Ballot Design and Unrecorded Votes in the 2002 Midterm Election

By

David Kimball
University of Missouri-St. Louis
Kimballd@umsl.edu

and

Martha Kropf
University of Missouri-Kansas City
Kropfm@umkc.edu

Paper Prepared for Presentation at the 99th Annual American Political Science Association Meeting, August 28-31, 2003, Philadelphia, PA.

***The authors wish to give special thanks to Amanda Baker, Lindsay Battles and Susan Mason for their helpful research assistance.**

****This research was supported by a grant from the University of Missouri System Research Board.**

ABSTRACT

The 2000 presidential election focused the attention of Americans on the problem of unrecorded votes. Approximately two million Americans failed to cast a valid vote in the 2000 presidential contest. While much recent research has evaluated voting technologies and their effects on unrecorded votes, there is little research on the effects of ballot design. In addition, while many election officials have consulted with graphic design experts in the wake of the 2000 elections, many ballot features remain unexamined. Political scientists can draw from the same theory used to evaluate self-administered surveys in order to analyze ballot features. Little published research exists comparing ballots actually used in different places to see which ballot features, if any, correlate with high levels of unrecorded votes.

In this study, we create a dataset for 2002 that considers factors of ballot design, voting technology and demographic factors such as race. Paper-based ballots from counties in six states are collected (Iowa, Kansas, Florida, Missouri, Tennessee and Illinois). We code the ballots in terms of several graphic design elements including whether ballot instructions are written in plain language, whether instructions are printed in a visible font, and whether lists of candidates running for the same office are clear and uncluttered. Our preliminary analyses suggest that several ballot features, including the content and location of instructions and the layout of candidate names, influence unrecorded votes. We expect that an understanding of not just voting equipment, but also ballot design, will play an important role in the implementation of the Help America Vote Act.

Introduction

The 2000 presidential election and the Florida recount controversy brought to light the phenomenon of unrecorded votes (in which some voters come to polling places but fail to cast a valid vote for a particular contest). Roughly two million voters in the United States (almost one out of every fifty to cast a ballot) failed to cast a valid vote for president in the 2000 election (Caltech/MIT 2001; Kimball and Owens 2002). The Florida imbroglio has prompted a new wave of research on election administration and a flurry of election reform laws in Congress and state governments.

A large part of the election reform effort is devoted to replacing outdated voting equipment, particularly the punch card ballots that were so problematic during the Florida recount. Similarly, much recent research has evaluated voting technologies and their effects on unrecorded votes. These studies generally agree that punch card ballots perform worse then other voting methods and that equipment with an error correction feature reduces the frequency of unrecorded votes (see Kimball 2003; Caltech/MIT 2001; Kimball and Owens 2002; Knack and Kropf 2002, 2003; Bullock and Hood 2002; Tomz and Van Houweling 2003). In contrast, aside from studies of the "butterfly ballot" used in Palm Beach County, Florida in 2000 (Wand et al. 2001; Jewett 2001; Kimball, Owens and Keeney 2003), there has been little research on the effects of ballot design features on unrecorded votes. As a result, not much is known about whether any ballot features (such as the readability of voting instructions, the font and size of ballot text, and other graphic design elements) are associated with voting errors, even as some election officials have begun consulting with graphic design experts on ballot layout (Omandam

2001) and political scientists have begun noting ballot features that might confuse voters (Niemi and Herrnson 2003).

This paper is a first effort to examine several ballot features and their impact on unrecorded votes. We use theories and concepts from several disciplines (including survey methodology, graphic design, human factors, and optometry) to identify ballot features that are expected to produce higher rates of unrecorded votes. We collected and coded paper and optical scan ballots used in over 280 counties and six states during the 2002 general elections. In preliminary analyses, we find that several features, including the location of voting instructions and layout of candidate names, appear to be associated with higher rates of unrecorded votes.

Previous Literature Examining Unrecorded Votes

Unrecorded votes are known by many names in the literature including roll-off, residual votes, lost votes or voided votes. However, they are defined commonly by the difference between total turnout and the number of valid votes cast in a particular contest. Unrecorded votes occur as the result of undervotes (where voters intentionally or unintentionally record no selection) or overvotes (where voters select too many candidates, thus spoiling the ballot). Theories explaining the incidence of unrecorded votes can be separated into at least three perspectives. First, scholars argue that some unrecorded votes are intentional for reasons such as voter fatigue, lack of a desirable choice or low information about a contest (e.g. Bullock and Dunn 1996; Knack and Kropf 2003; Kimball, Owens, and McAndrew 2001; Wattenberg et al. 2000).

Second, some researchers argue that accidental undervotes and overvotes occur due to faulty equipment or confusing ballot design (e.g. Knack and Kropf 2003; Kimball and Owens

2002; Darcy and Schneider 1989; Shocket *et al.* 1992; Nichols and Strizek 1995; Caltech/MIT, 2001; Jewett 2001). Several studies have examined the effects of different voting technologies (Knack and Kropf 2003; Kimball, Owens, and Keeney 2003; Kimball, *et al.* 2001; Caltech/MIT 2001; Kimball 2003; Bullock and Hood 2002). In response to problems identified by the 2000 election, several counties and states have replaced older voting technologies with optical scan methods or electronic voting machines (see Kimball 2003). In addition, the occasional practice of listing candidates for the same office in multiple columns or on multiple pages produces higher rates of unrecorded votes (Jewett 2001; Herron and Sekhon 2003; Kimball, Owens, and Keeney 2003).[1]

