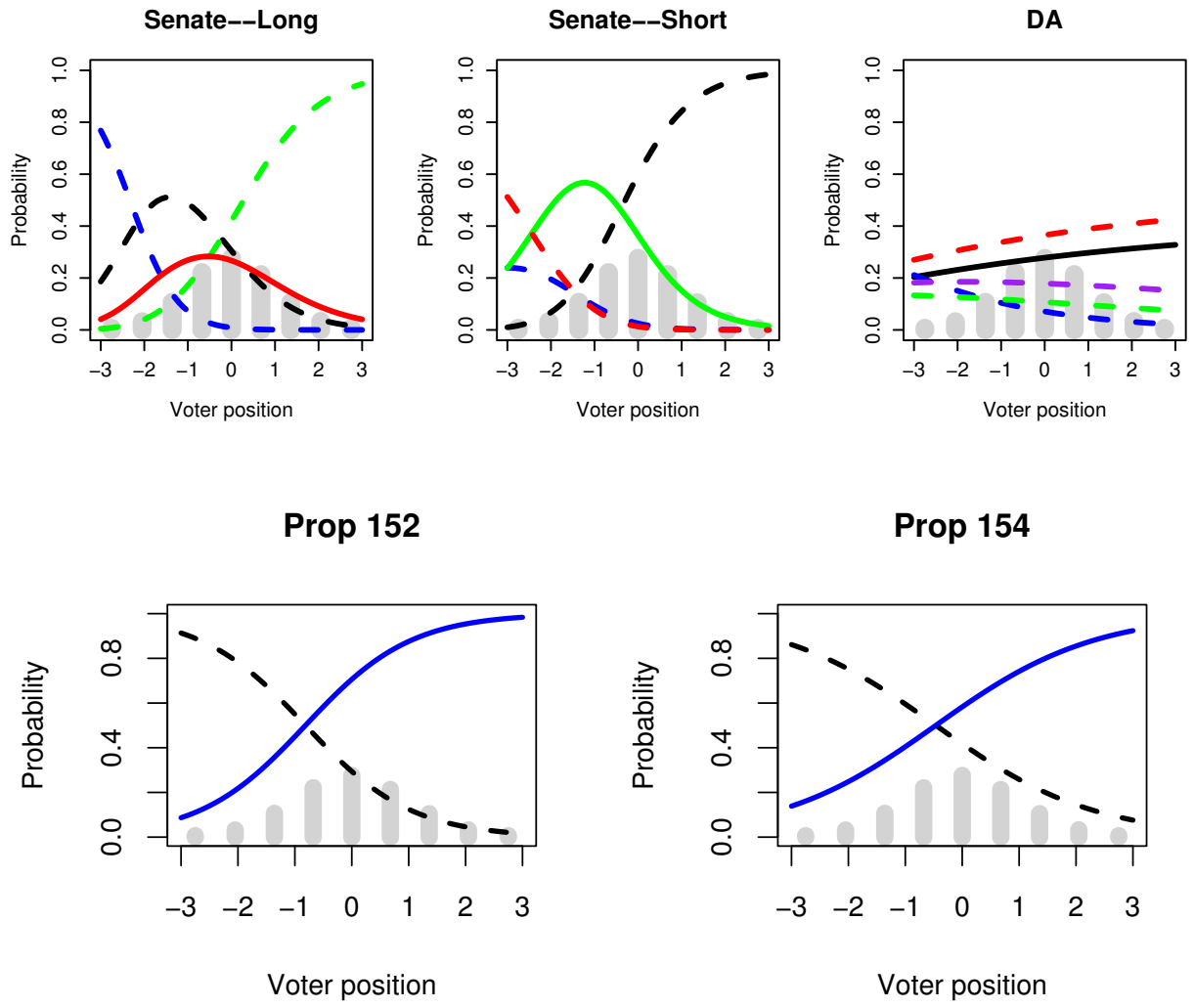

Figure 2: *Graphs show predicted support for each candidate in each race as function of the voter's spatial location. The solid lines in each panel represent the candidate endorsed by the Voter's Guide. The histogram bars show the distribution of spatial locations in the voter population.*

line in the "Senate–Long" panel, was supported by only 8 percent of the voters at the 25th percentile of the ideological dimension, but received 81 percent of the vote among those at the 75th percentile. Feinstein support followed a similar pattern. Her support is represented by the steeply rising dashed line in the "Senate–Short" panel. Feinstein captured 91 percent of among those at the 75th percentile and only 12 percent of the vote of those at 25th percentile.

