Ballot Formats, Touchscreens, and Undervotes: A Study of the 2006 Midterm Elections in Florida[1]

Laurin Frisina Michael C. Herron James Honaker Jeffrey B. Lewis

[1]Frisina is a Methodology Assistant at Harris Interactive in Rochester, NY, and she received a B.A. in Political Science from the University of Rochester in June, 2006 (LFrisina@HarrisInteractive.com); Herron is Associate Professor of Government, Dartmouth College (Michael.Herron@dartmouth.edu); Honaker is Assistant Professor of Political Science, University of California at Los Angeles (tercer@ucla.edu); and Lewis is Associate Professor of Political Science, University of California at Los Angeles (jblewis@ucla.edu). The authors thank Greg Huber and seminar participants at Dartmouth College and the University of Chicago for comments on an earlier draft of this article and thank election officials in Charlotte, Collier, DeSoto, Hardee, Hillsborough, Lee, Manatee, Pinellas, and Sarasota Counties for providing data and assistance. The initial draft of this article was circulated on the Internet on November 23, 2006, and subsequently Herron and Lewis were retained as expert witnesses by Election Systems & Software in the matter of *Jennings v. Elections Canvassing Commission of the State of the Florida, et al.*, 2nd Judicial Court, Leon County, Florida, 2006 CA 002973.

Abstract

The 2006 midterm elections in Florida focused renewed attention on undervotes, instances where no candidate is selected in a given race. This interest was sparked by the high undervote count—more than 21,000 out of approximately 240,000 ballots cast—in the race for Florida's 13th Congressional seat. This race was decided by a margin of only 369 votes, and the numerous undervotes cast in it were highly concentrated in Sarasota County, one of five counties that contributes to the 13th Congressional District. Using ballot- and precinct-level voting data we show that the exceptionally high Sarasota undervote rate in the 13th Congressional District race was almost certainly caused by the way that Sarasota County's electronic voting machines displayed on a single ballot screen the Congressional contest and the Florida gubernatorial race. We buttress this claim by showing that extraordinarily high undervote rates were also observed in the Florida attorney general race in Charlotte and Lee Counties, places where the attorney general and gubernatorial races were combined on a single ballot screen. Using a precinct-level, statistical imputation model we find that there is a 98 percent chance that the Sarasota undervotes were pivotal in the 13th Congressional District race. With more precise estimates based on ballot data from Charlotte and Sarasota Counties, we find that there is essentially a 100 percent chance that the 13th Congressional District election result would have been reversed in the absence of the large Sarasota undervote.

1 Introduction

Despite widespread concerns that electronic voting machines adopted in the wake of the 2000 presidential election would wreak havoc on the 2006 midterm elections, in the aftermath of November, 2006 there were relatively few reports of election-altering voting machine failures. This fortunate situation arose even though the 2006 elections included a number of closely contested United States Senate (e.g., Virginia), United States House (e.g., the 2nd Congressional District in Connecticut), and state-level (e.g., the Vermont state auditor) races.

One might argue that the relative lack of post-election contentiousness in 2006 resulted from the recent replacement of old voting systems by electronic voting machines and the consequent diminishing of two types of voter errors, undervotes (instances where no candidate is selected in a race) and overvotes (where too many candidates are selected). Electronic voting machines reduce undervotes, it is believed, by reminding voters of races they have overlooked (Celeste, Thornburgh & Lin 2005), and they prevent overvotes by not allowing too many candidates to be

selected (Kimball 2003, Carrier 2005). Critics fear, nonetheless, that electronic voting machines are susceptible to tampering (e.g., Feldman, Halderman & Felten 2006), and the debate surrounding electronic voting has thus been framed as balancing the benefits gained by reducing known sources of voter errors against the risk of tampering or technology failure.

The 2006 midterm elections have shown that this tradeoff is more complicated than had been realized and that perhaps one of the most important risks of electronic voting may be a new incarnation of an old problem. Indeed, the most notable controversy associated with electronic voting in November, 2006 was excessive undervoting, a problem closely associated with punchcard voting, a technology that in most venues has been discredited.

In Florida's 13th Congressional District (hereinafter CD 13), Democrat Christine Jennings lost the election night canvass to Republican Vern Buchanan by only 377 votes. An initial recount indicated that Buchanan beat Jennings by 119,142 to 118,741 votes—a margin of 401 votes—and that 21,303 ballots cast in CD 13 recorded no House candidate.[1] These undervotes were heavily concentrated in Sarasota County, one of the five counties that contributes to CD 13; the other four such counties are Charlotte, DeSoto, Hardee, and Manatee. Within Sarasota County, CD 13 undervotes were concentrated among early and election day, as opposed to absentee, voters.[2]

The certified margin between Buchanan and Jennings eventually shrunk to 369 votes, and Jennings formally contested the CD 13 election outcome, arguing that, "The vote totals in the certification are wrong because they do not include thousands of legal votes that were cast in Sarasota County but not counted due to the pervasive malfunctioning of electronic voting machines."[3] The CD 13 undervote also motivated a lawsuit by Sarasota voters who claim that they were disenfran-

[1] See "Second Unofficial Returns" at `http://election.dos.state.fl.us/pdf/certCanvasCom.pdf`. These recount totals include a handful of overvotes.