A final research perspective focuses on equal protection issues, analyzing the relationship between unrecorded votes and demographic variables such as ethnicity or age. For example, there is extensive evidence that unrecorded votes are more common in precincts and counties with large populations of racial and ethnic minorities, low-income residents, less-educated citizens, or elderly voters (Walker 1966; Vanderleeuw and Engstrom 1987; Darcy and Schneider 1989; Sheffield and Hadley 1984; Nichols and Strizek 1995; Herron and Sekhon 2003; Knack and Kropf 2003; Tomz and Van Houweling 2003). Furthermore, there appears to be an interaction between demographic variables and some voting methods and ballot features. The association between socioeconomic measures and unrecorded votes is weaker in places using equipment (such as error correction mechanisms) or ballot features (such as a straight-party option) that make it easier for voters to complete a valid ballot (Knack and Kropf 2003; Tomz and Van Houweling 2003; Kimball, Owens, and Keeney 2003). By the same token, the elevated

[1] In Florida in 2000, Jewett (2001) shows that counties who followed the state-prescribed ballot design had a much lower level of invalidated ballots than those did not. According to Jewett (2001: 8): "a legal, simple and easily understood design consists of placing all candidates in one column on one page that can be marked to the left of the candidate or issue."

rate of unrecorded votes associated with confusing ballots and voting technology tends to fall disproportionately on precincts and counties with high concentrations poor, elderly, or minority voters (Knack and Kropf 2003; Darcy and Schneider 1989; Nichols 1998; Kimball, Owens, and Keeney 2003; Herron and Sekhon 2003; Tomz and Van Houweling 2003).

Overall, there has been more research devoted to the impact of voting technologies than to the effects of ballot design. Many ballot features remain unexamined. A recent study by Niemi and Herrnson (2003) identifies several ballot features in different states that may be confusing. For example, in some jurisdictions ballot instructions include double negatives or other confusing language (Niemi and Herrnson 2003). In other jurisdictions, ballots are cluttered with items (such as a candidate's hometown or occupation) that may obscure the most critical information. However, to our knowledge no study has examined whether these ballot features are associated with higher rates of unrecorded votes. Thus, despite concern about features that might make ballots more confusing, little published research exists comparing ballots actually used in different places to see which ballot features, if any, correlate with high levels of unrecorded votes.[2]

Methodology

We create a dataset for the 2002 midterm elections that includes factors of ballot design, voting technology and demographic factors such as race and education. Election administration is very decentralized in the United States. This produces quite a bit of variation in voting methods and ballots even within the same states. The unit of analysis for the study is the county,

[2] One important ballot format issue is candidate order, which has received extensive treatment in political science literature (Bain and Hecock 1957; Darcy 1986; Krosnick et al. Forthcoming; Koppel and Steen forthcoming). Generally, there is a vote-getting advantage to being the first name listed on the ballot. However, candidate order is not necessarily a cause of unrecorded votes. We do not expect candidate order to affect unrecorded votes. We also

since voting technology and ballot-design decisions are typically made at the county level in all the states analyzed in this study. Ballots from counties in six states are collected (Iowa, Kansas, Florida, Missouri, Tennessee and Illinois[3]) from the November 2002 election.[4] These states were chosen because they use a variety of voting methods and ballots, and each state featured a competitive statewide race in 2002. We mailed surveys to each county or election district official in the these states for information including the number of ballots cast in the county, the number of votes cast for each candidate (Republican, Democratic and other), and the number of overvotes and undervotes.[5] On the premise that more training for poll workers may result in fewer unrecorded votes, we also asked each county about the number of hours of training offered to pollworkers. The survey also ascertained the type of voting equipment used for election day vote tabulating, as well as for early voting, voting absentee by mail and absentee early. Nonrespondent counties were contacted via telephone, but some missing data were filled in with state reports (Tennessee and Florida).[6] For this paper, we coded the paper-based ballots (including optical scan and hand-counted paper ballots) in terms of several graphic design elements as indicated in the next section (See Appendix B for coding list). All in all, voting data and ballots (paper and optical scan) were obtained for 285 counties.

---

do not deal with the issue of ballot length (Bain and Hecock 1957; Allen 1906). Our data concerns unrecorded votes at or near the top of the top of the ticket, rather than those further down on the ticket.

[3] One might argue that using data from only six states limits the generalizability of this study. However, based on 2000 census figures, our sample is quite similar to the rest of the country in terms of the percentage of African American residents (13.4% in our sample versus 12.6% in the rest of the country), the percentage of Hispanic residents (10.4% versus 12.9%), the percentage of citizens over the age of 65 (14.4% versus 12.0%), the percentage of adults with a college degree (19.0% versus 10.6%), and median household income ($37,126 versus $39,699).

[4] We requested that election officials submit one ballot to us that was most representative of the county or was from the largest precinct.

[5] Depending the type of races they had (Governor or Senate) and whether they had implemented early voting, each state received a slightly different survey. See Appendix A for two examples of the survey. Special thanks to Valley Renshaw, UMKC student, for his work in designing the election official surveys.

[6]See Florida Division of Elections, *Analysis and Report of Overvotes and Undervotes for the 2002 General Election*, Available at http://election.dos.state.fl.us/reports/pdf/03OverUnderVotes.pdf. See also http://www.state.tn.us/sos/election.htm.

To measure the number of unrecorded votes for a particular office in each county, we calculate the difference between the total number of ballots cast and the number of votes cast for that office. The distribution of unrecorded votes across counties is somewhat skewed, with outliers at the high end. In our sample, unrecorded vote percentages for Senate contests range from .6% to 22.0%, with a median of 2.2%, a mean of 3.3%, and a standard deviation of 3.0%. Unrecorded vote percentages for gubernatorial contests range from .3% to 6.8%, with a median of 1.4%, and mean of 1.6%, and a standard deviation of 1.1%.