Support for the District Attorney candidates show much less spatial influence. The Voter Guide endorsed candidate, Ira Reiner, and Gil Garcetti were somewhat more strongly supported by liberals and received greater support from Democrats than they did from Republicans. Two of the three candidates whose support is (somewhat) lower among liberals gained larger support among Republican voters. The third candidate whose support falls as we move from left to right in the DA panel, Howard Johnson (the steeply declining dashed line) drew 7.4 percent of the vote in both the Republican and Democratic primaries.[9]

4.6 Fit of the West LA data

Figure 3 shows the fit of the voting models applied to the West LA races. While the fit is considerably less exact than found in the countywide races, it must be noted the sample size is considerably smaller and the number of possible alternative is much larger in the West LA sample. Thus, the patterns are both harder to predict and we have less data on which to base our predictions. Nevertheless, the presence of large outliers suggests that something (perhaps another set of endorsements) is missing from the model. The spatial voting model again shows considerable advantage over the independent voting model. The spatial model is a particularly good fit for the proposition voting data. [ADD MORE HERE..]

5 Conclusions

Our analysis of voting patterns in the 1992 Los Angeles County Democratic primary election forms the starting point for a broader inquiry into the determinants of voter choice in

[9]District attorney is contested on a non-partisan basis. Candidates are list on all parties' primary ballots. The two highest polling candidates meet in a run off election in the general election.