[2] Pre-recount and in some cases certified canvasses can be downloaded from `http://www.charlottevotes.com/` (Charlotte County), `http://www.hardeecountyelections.com/` (Hardee County), `http://www.votemanatee.com/` (Manatee County), and `http://www.srqelections.com/` (Sarasota County). The pre-recount canvass from DeSoto County was faxed to the authors and is available from them.

[3] Complaint was downloaded from `http://www.heraldtribune.com/assets/pdf/SH81371120.PDF` on November 21, 2006.

chised by the failure of Sarasota County's touchscreen machines; plaintiffs in this lawsuit requested a re-vote.

The CD 13 legal complaints raise issues often associated with adoption of new technologies: while electronic voting purportedly reduces individual voter errors (a benefit), it does so at the risk of technology failure (a cost). Nonetheless, we show here that the reason for the high CD 13 undervote rate in Sarasota County almost certainly was a problematic ballot format and not a technology failure. The ballot format at the root of the CD 13 undervote led many Sarasota County voters to miss the CD 13 race while paging through their electronic ballots.

As we detail later, among the five CD 13 counties Sarasota County had a unique early and election day ballot format; and, only among Sarasota early and election day voters were exceptionally high CD 13 undervote rates observed. In addition, elevated early and election day Florida attorney general undervoting occurred in nearby Charlotte and Lee Counties where, for early and election day voting, the Charlotte and Lee ballot formats approximated the early and election day ballot format used in Sarasota County.

Beyond our demonstrating that the anomalous CD 13 undervote was caused by the ballot format used in Sarasota County, we can also say with very high confidence that many of the CD 13 undervotes cast in this county would have been valid votes had Sarasota early and election day voters been presented with the same voting machinery and ballot layouts used elsewhere in CD 13. Simply put, Christine Jennings would have been a member of the 110th Congress had Sarasota County early and election day voters not faced a problematic ballot format.

In Section 2 we provide background on the CD 13 race and present two competing explanations for the high undervote rate in this contest. In Section 3 we show that the cause of the undervote was a ballot format, and in Section 4 we show that the CD 13 outcome, Buchanan over Jennings by 369 votes, would almost certainly have been reversed in the absence of Sarasota undervotes. Section 5 concludes.

2 Background and Explanations for CD 13 Undervote

Florida has 25 Congressional Districts and 67 counties, and midway up the state's Gulf Coast sits Sarasota County. Sarasota and its 156 election day precincts lie completely within the 13th Congressional District which also includes parts of Charlotte County (eight election day precincts) and Manatee County (134 election day precincts, some of which are split between CDs 11 and 13) and all of DeSoto and Hardee Counties (15 and 12 election day precincts, respectively).[4] In November, 2006 the five CD 13 counties varied in their use of voting technologies: for early and election day voters, Charlotte and Sarasota used iVotronic touchscreens, manufactured by Election Systems & Software (ES&S), whereas DeSoto, Hardee, and Manatee used optical scan voting. All five counties used optical scan voting for absentee voters.

Two competing explanations have been offered for the anomalous (greater than 8%) undervote rate in CD 13 and in particular for Sarasota County's undervote rate in the CD 13 race. Deliberate election fraud is not among the common conjectures. Rather, attention has focused on voter confusion arising from a ballot format and voting machine malfunction.[5]

[4]What is typically called a "precinct" in the vernacular of United States elections is what we call an election day precinct, i.e., a place where voters physically go to vote. In contrast, one can also speak of "absentee precincts," which are collections of absentee voters albeit not physical voting locations.

[5]Immediately after the November, 2006 elections various observers speculated that CD 13 voters may have undervoted in the CD 13 race as a protest against the negativity that characterized the Buchanan versus Jennings contest—e.g., "Call for paper trail, new election," *Herald Tribune*, November 16, 2006. This allegation is not compelling, though, because the CD 13 undervotes were concentrated in a single county (Sarasota) as opposed to being distributed across all of CD 13. Moreover, Sarasota County is not in a unique media market; if it were, campaign negativity could conceivably have been concentrated there. Rather, Sarasota County is divided between the Tampa–Saint Petersburg–Sarasota DMA (which includes Sarasota, Manatee, and Hardee Counties) and the Fort Myers–Naples DMA (which includes Charlotte and DeSoto Counties). And, the set of broadcast television channels to which residents of CD 13 have access varies little across the five CD 13 counties. Of the 22 broadcast stations whose broadcast areas overlap CD 13, twelve can be seen by at least some of the voters in all five of the counties that constitute the district (cable providers generally carry local stations whose coverage area intersect the cable providers' service areas). At least one affiliate representing each of the four major networks (ABC, CBS, NBC, and FOX) reaches voters across all five counties in CD 13. Moreover, of the 21 stations that reach the center of Sarasota County, all but six also reach the center of the portion of Charlotte County that is contained in CD 13. Moreover, none of these six stations reach all of Sarasota County. Additional details on DMAs and CD 13 are available from the authors. In light of these details, we discount "protest undervotes" as a credible explanation for the CD 13 undervote.