It should be noted that previous studies examining the effects of voting equipment and demographics of individuals in the geographic area have made little attempt to distinguish between vote counting mechanisms for election day, early and absentee voting, despite the fact that many note that using different equipment is a potential source of measurement error in studies of unrecorded votes. This is a potential source of measurement error in studies examining unrecorded votes. In this study, where possible, absentee, early and precinct vote residuals are separated in order to reduce this potential source of measurement error. This particular study focuses on vote counting mechanisms used in the counties on Election Day.[7]

Information Processing and Ballots

In order to expand our understanding of what ballot features may be confusing, one may examine graphic art and information processing theory. Some scholars studying ballots and voting equipment have drawn on human factors and other graphic design research in order to evaluate ballots and equipment (Roth 1994, 1998; Callegaro and Peytcheva 2003), but the work

[7] For a certain percentage of counties, early, absentee and precinct ballots were counted together, or elections officials used the same equipment to count the votes. In those cases, we used total ballots.

needs to be expanded, both in what features are evaluated and the number of jurisdictions that are evaluated.

Survey methodologists have borrowed more heavily from graphic design and information processing theory than political scientists, using the theory to reduce item and survey non-response (see for example, Dillman 2000). We argue that ballots and self-administered questionnaires have many features in common in that the voter alone selects his or her choices on Election Day and must process both verbal language (the written word including instructions and candidate choices) and graphic language (numbers and other symbols and arrangements of choices) presented on the ballot when deciding how to complete the task facing them (Dillman and Jenkins 1995, 1997). The verbal and graphic language on a written form must be organized in a way so that all readers follow a consistent and efficient path to cover all of the information on each page.

In general, the features of a good ballot might include both information organization and navigational tools. Combining both the work of Dillman and his colleagues and some of the work in the area of political science, there are several possible criteria to use to evaluate the degree to which ballots make voting easier. Our research indicates that some issues to consider are ballot-specific, where others are candidate and office-specific.

Location of Instructions

Looking at the ballot as a whole, one first confronts the fact that voters usually look first at the upper left-hand corner of the ballot. In western culture we typically begin reading in the top left-hand corner of a document (Brandt 1945, c.f. Jenkins and Dillman 1997; Dillman 2000:

113).[8] Indeed, this is confirmed by Roth's study of voting (1994: 59). In addition, work in survey research indicates that people may not read directions (Gower and Dibbs 1989). Thus, instructions should be as close to the first response task as possible (Dillman 2000: 98-99; Zukerberg and Lee 1997; Dillman and Christian 2002). The theory here is based on Gestalt psychology's Law of Proximity: placing items close together encourages readers to view them as a group (Dillman 2000, 107; Wallshlaeger and Busic-Snyder 1992).

Instructions separated from the response task are coded in this study. For example, instructions may appear at the top of the ballot, rather than in the upper left quadrant, just before the first office to be voted on. We code ballots whether voting instructions begin in the top left corner of the survey right above the first response task.

In 52% of the ballots in our sample, voting instructions are in the top left corner, right above the first contest. In 44% of the ballots, instructions are spread across the entire width of the ballot at the top. In the other 4%, either no instructions appeared or the instructions appear somewhere in the middle of the ballot below some of the contests. We expect that unrecorded votes are less common in counties where voting instructions are in the upper left corner of the ballot.

Readability

Human factors research also indicates that instructions should be easy to read. Sentences and words should be short and simple, written in an active, affirmative style (Sanders and

[8] This is consistent with top-down processing, as noted by Dillman and Jenkins (1995, 1997). Individuals have certain expectations about how to approach potentially unfamiliar tasks. Thus, pattern recognition is key to the process of taking on a task such as completing a survey or voting. "…[T]op-down processing emphasizes the role of context and expectations in identifying a pattern. In this case, our knowledge about how the world is organized helps us to identify patterns" (Jenkins and Dillman 1995: 5).

McCormick 1993: 110; Dillman 1978: 111; Zukerberg and Lee 1997).[9] In general, readability describes the ease of processing the information content of written words.[10] To measure the readability of each ballot's voting instructions, we type them into Microsoft Word and compute Flesch-Kincaid Grade Level scores. The use and validity of the Flesch-Kincaid scores are supported by other studies (Heilke, Joslyn, and Aguado 2003; Sanders and McCormick 1993; Tefki 1987). The Flesch-Kincaid scores indicate the grade level needed to understand the text.[11] Higher scores indicate documents that are harder to read.

In our sample, the grade level scores for ballot instructions range from 4.0 (4th grade) to 12.0 (12th grade). The mean reading score for our sample is 8.8, suggesting that the average ballot in our sample required almost a ninth grade education to understand the text. We expect that unrecorded votes are more common in counties with high reading level scores for voting instructions.

Spacing Between Candidates

We also code for appropriate spacing between candidates. Appropriate space between lines improves legibility and readability (Roth 1994: 54; Sanders and McCormick 1993, 109). Spacing also highlights "appropriate groupings of visual elements" (Dillman 2000: 107). However, if there is too little space between candidate names, that may prevent a voter from seeing the candidate of his or her choice, and therefore skipping the office altogether.

---

[9] In examining ballots from all over the country, Niemi and Herrnson (2003) found confusing language in instructions on how many votes to cast: "vote once" etc.

[10] In addition, instructions need to be legible (Roth 1994: 54). Voters should be able to recognize the letters on the ballot.

[11] The Flesch-Kincaid Grade Level scores are based on the length of words and sentences. The formula for computing the score is (.39 x ASL) + (11.8 x ASW) – 15.59, where ASL is the average number of words per sentence and ASW is the average number of syllables per word.

In looking at the spacing between candidates, we find that 89 percent of the ballots have at least one line of space between the candidates; 11 percent have no space between the candidates. It is also possible to separate candidates with lines between their names—only about 11 percent of the ballots had lines between the candidates.