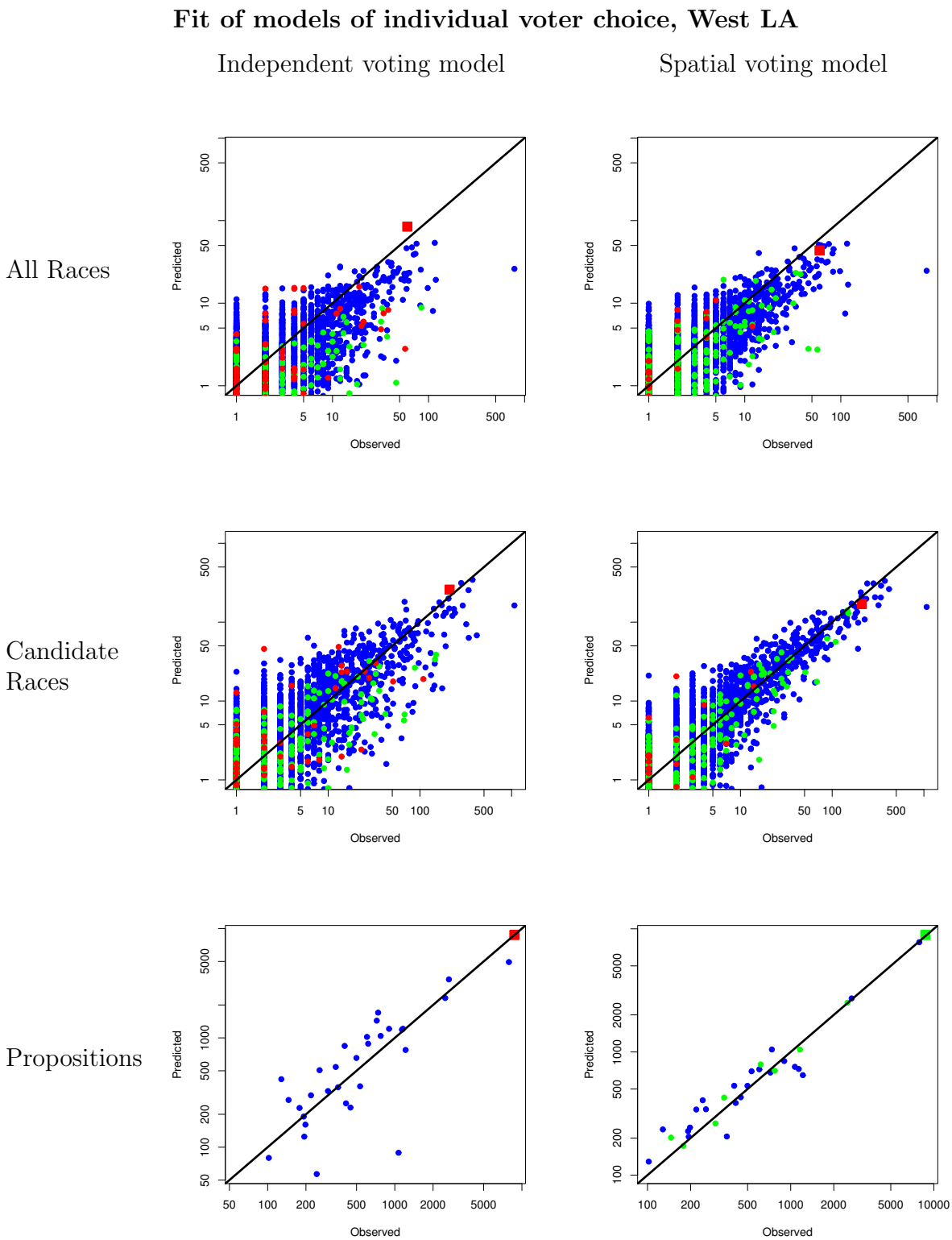

Figure 3: *Each panel plots the observed frequency with which each pattern of voter across the set of races considered occurs versus the frequency with which those patterns are predicted to occur by each of the two voting models developed in the text. The larger square symbol indicates the voting pattern which matches all of the Voter Guide endorsements.*

low information elections. We believe the current analysis contributes to our understanding of elections and electoral behavior in a number of ways. First, we extend theoretical work on voting by developing hypotheses about voting patterns that derive from several common models of electoral behavior. Second, we contribute to our substantive understanding of voting by identifying potentially important sources of voter cues.[10] Third, we make a methodological contribution by constructing and estimating a simple mixture model of voter types that may apply to other choice problems as well. And fourth, we add to our empirical understanding of elections by obtaining initial estimates of the effects of slate mailers on voter choice.

We envision of the following extensions of this work. Most immediately, we plan to incorporate more slates, and hence more voter types, into our mixture models. In the models above, we only allow for the possibility of two voter types - slate followers and "others." To the extent that additional voter types exist in the electorate, then only specifying two types undoubtedly reduces the precision of our estimates, since it forces the model to classify other types of voters into one of the posited types. Subsequently, we hope to also allow for other voter types, such as those who vote for the candidates who spend the most money or who are listed first on the ballot. We are also currently developing an explicit model of abstention.

A Appendix

Suppose an electorate comprised of voters with political preferences over a single policy dimension. Assume each voter i has quadratic preferences over the dimension. Each voter casts a vote in contests $k = 1, 2, \ldots, K$. Each contest involves J_k candidates. Index the candidates in each contest by $j = 1, 2, \ldots, J_k$. Let c_{jk} be the location of candidate j in race k. The utility of candidate j to voter i with ideal point θ_i in race k is

$$U(\theta_i, c) = v_{jk} - (\theta_i - c_{jk})^2 + \epsilon_{ijk}$$

[10] See also recent work by Iyengar, Lowenstein, and Masket 1999.

where v_{jk} reflects non-policy utility provided by candidate j (the value of incumbency for example) and ϵ_{ijk} is an idiosyncratic utility shock which follow a type I extreme value distribution. Voter i selects candidate j in race k if

$$U(\theta_i, c_{jk}) \geq U(\theta_i, c_{j'k}) \text{ for } j' = 1, 2, \ldots, J_k.$$

Note that the continuity of the distribution of idiosyncratic shocks (ϵ) insures that $U(\theta_i, c_{jk}) = U(\theta_i, c_{j'k})$ for $j \neq j'$ is a zero probability event and thus the possibility that a voter is indifferent between two candidates can be ignored in the derivation that follows. As shown by XXX, the assumption of independent type I extreme value shocks, the probability that a voter with ideal point θ support candidate j in race k is

$$\text{Prob}(Y_k = j|\theta) = \frac{\exp(v_{jk} - (c_{jk} - \theta)^2)}{\sum_{j'=1}^{J} \exp(v_{j'k} - (c_{j'k} - \theta)^2)} \text{ for } j = 1, 2, \ldots, J.$$

Rearranging the above we find that

$$\text{Prob}(Y_k = j|\theta) = \frac{\exp(v_{jk} - c_{jk}^2 + 2c_{jk}\theta) exp(-\theta^2)}{\exp(-\theta^2) \sum_{j'=1}^{J} \exp(v_{jk} - c_{jk}^2 + 2c_{jk}\theta)} = \frac{\exp(v_{jk} - c_{jk}^2 + 2c_{jk}\theta)}{\sum_{j'=1}^{J} \exp(v_{jk} - c_{jk}^2 + 2c_{jk}\theta)}.$$

Letting $\alpha_{jk} = v_{jk} - c_{jk}^2$ and $\beta = 2c_{jk}$, we find the familiar multinomial logit model:

$$\text{Prob}(Y_k = j|\theta, \boldsymbol{\alpha}_k, \boldsymbol{\beta}_k) = \frac{\exp(\alpha_{jk} + \beta_{jk}\theta)}{\sum_{j'=1}^{J} \exp(\alpha_{jk} + \beta_{jk}\theta)}.$$