2.1 Ballot Format

The claim that ballot format caused the large CD 13 undervote in Sarasota County is best understood by comparing the early and election day iVotronic ballot formats used in Sarasota and Charlotte Counties. When a voter confronts an iVotronic, her "ballot" consists of a sequence of pages of races and corresponding candidate choices. In both Sarasota and Charlotte Counties, the first ballot page contained a single race, the Florida United States Senate contest.

Beyond their first pages, the Sarasota and Charlotte County ballots differed. The second page of the Sarasota County ballot featured two races, the CD 13 race and the Florida gubernatorial contest. See Figure 1 for pictures of the first two Sarasota ballot pages.[6]

*** Figure 1 about here ***

In contrast, the Charlotte County iVotronic ballot featured a second page that contained the CD 13 race only; a third page with the Florida gubernatorial race along with the Florida attorney general race under it; and a fourth page with two Florida-wide races and a state legislative race. Figure 2 displays the first four pages of the Charlotte County ballot.[7]

*** Figure 2 about here ***

In various media accounts, Sarasota County voters complained about the ballot format they faced, and that "the touch-screen ballot design concealed the candidates for the 13th District."[8] In what sense might Buchanan and Jennings have been "concealed?" The most commonly cited mechanism that may have confused Sarasota voters involved the vertical stacking of the CD 13 contest (two candidates) above the gubernatorial contest (six candidates plus a write-in).[9]

The precise mechanism by which the stacking of two races, one with many candidates and one

[6]Source for the ballot pictures is personal communication on June 26, 2007 between the authors and Barbara Bain in the office of the Sarasota County Supervisor of Elections. The original ballot pictures received by the authors contained color, but the pictures are rendered here without color. In addition, the pictures received by the authors contained mouse pointers; these pointers were removed for clarity.

[7]Source for the Charlotte ballot pictures is electronic mail received on November 15, 2006 from the office of the Charlotte County Supervisor of Elections. The original Charlotte ballot pictures received by the authors contained green highlighting over the words "Next Page." This highlighting is rendered grey in Figure 2.

[8]"Elections office braces for 'intense' recount," *Herald Tribune*, November 10, 2006.

[9]See "Analysis suggests undervote caused by ballot design," *Herald Tribune*, November 15, 2006.

with two, may lead voters to miss the latter is beyond the scope of our analysis. Nonetheless, the extensive CD 13 undervote in Sarasota County is roughly consistent with Kimball (2007), who argues that full-face touchscreen voting machines, as compared to scrolling touchscreens, may overwhelm voters and lead to excessive numbers of residual votes (undervotes plus overvotes). Insofar as the second Sarasota ballot page contained two races whereas the first page contained only one, the second page was relatively overwhelming and cluttered. Janiszewski's (1998) experimental study of corneal movements shows that, the more cluttered a screen of information, the less respondents are able to learn from it.

Note as well the use of different colors—rendered in grey here—in Sarasota's second ballot page (Figure 1(b)) and in particular the grey surrounding the word "State." In the original Sarasota ballot, the word "State" was teal. If a voter's eyes were drawn to this swath of color, it seems quite plausible that this would lead the voter to look down and hence not be aware of the CD 13 race above the colored area.

We suspect that the extent to which a given ballot page is perceived, consciously or not, as overwhelming or cluttered depends on a comparison of the page to the ballot pages that preceded it. By this logic, the CD 13 undervote in Sarasota County reflects the use of a relatively cluttered and complicated ballot page that follows a much less cluttered page.

If this allegation holds, then Charlotte County's third ballot page should have generated excessive undervotes as well. We turn to this possibility shortly.

It is known that ballot formats are not neutral in terms of affecting voter behavior. Kimball & Kropf (2005) show that the way optical scan ballots are shaded and filled out (either with circles or arrows to be connected) affect vote choices; Wand, Shotts, Sekhon, Mebane, Jr., Herron & Brady (2001) show that the butterfly ballot format used in Palm Beach County, Florida, in November, 2000 altered the winner of the 2000 presidential election; and Koppell & Steen (2004) and Ho & Imai (2006) show that the order in which candidate names appear on ballots affects vote totals. Finally, Herrnson, Bederson, Niemi, Conrad, Hanmer & Traugott (2006) show that different

touchscreen formats strongly affect the error rates voters make when selecting candidates.