Font Size

The size of the ballot text is another factor to consider, especially a font size large enough for voters to read - 12 or 14 point is recommended for elderly readers (Roth 1994). Since visual acuity declines with age, there may be an interaction between font size and age (Roth 1994). One may also vary the font size within the questionnaire/ballot to attract attention and guide the respondent/voter (Dillman 2000: 106). In this study, instructions, office titles and candidate names are coded for font size.[12] We hypothesize that smaller font size will lead to greater levels of unrecorded votes. We find that the mean of the instruction font size to be 9.33 (ranging from 7-point to 14-point font),[13] the mean of the candidate font size is 10.27 (ranging from 7-point to 16-point font) and the mean of the office font size is also 10.27 (ranging from 7- to 14-point font). The vast majority of ballots had font sizes much below that recommended by graphic design principles (91 percent of instruction font sizes, 82 percent of candidate fonts and 79 percent of office fonts were 11-point font or below).

Spoiled Ballot Instructions

[12] It should be noted that for several of the ballots, copies were obtained that may or may not be a good representation of the actual size of the ballot. Font size for these ballots cannot be determined and were coded as "missing." However, for some ballots (such as Iowa's), it was determined that one could estimate the font size based on the fact that there are only about seven different optical scan ballot styles in the state.

[13] There may be some measurement error in font size for the instructions. Many ballots had more than one font size listed, so we coded the "main" instruction font size: that which told the voters how to vote (i.e., darken the oval

We also examine whether ballots include instructions about spoiling the ballot, including a warning about the consequences of spoiling a ballot and how to correct ballot errors. We examine this ballot feature because the recently passed Help America Vote Act of 2002 (HAVA) requires local election officials to implement more vigorous voter education programs. In particular, HAVA requires jurisdictions with paper ballots to include instructions on the effect of multiple votes for a single office and how to correct a spoiled ballot (HAVA 2002, Title III, section 301.a.1.B).

About 20 percent of the ballots we examined did not warn voters not to spoil their ballots or inform the voters what to do if they did spoil their ballots. Sixty-three percent of ballots told voters they could turn in their ballots to an election judge or poll worker and obtain a new ballot if they made a mistake or voted for an incorrect candidate. Only about seventeen percent of ballots include a warning that votes will not count if the ballot is spoiled and explain what to do in the event they spoiled their ballots.

The Use of Shading and Bolding

An issue that is both ballot-related and candidate related is the use of shading and boldface fonts to guide the voter. According to Jenkins and Dillman (1997), not all of what we look at is of equal visual interest. Jenkins and Dillman point to Kahneman (1973) who says that people will focus on areas that are physically informative: "high contrast areas," which can be used to a designer's advantage (Jenkins and Dillman 1997: p. 11; See also Dillman 2000: 106). The appropriate response choices (e.g., the list of candidates on a ballot) can be identified more quickly if they are differentiated from other sections (Redline and Dillman cite Foster 1979).

completely to the left of the candidate you wish to select). Other instructions included how to vote a straight ticket or told voters the location of the judicial offices.

Shading can encourage the appropriate grouping of information and guide the respondent to the correct spot on the ballot (Dillman 2000; Design for Democracy; Omandam 2002). Reverse shading could also be used to emphasize certain parts of the ballot (Redline and Dillman 2001: 8). Bolding may be used to highlight the questions or office sections on which a voter is working; light print may be used for the candidates, to make the candidates different from the office names and highlight the response task (Dillman 2000: 118). Thus, we examine shading and bolding to highlight each contest on the ballot and shading to draw attention to the candidates listed for each office.

Most ballots (69 percent) do not have any shading of any kind at all; 29 percent employ shading to highlight groups of offices, such as federal offices or state offices. Only three percent use shading to highlight the different offices themselves. Also, most ballots do not use bolding to highlight candidates for office or make the candidates stand out from the office for which they are running. In only about one-third of cases do the candidates stand out from the office titles (35 percent) in terms of boldfaced text or shading.

Multiple Columns for Candidates

In terms of candidate-level ballot issues, the Law of Proximity as well as recent political science research indicates that candidates for the same office should not be listed in two or more columns (Darcy and Schneider 1989; Jewett 2001; Cauchon 2001; Wand et al. 2001; Kimball, Owens, and Keeney n.d.; Kimball 2003). Each office we are concerned with is coded as such. Only four ballots used multiple columns for senate candidates (1.6 percent); none did for the governor's races we coded. By and large, election officials have learned from Florida about the need to list candidates in a single column.

Justification

Other work in survey research sheds light on the horizontal location (justification) response choices. Generally, survey methodology research recommends that response categories be left justified, with the appropriate box or circle located to the left of the each category (de Vaus 2002; Dillman 1978: 133). Building on the work of Dillman (1978: 133), we expect that response categories are best justified to the left, indented from the question. In our sample, ninety percent of the ballots had the box or circle to mark and the candidate both left justified on the ballot.

Making it Easy to Mark the Correct Box or Oval

Indeed, one possible reason for high levels of unrecorded votes may be that voters are marking the wrong box for the candidate. Aside from the simple justification of the choice, the survey methodological research indicates that there should be no ambiguity about which box or circle corresponds to which candidate, in keeping with the Law of Proximity. In particular, "relative closeness" of items to each other can cause one to see the items as one group (Roth 1994, 63; see also Jenkins and Dillman 1997.) Thus, the ballots in our sample are coded as to whether there is a circle, arrow, or square on both sides of the candidate names and whether there is any potential confusion about which circle belongs with which candidate. Even though most ballots have the response location and the candidates left-justified, approximately 29 percent of senate races and 28 percent of gubernatorial races left some confusion about where to mark a vote in that race. That is to say, there may have been a box or an oval on both sides of the candidate’s names. This, of course, includes cases where the ovals/boxes/circles are in a separate column.