Conditional on θ, votes for candidates across races are independent. Thus, the probability of voting for a particular set of candidates can be written as

$$P(j_1, j_2, \ldots, j_K|\theta, \boldsymbol{\alpha}, \boldsymbol{\beta}) = \text{Prob}(Y_1 = j_1, Y_2 = j_2, \ldots, Y_K = j_K) = \prod_{k=1}^{K} \text{Prob}(Y_k = j|\theta, \boldsymbol{v}_k, \alpha_k, \boldsymbol{\beta}_k).$$

These probabilities could be evaluated if θ were observed. It is not. Treating θ as a random variable, we marginalize the distribution of vote choices with respect to θ so that

$$EP(j_1, j_2, \ldots, j_K|\boldsymbol{\alpha}, \boldsymbol{\beta}) = \int P(j_1, j_2, \ldots, j_K|\theta, \boldsymbol{\alpha}, \boldsymbol{\beta}) f(\theta) d\theta.$$

Assuming that θ follows as predetermined distribution—the standard normal in the application presented here—the parameter matrices $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are estimated by marginal maximum likelihood. Grouping the data into common patterns of votes cast across the k offices, $p = 1, 2, \ldots, J_1 J_2 \ldots J_K$, letting n_p be the number of voters casting vote pattern p, and $p(k)$ represent the candidate in the kth race chosen by a voter casting the pth voter pattern, the log likelihood is

$$L(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \sum_p n_p \ln EP\left(p(1), p(2), \ldots, p(K)|\boldsymbol{\alpha}, \boldsymbol{\beta}\right).$$

This likelihood can be maximized by standard numerical techniques (using Gauss-Hermite quadrature to approximate the normal integral) or by an EM approach. Given estimates of the parameter vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, estimates of the candidate valance and location parameters, c_{jk} and v_{jk} for $j = 1, 2, \ldots, J_k$ and $k = 1, 2, \ldots, K$, can be solved for directly.

In order to identify the model, the valance and position of the first candidate in each race is normalized to zero. That is, $v_{1k} = 0$ and $c_{1k} = 0$ (and consequently $\alpha_{1k} = 0$ and $\beta_{1k} = 0$) for $k = 1, 2, \ldots, K$. This is the standard normalization used in multinomial logit modelling. The candidate valance and locations can then be thought of as relative to the valance and location of the first candidate in each race. It should be noted to that this preclude the possibility of directly comparing the valances and locations of candidates across races. For such comparisons additional identifying restricts are required. For example, the average valance might be assumed to be zero each race.

References

Abramowitz, Alan I. 1989. "Viability, Electability, and Candidate Choice in a Presidential Primary Election: A Test of Competing Models." *The Journal of Politics* 51(4): 977-992.

Alwin, Duane F. and Jon A. Krosnick. 1991. "Aging, Cohorts, and the Stability of Sociopolitical Orientations Over the Life Span." *American Journal of Sociology* 97(1): 169-195.

Bartels, Larry M. 1996. "Uninformed Votes: Information Effects in Presidential Elections." *American Journal of Political Science* 40(1): 194-230.

Beck, Paul Allen and M. Kent Jennings. 1991. "Family Traditions, Political Periods, and the Development of Partisan Orientations." *The Journal of Politics* 53(3): 742-763.

Bock, and Aitken. 1981.

Bowler, Shaun and Todd Donovan. 1998. *Demanding Choices: Opinion, Voting, and Direct Democracy.* Ann Arbor: University of Michigan Press.

Brown, Robert D. and James A. Woods. 1991. "Toward a Model of Congressional Elections." *The Journal of Politics* 53(2): 454-473.

Cain, Bruce E., D. Roderick Kiewiet, and Carole J. Uhlaner. 1991. "The Acquisition of Partisanship by Latinos and Asian Americans." *American Journal of Political Science* 35(2): 390-422.

Dalton, Russell J., Paul A. Beck and Robert Huckfeldt. 1998. "Partisan Cues and the Media: Information Flows in the 1992 Presidential Election." *The American Political Science Review* 92(1): 111-126.

Enelow, James and Melvin J. Hinich. 1984. *Spatial Analysis of Elections.* New York: Cambridge University Press.

Fiorina, Morris P. 1981. *Retrospective Voting in American National Elections.* New Haven: Yale University Press.

Gerber, Alan and Donald P. Green. 1998. "Rational Learning and Partisan Attitudes." *American Journal of Political Science* 42(3): 794-818.