Early and election day voters in the remaining counties in CD 13, DeSoto, Hardee, and Manatee, voted with paper-based, optical scan ballots. There is no sense in which optical scan voting involves repetitive page-turning as is required by iVotronic touchscreens. An example Manatee ballot is pictured in Figure 3.[10] Absentee voters in all five CD 13 counties used optical scan ballots roughly similar to Manatee's election day ballot.

*** Figure 3 about here ***

2.2 Voting Machine Malfunction

An alternative explanation for the high Sarasota undervote rate in CD 13 is voting machine malfunction.[11] A malfunction could reflect software counting bugs that caused undervotes; touchscreen misalignment or calibration problems that caused voters to "miss" touching the names of candidates they intended to support; problems in touchscreen memory devices that caused corruption of voting data; and so forth. The list of possible machine failures is extensive, and see Wallach (2006) for a thorough discussion.

3 Evidence of Ballot Format-Induced Undervotes in Sarasota County and Beyond

Which of the two preceding explanations, ballot format or voting machine malfunction, explains the CD 13 undervote in Sarasota County? Our answer to this question is based on a series of plots that highlights the connection between ballot format and undervote rates. We focus first on CD 13 and then expand our analysis to two neighboring Congressional Districts.

[10] Source for the Manatee County ballot is `http://www.votemanatee.com/PollLocationPub.asp?PollingId=856`.

[11] See "Buchanan declared winner; rival Jennings sues," *Herald Tribune*, November 21, 2006.

Insofar as the plots that follow document the impact on undervote rates of an election administration practice, the conclusions that follow from the plots are germane to existing studies of undervoting and residual votes. Literature on residual votes has demonstrated that high residual vote rates are associated with particular kinds of voters and are often concentrated in low education and minority precincts (Brady, Buchler, Jarvis & McNulty 2001, Tomz & van Houweling 2003, Mebane 2004, Ansolabehere & Stewart III 2005, Herron & Sekhon 2005).

3.1 Undervoting in the Five CD 13 Counties

Figure 4 describes undervote rates for the five counties that contribute to CD 13.[12] The point in Figure 4 is simple: among CD 13 counties, Sarasota had the largest election day-absentee (and early-absentee) difference in undervote rates. In particular, on election day the CD 13 undervote rate in Sarasota County was approximately 0.139 (i.e., 13.9%, $\hat{\sigma} = 0.00116$ where $\hat{\sigma}$ denotes an estimated standard deviation), but Sarasota absentee voters had an undervote rate of approximately 0.0250 ($\hat{\sigma} = 0.00104$). This translates to a difference of 0.114 ($\hat{\sigma} = 0.00239$). Election day voters in Sarasota County were over five times as likely to cast an undervote in the CD 13 race than their absentee counterparts.

*** Figure 4 about here ***

Figure 4 cannot be explained by a generic proclivity of Sarasota election day voters to undervote. In 2004, Republican Katherine Harris defeated Democrat Jan Schneider by 55.3 percent to 44.7 percent. Although not pictured, a plot of election day, early, and absentee undervote rates in CD 13 in 2004 shows no dramatic differences between undervote rates in various counties and across different times of voting.[13]

[12] Figure 4 uses certified precinct returns except where they remain unavailable to the authors. As of this article's printing the publicly-available canvass for DeSoto County does not break down CD 13 undervote counts by time of voting, i,e., election day, early, and absentee; moreover, we do not have access to a certified canvass from DeSoto County. Finally, the Figure 4 does not include Manatee County precincts that were split between CD 11 and CD 13.

[13] A figure for 2004 can be found in a web appendix for this article. See `http://www.dartmouth.edu/~herron/cd13-webappendix.pdf`.

3.2 Undervoting in CD 13 Precincts

Figure 4 highlights a striking disparity between election day and absentee undervote rates within CD 13 counties, and we build on this finding by displaying precinct-level undervote rates in Figure 5. This figure plots for top races (two federal contests plus Florida cabinet races) election day undervote rates versus absentee undervote rates. The order of the races in the figure—United States Senate, CD 13, governor, attorney general, chief financial officer, and commissioner of agriculture—mirrors the official Florida order that is prescribed by law. Figure 4 could have been made with early and absentee undervote rates, and its results would have been qualitatively identical.

*** Figure 5 about here ***

For the Senate race, Figure 5(a) shows that all the CD 13 precincts were roughly similar insofar as they are clustered. Moreover, the precincts fall around a 45-degree line, and precincts on this line have identical election day and absentee undervote rates.

In contrast, there is a preponderance of Sarasota precincts, and only Sarasota precincts, above the 45-degree line in Figure 5(b). The location of the Sarasota cloud implies that Sarasota precincts had election day CD 13 undervote rates greater than corresponding absentee CD 13 undervote rates.

On the other hand, Sarasota County's dots are not anomalous in Figure 5(c), which describes undervoting in the Florida governor race. Indeed, this figure, in which all precincts are roughly clustered, looks very much like the Senate undervoting plot.

In the next race, attorney general, anomalous precincts again appear except the culprit now is Charlotte County. After the attorney general race, the chief financial officer and commissioner of agriculture plots have no anomalous features.