Clutter Around the Candidates

Finally, both political scientists and survey researchers have indicated that ballots (and self-administered surveys) must be spread out and uncluttered, especially around candidate names (Niemi and Herrnson 2003; Babbie 1990). Thus, we code ballots for whether they avoid clutter around candidate names, such as occupations, incumbency, hometown (Niemi and Herrnson 2003; Babbie 1990). About 55 percent of gubernatorial races had excessive clutter around the candidates, while only about 35 percent of senate races had this problem. We hypothesize that more clutter will result in more errors in voting.

Findings

To assess the impact of ballot features, we estimate a model of unrecorded votes in the 2002 gubernatorial and U.S. Senate elections for the counties in our sample. The model includes the ballot features described above, voting technology, and demographic measures as explanatory variables.

In this preliminary analysis, we examine seven ballot features as explanatory variables.[14] The analysis contains three explanatory variables associated with the voting instructions. One explanatory variable is the Flesch-Kincaid Grade Level score for the voting instructions on each ballot. Instructions requiring higher reading ability should be associated with higher levels of unrecorded votes. Second, we include a measure indicating whether or not the voting instructions contain specific sentences about spoiled ballots (as required by HAVA). Third, we include a measure indicating whether the instructions are located at the top left corner of the ballot, the spot where most voters will train their eyes first. We expect the latter two instruction variables to be associated with lower levels of unrecorded votes.

The analysis contains four explanatory variables dealing with the layout of candidate choices. First, we include a dichotomous variable indicating whether or not the candidate choices are cluttered with other information, text, or graphics. A second variable indicates whether circles or squares are located on both sides of candidate names, which may confuse voters about which one to mark for their chosen candidate. We expect higher levels of unrecorded votes with cluttered ballots and ballots with marking options on both sides of candidate names. A third measure indicates whether shading is used to guide the voter to each office or group of similar offices. A fourth measure indicates whether boldfaced text is used to differentiate candidate names from the office for which they are competing. According to the design principles described in the previous section, we expect fewer unrecorded votes where shading and boldfaced text are used in these ways.

The counties in our sample use hand counted or optically scanned paper ballots. We include a dummy variable for counties using precinct-count optical scan systems, since they have an error-correction feature that reduces unrecorded votes. Finally, the regression model includes a number of demographic variables that are often correlated with unrecorded votes. These control variables include the percentage of a county's residents who are African-American, the

[14] Given the size of our sample and strong correlations between some of the ballot features, we could not include each element as a separate explanatory factor in a multivariate analysis. In the future, we will probably develop an index to measure how well a ballot satisfies each of the criteria described above.

percentage over the age of 65, and the percentage of adults with a high school degree.[15] We expect unrecorded votes to be positively correlated with the size of the African-American and elderly populations, and unrecorded votes should be negatively correlated with the percentage of high school graduates.

We estimate a regression model to calculate the impact of each explanatory variable on unrecorded votes. We estimate separate models for governor contests, U.S. Senate contests, and the top contest on the ballot (which varies from state to state).[16] Since the number of voters in each county varies dramatically, we weight each county by the number of ballots cast. In addition, we estimate robust standard errors to correct for heteroskedasticity likely in data with a skewed dependent variable (White 1980). Finally, the model is modified to account for the fact that the count of unrecorded votes in a county is a function of the total number of ballots cast. Even with the greatest voting equipment and the most helpful poll workers, a county with 100,000 voters will probably have more unrecorded votes than a county with 1,000 voters.

In general, our preliminary analyses suggest that several ballot design features affect the rate of unrecorded votes in the way we hypothesize (see Table 1).[17] Ballot features are more successful predicting unrecorded votes in gubernatorial races and the top contest, but ballot features are less successful in the Senate model. An obvious question is why the ballot features do not fare as well in the Senate model.

The location and content of ballot instructions are correlated with unrecorded votes in the top contest and the governor's contest, as expected. In two out of three models, placing response

---

[15] We obtained demographic data from the U.S. Census Bureau's American FactFinder web site (http://factfinder.census.gov/servlet/BasicFactsServlet).

[16] Even though the senate race is the top contest in Kansas, for the purpose of this analysis, we used the gubernatorial unrecorded vote rate as the top contest dependent variable. There was no Democratic candidate in the U.S. Senate race in Kansas, creating more intentional undervoting.

[17] We also conducted the multivariate analyses using a negative binomial regression model (Long 1997). The results are substantively similar to those presented in Table 1.

options on both sides of candidate names is associated with higher rates of unrecorded votes (the effect falls just short of statistical significance in the Senate model). Furthermore, the use of shading and bold text to set candidate names apart from office titles seems to reduce the frequency of unrecorded votes, at least for the top contest and the governor's race.

Finally, the controls for voting technology and demographics perform as expected. Unrecorded votes tend to be less common in counties using precinct-count optical scan systems that have an error correction mechanism. In addition, unrecorded votes are more common in counties with large concentrations of elderly and African-American voters. Higher percentages of high school graduates in a county is negatively related to unrecorded votes, though the relationship is not statistically significant.

Conclusion

It is ironic that survey researchers have spent so many resources in determining survey features that will decrease survey and item non-response, yet election ballots in the United States receive very little of the same type of attention. Before the 2000 election, ballot non-response was not a major concern for election officials. In the wake of the controversy wrought by unrecorded votes in Florida and subsequent legislation intended to reduce the frequency of unrecorded votes, election officials have pursued a number of reforms. Since the 2000 election and the increase in awareness that formats such as the butterfly ballot are confusing, election officials and scholars have spent more time analyzing voting equipment, ballots and other sources of error on Election Day. In particular, much of the reform effort has focused on voting technology upgrades. While new voting technology is likely to help, voting equipment is not the only source of voting error. This study describes several ballot features that appear to be associated with unrecorded votes. Purchasing new voting equipment can be very expensive and difficult during the current retrenchment in state and local government budgets. In contrast, modifying ballot features to create a layout that is easier to read and follow may be relatively inexpensive. Ballot design deserves closer inspection than it has received thus far.