Gerber, Elisabeth R. and Jeffrey B. Lewis. 2001. "Representing Heterogeneous Districts." Working paper, Ford School of Public Policy, University of Michigan.

Hinich, Melvin J. and Michael C. Munger. 1994. *Ideology and the Theory of Political Choice.* Ann Arbor: University of Michigan Press.

Iyengar, Shanto, Daniel H. Lowenstein, and Seth Masket. 1999. "The Stealth Campaign: Experimental Studies of Slate Mail in California." Paper presented at the Annual Meeting of the American Political Science Association, Atlanta, GA.

Kahn, Kim Fridkin and Patrick J. Kenney. 1997. "A Model of Candidate Evaluations in Senate Elections: The Impact of Campaign Intensity." *Journal of Politics* 59(4): 1173-1205.

Krebs, Timothy B. 1998. "The Determinants of Candidates' Vote Share and the Advantages of Incumbency in City Council Elections." *American Journal of Political Science* 42(3): 921-935.

Lewis, Jeffrey B. 1998. "Who Do Representatives Represent? The Importance of Electoral Coalition Preferences in California." Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA.

Lewis, Jeffrey B. 2001. "Estimating Voter Preference Distributions from Individual-Level Voting Data (with Application to Split-Ticket Voting)." Working paper, Department of Political Science, UCLA.

Lieske, Joel. 1989. "The Political Dynamics of Urban Voting Behavior." *American Journal of Political Science* 33(1): 150-174.

Lupia, Arthur. 1994. "Shortcuts Versus Encyclopedias: Information and Voting Behavior in California Insurance Reform Elections." *American Political Science Review* 88(1): 63-76.

Marra, Robin F. and Charles W. Ostrom, Jr. 1989. "Explaining Seat Change in the U.S. House of Representatives, 1950-86." *American Journal of Political Science* 33(3): 541-569.

McLachlan, Geoffrey and David Peel. 2000. *Finite Mixture Models.* New York: John Wiley & Sons.

Norrander, Barbara. 1986. "Correlates of Vote Choice in the 1980 Presidential Primaries." *Journal of Politics* 48(1): 156-166.

Norrander, Barbara. 1993. "Nomination Choices: Caucus and Primary Outcomes, 1976-88." *American Journal of Political Science* 37(2): 342-364.

Poole, Keith and Howard Rosenthal. 1985. "A Spatial Model for Legislative Roll Call Analysis." *American Journal of Political Science* 29: 355-384.

Poole, Keith and Howard Rosenthal. 1997. *Congress: A Political-Economic History of Roll Call Voting.* Oxford: Oxford University Press.

Popkin, Samuel L. 1991. *The Reasoning Voter.* Chicago: The University of Chicago Press.

Rahn, Wendy. 1993. "The Role of Partisan Stereotypes in Information Processing about Political Candidates." *American Journal of Political Science* 37(2): 472-296.

Riker, William H. and Peter C. Ordeshook. 1968. "The Calculus of Voting." *American Political Science Review* 62: 25-42.

Sears, David O. and Nicholas A. Valentino. 1997. "Politics Matters: Political Events as Catalysts for Preadult Socialization." *American Political Science Review* 91(1): 45-65.

Shanks, J. Merrill and Warren E. Miller. 1991. "Partisanship, Policy and Performance: The Reagan Legacy in the 1988 Election." *British Journal of Political Science* 21(2): 129-197.

Stone, Walter J., Ronald B. Rapoport and Alan I. Abramowitz. 1992. "Candidate Support in Presidential Nomination Campaigns: The Case of Iowa in 1984." *The Journal of Politics* 54(4): 1074-1097.

Stone, Walter J., Ronald B. Rapoport and Lonna Rae Atkeson. 1995. "A Simulation Model of Presidential Nomination Choice." *American Journal of Political Science* 39(1): 135-161.

Trevor, Margaret C. 1999. "Political Socialization, Party Identification, and the Gender Gap." *Public Opinion Quarterly* 63(1): 62-89.

Wattier, Mark. 1983. "Ideological Voting in 1980 Republican Presidential Primaries." *The Journal of Politics* 45(4): 1016-1026.

Whiteley, Paul F. 1988. "The Causal Relationships between Issues, Candidate Evaluations, Party Identification, and Vote Choice–the View From "Rolling Thunder." *The Journal of Politics* 50(4): 961-984.

Williams, Daniel C., Stephen J. Weber, Gordon A. Haaland, Ronald H. Mueller and Robert E. Craig. 1976. "Voter Decision-making in a Primary Election: An Evaluation of Three Models of Choice." *American Journal of Political Science* 20(1): 37-49.