Figure 5, then, highlights two anomalous counties (Sarasota and Charlotte) in two different races (CD 13 and Florida attorney general, respectively). The Sarasota and Charlotte clouds of points in Figures 5(b) and 5(d) are quite similar, and the presence of these two clouds, in conjunction with the absence of other clouds, is consistent with the conjecture that, for touchscreen

counties, the first ballot page that combines races and follows single-race pages produces excessive undervotes.

It is worthwhile noting that there is a large literature on voter "roll-off" (e.g., Vanderleeuw & Utter 1993, Knack & Kropf 2003), and it shows that, the lower down on a ballot a race appears, the greater the number of residual votes expected. Because the CD 13 and attorney general races are at the tops of the Sarasota and Charlotte County ballots, the undervote patterns noted for these two races cannot be attributed to fatigue-induced undervotes.

3.3 Undervoting in CD 11 and CD 14 Precincts

Thus far, when we observe among CD 13 counties a potentially confusing ballot format (e.g., the Sarasota County format in the CD 13 race and Charlotte County format in the attorney general race) we also observe anomalous precinct-level undervoting; and, where ballot formats are not ostensibly confusing, we do not observe anything anomalous. We now continue to probe the connection between ballot formats and undervoting by turning to the top races in two neighboring Congressional Districts to CD 13, namely, Florida CDs 11 and 14.[14] See Table 1 for a description of election day and early voting technologies used in the counties that comprise CDs 11 and 14. All such counties used optical scan balloting for absentee voting.

Table 1: Election Day Voting Technologies Across Congressional Districts 11 and 14

District	County	Technology	Type	Notable Features
11	Hillsborough	Touchscreen	AVC Edge	Two column ballot
	Manatee	Optical Scan	Accuvote	Paper, multiple pages
	Pinellas	Touchscreen	AVC Edge	Two column ballot
14	Charlotte	Touchscreen	iVotronic	Governor, attorney general paired
	Collier	Touchscreen	iVotronic	Three cabinet races grouped
	Lee	Touchscreen	iVotronic	Governor, attorney general paired

[14]Our ignoring races that are far down various ballots (non-top races) is intentional. There are, we suspect, many sources of variance across counties and ballot styles that affect down-ballot voter participation rates, and these will confound any analysis of undervoting in local races, among others. Hence we focus solely on top races common across all Florida counties.

If ballot format explains the CD 13 undervote in Sarasota County, then by extension Table 1 has several implications. First, Charlotte and Lee Counties should have anomalous undervotes in the attorney general race and nowhere else. Second, we should not observe anomalous undervote rates in Hillsborough, Manatee, and Pinellas Counties. And, third, Collier County may have anomalous undervoting insofar as it combines on a single page three Florida races. This grouping is different than the Charlotte and Lee County governor-attorney general grouping because the Collier grouping of three races does not have any major asymmetries in candidate counts. Recall that the Florida governor race had seven candidate choices whereas only two candidates each contested the attorney general, chief financial officer, and commissioner of agriculture races.

In light of these implications, Figure 6 is a collection of plots for CDs 11 and 14 that parallel the earlier plots for CD 13 precincts. Namely, each component of Figure 6 describes the election day and absentee undervote rates for a large set of precincts; precincts that are close to pictured 45-degree lines have election day and absentee undervote rates that are roughly similar; and, clouds of points either above or below a 45-degree line indicate anomalous undervoting.

*** Figure 6 about here ***

We see from Figure 6(a) that among CD 11 and CD 14 precincts there was nothing systematically anomalous about undervoting in the United States Senate contest. For the CD 11 and CD 14 contests themselves—see Figure 6(b)—a similar conclusion applied. Neither the CD 11 nor the CD 14 race was particularly competitive, and thus one cannot compare absolute undervote rates from these two races to the undervote rate from the extremely competitive CD 13 race.

Figure 6(c) highlights the absence of anomalous undervote rates in the Florida governor race. But, Figure 6(d) is different: note that the Charlotte and Lee County precincts are elevated well above the pertinent 45-degree line. Such a Charlotte cloud appeared before (see Figure 5(d)), but its consistency with Lee County is noteworthy because Lee County is not part of CD 13. After the attorney general race, Figures 6(e) and 6(f) do not contain any anomalous clouds of precincts.

Figure 6 is consistent with our conjectures about confusing ballot formats. Charlotte and Lee

Counties were anomalous in the attorney general race and nowhere else. And, just as importantly, no other CD 11 or CD 14 counties suffered from anomalous undervoting in any of the top races in November, 2006.

With respect to Collier County, there is a hint of evidence in Figure 6 that this county had elevated election day undervote rates in the chief financial officer and commissioner of agriculture races. Close inspection of Figures 6(e) and 6(f) shows that, in general, Collier County's precincts are above pictured 45-degree lines for these two races and that this pattern does not appear in the attorney general race. While our evidence for elevated undervoting in Collier County is limited, this is consistent with the fact that Collier County's ballot format should not have been nearly as confusing as the ballot formats used in Charlotte and Lee Counties.