Appendix A: Copies of Surveys Sent to Election Officials

- Copy of survey sent to Missouri election officials
- Copy of survey sent to Tennessee and Kansas election officials

We appreciate you taking the time to answer the following questions.

1. **Please send** a copy of one of the *November 5, 2002 ballots* for your county in the enclosed self-addressed, stamped envelope. Also, if available, please send a copy of the November 7, 2000 ballot.

2. **What type** of voting technology was used on Election Day, 2002 in your county? (Please circle one)

Paper Ballots | Lever Machines | Punch card | Data Vote

Optical Scan with Central Count | Electronic or DRE Technology | Optical Scan with Precinct Count

3. **Please specify** the make and model number of the voting machine used in your county on Election Day, 2002.

4. **If you use** precinct count optical scan equipment, were voters allowed to check their ballots for errors at the poll?

5. **Was different** equipment used for early/absentee voting by mail? **If so**, what technology was used for early/absentee voting by mail?

6. **How many** hours of training must election judges complete (if any)?

7. **We are also** interested in the vote totals for ballots cast on Election Day and ballots cast by absentee voting received in *November 2002* in your county.

Vote Type	Election Day	Absentee
Total ballots cast (including those on which no valid vote was cast)		
Votes for Democratic Senate candidate		
Votes for GOP Senate candidate		
Votes for other Senate candidates		
Overvotes for Senate		
Undervotes for Senate		

Thank you for your cooperation.

We appreciate you taking the time to answer the following questions.

1. **Please send** a copy of one of the *November 5, 2002 ballots* for your county in the enclosed self-addressed, stamped envelope. Also, if available, please send a copy of the November 7, 2000 ballot.

2. **What type** of voting technology was used on Election Day, 2002 in your county? (Please circle one)

Paper Ballots   Lever Machines   Punch card   Data Vote

Optical Scan with Central Count   Electronic or DRE Technology   Optical Scan with Precinct Count

3. **Please specify** the make and model number of the voting machine used in your county on Election Day, 2002.

4. **If you use** precinct count optical scan equipment, were voters allowed to check their ballots for errors at the poll?

5. **Was different** equipment used for early/absentee voting by mail? **If so**, what technology was used for early/absentee voting by mail?

6. **How many** hours of training must election judges complete (if any)?

7. **We are also** interested in the vote totals for ballots cast on Election Day and ballots cast by absentee and early voting received in *November 2002* in your county.

Vote Type	Election Day	Absentee	Early
Total ballots cast (including those on which no valid vote was cast)			
Votes for Democratic Senate candidate			
Votes for GOP Senate candidate			
Votes for other Senate candidates			
Overvotes for Senate			
Undervotes for Senate			
Votes for Democratic Governor candidate			
Votes for GOP Governor candidate			
Votes for other Governor candidates			
Overvotes for Governor			
Undervotes for Governor			

Thank you for your cooperation.

Appendix B: Coding for ballot study

County and state:____
Coder:______________________________________
[including changes made August 6, 2003]

BALLOT-SPECIFIC VARIABLES

Variable #1: BLOCK:
Party column v. office block
1--office block
0--party column

INSTRUCT: Are the instructions in the top left of the first column?

1—top left of first column
2—spread across top
3--instructions on a separate page
4--instructions come after response task
5--Doesn't tell you what to do to vote (may include for example, straight ticket instructions, but nothing else
9—Not available (instructions missing)

SHADE: Location of shading

0—no shading on ballot
1—shading for groups of offices
2—shading for each office (includes reverse shading)

REVSHADE: Is there any reverse shading?

0—no
1—yes

CANRESH: Is there reverse shading of candidates?

0—no
1—yes

__NAMEOUT__: Do candidates' names stand out from the office? (is one more bold? Is one shaded?)

0—no
1—yes

BOLD: Boldness of candidates v. office they are running for

1—is the same
2—office is bolder than candidates
3—candidates is bolder than office

LINEBT: Horizontal lines between candidates for office

0—no
1—yes

SPACEBT: Is there space between the candidate lines?

0—no, squished
1—yes, not squished

TYPEBAL: Ballot type

0—circle
1—arrow
2—square
9—Missing, not available or not relevant (punchcard, DRE)

__GRADEINS__: Grade level of instructions

- Type the ballot instructions in the "ballot instruction file" and then analyze grade level using MS Word

K. __READABLE__: Readability score of instructions

- Type the ballot instructions in the "ballot instruction file" and then analyze readability using MS Word

L. __GRADEST__: Grade level of __straight ticket__ instructions (include any split ticket instructions as well)

- Type the straight ticket instructions in the straight ticket instruction file and analyze grade level MS Word

M. __READST__: Readability score of __straight ticket__ instructions (include split ticket instructions as well)

*N. **SPOIL**: Instructions of what to do in the case of a spoiled ballot (HAVA!)*

0—nothing
1—Warning not to spoil ballot
2—What to do if you spoil your ballot
3—Both 1 and 2
9—No instructions at all available

*O. **PARTY**: Is party spelled out? (as opposed to being abbreviated)*

0—no it is abbreviated
1—yes, it is spelled out
9--Not included right with the candidate

P. **PARTYPOS**: Is party off to the side (right) or under SENATE candidate name?

0—side
1—under
2—below and to the right
3--neither

P#1: **PARTPOSG**: Is party off to the side (right) or under GUBERNATORIAL candidate name?

0—side
1—under
2—below and to the right
3--neither

P#2: **PARTPOST**: Is party off to the side (right) or under TOP CANDIDATE WE CONSIDERS

candidate name?