We have estimated a variety of precinct-level regression models that cover CDs 11, 13, and 14, and in these models election day minus absentee undervote rates are regressed against various precinct demographics and other features.[15] Even controlling for precinct characteristics, the Charlotte, Lee, and Sarasota undervote anomalies remain and, notably, no other anomalies appear.

3.4 Summary of Evidence

We have now shown that a conjectured mechanism that induces high undervote rates—an ostensibly confusing ballot format—is present in our primary place of interest (Sarasota County); is present in other places with high undervote rates (Charlotte and Lee Counties); and is absent in places without high undervote rates. That is, county-race combinations with our mechanism have high election day-absentee undervote disparities, and county-race combinations without the mechanism lack corresponding disparities. This is compelling evidence in favor of the conjecture that the CD 13 undervote in Sarasota County was the result of a ballot format effect.

If one were to take seriously the notion that voter indifference induced by campaign negativity drove the Sarasota CD 13 undervote, one would have to argue that indifference over Florida

[15] These results are available in the article's web appendix. See fn. 13.

attorney general candidates affected voters in Charlotte and Lee Counties. If Florida voters were indifferent between attorney general candidates, why would they act on this indifference in two counties only and why on election day but not during absentee voting?

Similarly, the evidence we have adduced makes a technology malfunction explanation for the Sarasota CD 13 undervote difficult to sustain. Simply put, any explanation for the CD 13 undervote in Sarasota must be capable of explaining the attorney general undervotes in Charlotte and Lee Counties and the lack of an attorney general undervote elsewhere (not to mention the lack of high undervotes in other races). If Sarasota experienced a voting machine malfunction that caused Congressional race undervotes (e.g., Dill & Wallach 2007, Mebane & Dill 2007), then this same malfunction appears to have affected two other neighboring counties but only in one particular race and not in the same race that the malfunction appeared in Sarasota. To the extent that ballot format explains the observed the Sarasota CD 13 undervote, the conjecture that engineering flaws are responsible for the undervote becomes more difficult to sustain.[16]

4 Consequences of the Sarasota Undervote in CD 13

What would have happened if the Sarasota County undervotes cast in the CD 13 race had followed patterns similar to those in the rest of CD 13?[17] This question turns on the issue of reallocating Sarasota undervotes between the CD 13 candidates, Buchanan and Jennings, and intentional undervotes. We consider undervote reallocation in two ways, first with precinct returns and second with ballot data.

[16]The possibility of hardware and/or software problems among Sarasota County's voting machines led the Florida Division of Elections to devise a testing procedure with the intention of determining if engineering problems were the sole cause or a contributing factor to the CD 13 undervote in Sarasota County. See "Parallel Test" design at `http://election.dos.state.fl.us/pdf/parallelTestPlan.pdf`. The initial phase of the voting machine audit did not uncover any significant software or hardware issues with Sarasota County's voting machines (e.g., "State says all discrepancies in test election due to 'human error'," *Herald Tribune*, November 30, 2006) and this conclusion was sustained in the final software analysis (see `http://election.dos.state.fl.us/pdf/FinalAudRepSAIT.pdf`).

[17]A similar question could be asked about attorney general undervotes cast in Charlotte and Lee Counties, but this question is not particularly noteworthy because the Florida attorney general race was not close, unlike the CD 13 race.

4.1 Undervote Allocation with Precinct Returns

Our initial approach to Sarasota undervote reallocation relies on precinct-level election returns from Charlotte, Hardee, Manatee, and Sarasota Counties. As explained below, we use observed correlations between election results in CD 13 precincts outside Sarasota County to estimate CD 13 election results within Sarasota precincts.

Our precinct-level model for allocating Sarasota undervotes makes two key assumptions:

- *The relationships between races are the same across all precincts in CD 13.* That is, we assume that our ability to predict the outcome of one race (e.g., Florida attorney general) given knowledge of a set of other race outcomes applies to precincts inside Sarasota County in the same way that it applies to precincts outside of it.

- *Both CD 13 candidates, Buchanan and Jennings, would not have received fewer votes had the Sarasota ballot format problem not existed.* That is, the ballot format effect we have identified did not add votes to either Buchanan or Jennings but only changed intended votes into undervotes.

From these assumptions we derive a statistical model for vote shares across precincts in CD 13 assuming that such shares as well as the proportion of voters intentionally undervoting are additive-logistic-normally distributed as in Katz & King (1999). Once we specify a model for vote outcomes at the precinct level, we then treat CD 13 outcomes in Sarasota as missing and use multiple imputation to estimate what these outcomes would have been had Sarasota voters behaved like other voters in CD 13. Full technical details are in Appendix A.

Our imputation model generates a probability density over precinct voting outcomes and thus over election outcomes. We take one thousand random draws from this density and calculate one thousand predicted election outcomes for the CD 13 race (recall, our model treats CD 13 vote totals in Sarasota precincts as missing data). From these random draws we can answer key questions

such as, what fraction of Sarasota undervotes were intended to have been cast for Buchanan and Jennings, and, how might these intended votes have influenced the election outcome if recorded?

Our imputation results are summarized in Table 2, which provides estimated election results among early and election day voters in Sarasota County. Key here is how many more votes Buchanan and Jennings would have received if Sarasota undervotes had not been influenced by ballot design. This is expressed as the *pickup* for each candidate and for undervotes as well. The undervote pickup is negative because, according to our model, most undervotes were intended to have cast for Buchanan or Jennings. Vote totals and pickups reported in Table 2 are expected values, and associated 95% confidence intervals provide measures of uncertainty.

Table 2: Summary of Allocation Results

EARLY VOTING:			
	Jennings	Buchanan	Undervote
Estimated Vote:	16944	12939	830
Estimated Pickup:	2477	2106	-4582
Lower bound of 95% Confidence Interval:	2159	1811	-4750
Upper bound of 95% Confidence Interval:	2779	2400	-4401
ELECTION DAY VOTING:			
	Jennings	Buchanan	Undervote
Estimated Vote:	45387	40774	2511
Estimated Pickup:	5541	4298	-9840
Lower bound of 95% Confidence Interval:	5019	3837	-10165
Upper bound of 95% Confidence Interval:	6064	4811	-9509
EARLY + ELECTION DAY VOTING:			
	Jennings	Buchanan	Undervote
Election Returns:	54313	47309	17763
Estimated Totals:	62331	53713	3341
Projected Pickup:	8018	6404	-14422
Lower bound of 95% Confidence Interval:	7377	5853	-14792
Upper bound of 95% Confidence Interval:	8612	6969	-14040
Predicted probability Jennings pickup $>$ 369: 0.984			

Looking first at early voting, the confidence interval for the undervoting column is -4,750 to -4,401; therefore, we are 95 percent confident that there would have been between 4,750 and 4,401 fewer early voting undervotes if there had been no ballot design problem in Sarasota County. We estimate that 2,477 voters intended to vote for Jennings and 2,106 voters intended to vote for Buchanan in early voting.

With respect to election day voting, we are 95 percent confident that Jennings would gain between 5,019 and 6,064 more votes in Sarasota had there been no ballot design issues. Similarly, we are 95 percent confident that between 3,837 and 4,811 voters intending to vote for Buchanan were instead counted as undervotes.

Looking to combined results, we are 95 percent confident that *between 14,040 and 14,792 voters in Sarasota County intended to have cast a valid vote in the CD 13 race but did not*. Of these more than 14,000 votes, we estimate that Jennings would have received 8,018 and Buchanan, 6,404. This would cause Jennings to gain 1,614 votes over Buchanan thus overcoming her 369 vote deficit and changing the outcome of the CD 13 race.

Moreover, from our simulated election results we calculated the total number of votes gained by Jennings over Buchanan. When this number is greater than 369, then Jennings would have won the election; when this number is less than 369 (and is possibly negative), Jennings would have lost. Crucially, we are interested in the fraction of the time the pickup for Jennings is greater than 369. This probability, according to our precinct-level reallocation model, is 98.4 percent.

4.2 Undervote Allocation with Ballot Records

Another way to allocate Sarasota County's CD 13 early and election undervotes to Buchanan and Jennings is to compare patterns of undervotes across individual Charlotte County and Sarasota County ballots. Here, that is, we use individual-level ballot records to reallocate undervotes, whereas previously we used aggregate, precinct-level data. Intuitively speaking, we consider what would have happened if each individual Sarasota County voter had voted in Charlotte County

using Charlotte's non-confusing ballot format. For this analysis we set aside the governor and attorney general races because, as we have seen, Charlotte County had a confusing ballot format with respect to the latter.

To estimate the probability that a given Charlotte voter would vote in the CD 13 race given her votes and undervotes for other offices (specifically, United States Senate, chief financial officer, commissioner of agriculture, Amendment 8, and the retention of Supreme Court Justice Lewis), we use the logit models in Table 3, column one. Because only 7,326 of Charlotte's election day and early voters lived in CD 13 and only 174 of those undervoted, we cannot confidently estimate the probability of participation in the CD 13 race conditional on voting in other races in a fully non-parametric way. However, by including a number of interaction terms in our logit specification, we allow for a good deal of flexibility. As a robustness check, several non-parametric allocations are provided in Appendix B. For reasons that will be become clear shortly, we also require estimates of the probability that each voter in Sarasota County votes or undervotes in the CD 13 race given her other voting decisions—see Table 3, second column.