0—side
1—under
2—below and to the right
3--neither

Q. **INSTFONT**: Instruction font size

R. ***CANFONT****: Candidate font size*

S. ***OFFFONT****: Office font size*

T. ***PARTFONT****: Party font size*

U. ***JUST****: Justification of candidate names and circles/arrows/squares for vote*

1—circle/arrow/square and candidate names left justified and next to each other
2—circle/arrow/square and candidate names left justified but **not** next to each other
3—circle/arrow/square and candidate names right justified and next to each other
4—circle/arrow/square and candidate names right justified but **not** next to each other
9—not relevant (DRE and punchcard?)

Office-Specific Variables

V. ***TOP****: What is top contest?*

1—governor
2—senate
3—other state office
4—other federal office

V#1: ***TOPDV****: What is the top contest we consider?* ***(TOP DV IS TOP ONE WE CONSIDER)***

1—governor
2—senate

W. **COLSEN**: Candidates in same column for Senate

1—single column
2—multiple columns (includes party column ballot)
9--no Senate contest in state

X. **COLGOV**: Candidates in same column for Governor

1—single column
2—multiple columns (includes party column ballots)
9--no Governor's race in state

Y. **COLTOP**: Candidates in same column for top race (top dependent variable or race that we consider)

1—single column
2—multiple columns (includes party column ballots)

Z. ***CLUTTSEN****: Clutter for Senate (more than candidate and party listed)(Party column ballots usually are cluttered by our judgement)*

0—no
1—yes
9--no senate race in state

AA. ***CLUTTGOV****: Clutter for Governor (more than candidate and party listed))(Party column ballots usually are cluttered by our judgement)*

0—no
1—yes
9--no governor's race in state

*BB.****CLUTTTOP****: Clutter for top race or dependent variable that we consider (more than candidate and party listed))(Party column ballots usually are cluttered by our judgement)*

0—no
1—yes

CC. ***NOBEFSEN****: How many offices come before Senate?*

(9=No senate race in state)

DD. ***NOBEFGOV****: How many offices come before Governor?*
(9=no governor's race in state)

EE. ***NOBEFTOP****: How many offices come before top dependent variable?* ***(TOP DV IS TOP ONE WE CONSIDER***

FF. ***CIRLCSEN****: Is there a circle/arrow/square on both sides of Senate candidate names?*

0—no
1—yes

9--no senate race in state

GG. **CIRLCGOV**: Is there a circle/arrow/square on both sides of Governor candidate names?

0—no
1—yes
9--no governor's race in state

HH. **CIRCLTOP**: Is there a circle/arrow/square on both sides of top race candidate

names?(TOP DV)

0—no
1—yes

*II. **MARKSEN**: Any possible confusion about which circle/arrow/square to mark for Senate candidates?*

0—no
1—yes
9--No senate race on ballot

*JJ. **MARKGOV**: Any possible confusion about which circle/arrow/square to mark for Governor candidates?*

0—no
1—yes
9--no governor's race on ballot

*KK. **MARKTOP**: Any possible confusion about which circle/arrow/square to mark for top race candidates? (top DV)*

0—no
1—yes

*LL. **SPACESEN**: Space between office/directions and first Senate candidate*

0—no
1—yes, but only one line
2--yes, more than one line of space
9--no senate race on ballot

MM. ***SPACEGOV****: Space between office/directions and first Governor candidate*

0—no
1—yes, but only one line
2--yes, more than one line of space
9--no governor's race on ballot

NN. ***SPACETOP****: Space between office/directions and first top race candidate (top DV that we consider)*

0—no
1—yes, but only one line
2--yes, more than one line of space

Sources Cited

Allen, Philip L. 1906. "American Ballot Laws and Their Working." *Political Science Quarterly* 21:38-58.

Bain, Henry M., and Donald S. Hecock. 1957. *Ballot Position and Voter's Choice.* Detroit: Wayne State University Press.

Bederson, Benjamin B., Bongshin Lee, Robert M. Sherman, Paul S. Herrnson, and Richard G. Niemi. 2002. "Electronic Voting Usability Issues," Available at http://www.capc.umd.edu/research.asp. Accessed 1/30/03.

Bullock, Charles S., III, and Richard E. Dunn. 1996. "Election Roll-off: A Test of Three Explanations." *Urban Affairs Review* 32:71-86.

Bullock, Charles S., III and M. V. Hood, III. 2002. "One Person – No Vote; One Vote; Two Votes: Voting Methods, Ballot Types, and Undervote Frequency in the 2000 Presidential Election." *Social Science Quarterly* 83:981-993.

Callegaro, Mario and Emilia Peytcheva. 2003. "Electronic Voting Machines: A Comparison of the Principles of Human Computer Interaction and Computer-Assisted Questionnaire Design." Paper Prepared for Presentation at the 58th Annual Meeting of the American Association for Public Opinion Research, Nashville, TN.

Cauchon, Dennis. 2001. "Errors Mostly Tied to Ballots, not Machines." *USA Today* November 7, 2001, p. 6A.

Darcy, R. 1986. "Position Effects With Party Column Ballots." *Western Political Quarterly* 39:648-662.

Darcy, R., and Anne Schneider. 1989. "Confusing Ballots, Roll-Off, and The Black Vote." *Western Political Quarterly* 42:347-364.

De Vaus, David. 2002. *Surveys in Social Research*, 5th ed. London: Routledge.

Dillman, Don A. 2000. *Mail and Internet Surveys: The Tailored Design Method*. New York: John Wiley & Sons.

Dillman, Don A. 1978. *Mail and Telephone Surveys: The Total Design Method*. New York: John Wiley & Sons.

Dillman, Don A. and Leah Christian. 2002. "The Influence of Words, Symbols, Numbers and Graphics on Answers to Self-Administered Questionnaires: Results From 18 Experimental Comparisons." Unpublished Manuscript available at http://survey.sesrc.wsu.edu/dillman/papers.htm, accessed 6-11-03.

Foster, J.J. 1979. The Use of Visual Clues in Text. *Processing of Visual Language* 1, 189-203. (check cite)

Gower, A.R. and R. Dibbs. 1989. "Cognitive Research: Designing a 'Respondent Friendly' Questionnaire for the 1991 Census." Proceedings from the 5th Annual Census Bureau Research Conference. Washington: Census Bureau.

Help America Vote Act of 2002. U.S. Public Law 252. 107th Congress, 2nd session, 29 October 2002.

Heilke, Thomas, Mark R. Joslyn, and Alex Aguado. 2003. "The Changing Readability of Introductury Political Science Textbooks: A Case Study of Burns and Peltason, *Government by the People*." *PS: Political Science & Politics* 36:229-232.

Jewett, Aubrey. 2001. "Explaining Variation in Ballot Invalidation Among Florida Counties in the 2000 Election." Presented at the annual meeting of the American Political Science Association, San Francisco, August, 2001.

Kimball, David C. 2003. "Voting Methods Two Years After Florida." Presented at the annual meeting of the Midwest Political Science Association, Chicago, April 3, 2003.

Kimball, David C., Chris T. Owens, and Katherine McAndrew Keeney. Forthcoming, 2004. "Unrecorded Votes and Political Representation." In *The State of Elections: From Florida to Reform*, Robert P. Watson, ed. Gainesville: University Press of Florida.

Kimball, David C., Chris T. Owens, and Katherine M. Keeney. 2003. "Unrecorded Votes and Election Reform." *Spectrum: The Journal of State Government* 76(1):34-37.

Knack, Stephen and Martha Kropf. 2003. "Invalidated Ballots in the 1996 Presidential Election: A County-Level Analysis." *Journal of Politics*. 65(3): 881-897.

Knack, Stephen and Martha Kropf. 2002. "Who Uses Inferior Voting Technology?" *PS: Political Science and Politics*. September: 541-548.

Koppel, Jonathan GS, and Jennifer A. Steen. Forthcoming. "The Effects of Ballot Position on Election Outcomes." *Journal of Politics.*

Krosnick, Jon A., Joanne M. Miller, and Michael P. Tichy. 2003. "An Unrecognized Need for Ballot Reform: The Effects of Candidate Name Order on Election Outcomes." In *Rethinking the Vote: The Politics and Prospects of American Election Reform*, ed. Ann N. Crigler and Marion R. Just. New York: Oxford University Press.

Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks, CA: Sage.

Niemi, Richard G. and Paul S. Herrnson. 2003. "Beyond the Butterfly: The Complexity of U.S. Ballots." *Perspectives on Politics* 1(2): 317-326.

Omandam, Pat. 2002. "Ballot may Confuse, Design Expert Says." *Honolulu Star-Bulletin*,September 19 (http://starbulletin.com/2002/09/19/news/story3.html).

Redline, Cleo D. and Don A. Dillman. 2001. "The Influence of Alternative Visual Designs on Respondents' Performance with Branching Instructions in Self-Administered Questionnaires." In Groves, R., D. Dillman, J. Eltinge and R. Little, eds. *Survey Nonresponse*. New York: John Wiley and Sons.

Roth, Susan King. 1994. "The Unconsidered Ballot: How Design Affects Voting Behavior." *Visible Language* 28:48-67.

Sanders, Mark S., and Ernest J. McCormick. 1993. *Human Factors in Engineering and Design*, 7th ed. New York: McGraw-Hill.

Tefki, Chafai. 1987. "Readability Formulas: An Overview." *Journal of Documentation* 43:261-273.

Wallschlaeger, C., and Busic-Snyder, C. 1992. *Basic Visual Concepts and Principles for Artists, Architects and Designers*. Dubuque, IA: William C. Brown Publishers.

Wand, Jonathan N., 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 County, Florida." *American Political Science Review* 95:793-810.

Zukerberg, Andrew and Meredith Lee. 1997. "Better Formatting for Lower Response Burden." Census Bureau Research Report Series, Available at http://www.census.gov/srd/www/byyear.html, accessed 6-30-03.

Table 1
Multivariate Analyses of Unrecorded Votes in the 2002 Elections

Explanatory Variable	*Governor*	*U.S. Senate*	*Top Contest*
Ballot Features			
Cluttered ballot	.08	.98*	.17
	(.13)	(.54)	(.16)
Grade level of instructions	.10**	-.02	.002
	(.03)	(.05)	(.04)
Spoiled Ballot Instructions	-.16*	.39*	-.27*
	(.08)	(.17)	(.11)
Response options on both sides of candidate names	.42**	.39	.55**
	(.13)	(.31)	(.15)
Instructions at top left	-.36*	-.30	-.50**
	(.14)	(.24)	(.17)
Shading to identify different offices	-.26*	.25	-.22
	(.11)	(.25)	(.17)
Candidates visually distinct from office titles	-.20*	.02	-.42**
	(.09)	(.22)	(.15)
Voting Technology			
Precinct Count Optical Scan	-1.01***	-.10	-.81***
	(.15)	(.24)	(.18)
Demographic Controls			
Percent Black	.017*	.044*	.017
	(.010)	(.017)	(.016)
Percent 65 or older	.018**	.111**	.027*
	(.008)	(.033)	(.013)
Percent with a high school degree	-.010	-.017	-.004
	(.007)	(.015)	(.009)
Constant	1.86*	.98	2.07
	(.70)	(1.67)	(.98)
Number of Cases	220	158	282
R-Squared	.59	.38	.36
Root MSE	.51	.92	.79

The dependent variable is the percentage of ballots cast that failed to record a valid vote for the office listed at the top. Cell entries are regression coefficients with robust standard errors in parentheses.
Observations (counties) are weighted by the number of ballots cast in the 2002 election.
Counties in Kansas are excluded from the Senate model because there was no Democratic candidate for Senate in Kansas in 2002.
*** p < .001, ** p < .01, * p < .1, two-tailed