To estimate how Sarasota CD 13 undervotes that should have been valid votes would have been divided between Buchanan and Jennings, we assume that, conditional on votes in other races on the ballot, those Sarasota voters undervoting in the CD 13 race were no more or less likely to support Buchanan or Jennings than those who did not undervote in this race. Under this assumption, we can estimate the probability that any Sarasota voter would support Buchanan or Jennings given her support for other candidates on the ballot by analyzing the Sarasota ballots which contain a valid CD 13 vote. For this estimation problem, we employ the logit model in Table 3, column three. Here we model the probability of supporting Jennings as a function of variables indicating whether Democrats were supported in the given races (or whether Amendment 8 was opposed).

We pose the following counterfactual: what if each early and election day Sarasota voter had voted using Charlotte's machines and ballot formats? We assume the model governing Charlotte

Table 3: Logit Models Predicting Voting and Candidate Choice in CD 13

Variable	Charlotte (vote=1)	Sarasota (vote=1)	Sarasota (Dem. vote|vote=1)
US Senate	2.00	0.55	2.45
	(0.41)	(0.09)	(0.03)
Chief Financial Officer	1.24	0.26	2.23
	(0.53)	(0.12)	(0.04)
Agriculture Commissioner	1.17	0.97	1.89
	(0.25)	(0.04)	(0.06)
Amendment 8	0.38	0.16	0.38
	(0.33)	(0.04)	(0.04)
Supreme Court: Lewis	-0.15	-0.10	
	(0.49)	(0.05)	
Amendment 8 (No) $\times$ Supreme Court	0.22	0.27	
	(0.53)	(0.06)	
US Senate $\times$ CFO	0.23	0.53	-0.23
	(0.56)	(0.12)	(0.05)
US Senate $\times$ Comm. Agric.			-0.44
			(0.07)
US Senate $\times$ Amendment 8			-0.04
			(0.05)
Comm. Agric. $\times$ Amendment 8			-0.29
			(0.06)
Intercept	-0.86	-0.70	-2.71
	(0.41)	(0.09)	(0.02)
N	7,326	119,898	102,073
Log likelihood	-704.43	-483,329	-29,833
Geometric Mean Probability	0.91	0.67	0.75

voters' undervoting would apply to Sarasota voters if they had voted in Charlotte.[18] Thus, for any voter i we can consider the event C_i in which the voter undervotes using a Charlotte machine and the event S_i in which the voter undervotes using a Sarasota machine. We assume that any voter who undervotes in Charlotte would undervote if voting in Sarasota; formally, $S_i \cap C_i = C_i$. To allocate Sarasota undervotes, we require estimates of the probability of C_i given S_i. That is, we need to know the probability that each Sarasota undervoter would continue to undervote if using

[18] Among other things this assumption requires that there were no other ballot format issues in Sarasota or Charlotte that we are not accounting for.

Charlotte machines. By definition, $\mathrm{P}\left(\mathrm{C_i}|\mathrm{S_i}\right) = \mathrm{P}\left(\mathrm{C_i} \cap \mathrm{S_i}\right)/\mathrm{P}\left(\mathrm{S_i}\right)$, which in this case is simply $\mathrm{P}\left(\mathrm{C_i}\right)/\mathrm{P}\left(\mathrm{S_i}\right)$. Estimates of $\mathrm{P}\left(\mathrm{S_i}\right)$ and $\mathrm{P}\left(\mathrm{C_i}\right)$ are obtained from the logit regressions described above. Given these estimates of $\mathrm{P}\left(\mathrm{C_i}|\mathrm{S_i}\right)$ and the probability of voting for Buchanan or Jennings conditional on voting at all, it is a simple matter of multiplication to assign to each Sarasota CD 13 undervoter a probability of continuing to undervote, supporting Buchanan, or supporting Jennings. Summing these probabilities across all Sarasota CD 13 undervoters yields the reallocation in Table 4. Confidence intervals were obtained via a bootstrap.

Table 4: Summary of Allocation Results from Ballot Data

	Jennings	Buchanan	Undervote
Election returns	54313	47567	17825
Estimated Vote:	63261	53420	3024
Estimated Pickup:	8948	5853	-14801
Lower bound of 95% Confidence Interval:	8710	5635	-14347
Upper bound of 95% Confidence Interval:	9162	6059	-15266
Predicted probability of Jennings's pickup $>$ 369: ≈ 1			

As compared to our precinct-level results, here we find stronger evidence that Jennings would have won the CD 13 election if Sarasota had used the same machines as were used in Charlotte County. Our individual ballot data reallocation suggests the Sarasota undervotes would have produced a net swing of between 2,856 and 3,297 votes in favor of Jennings with a point estimate of 3,086. This is roughly eighty percent more votes than is implied by our precinct-level reallocation model. Note the precinct-level confidence interval was plus or minus roughly 1111 votes. Here the confidence interval is only plus or minus 221 votes. Thus, through the use of ballot-level data, we improved the precision of our estimates five-fold.

5 Conclusion

The motivation for this article is the unusually high undervote rate in the November, 2006 13th Congressional District race in Florida. This race featured Republican Vern Buchanan versus Democratic Christine Jennings, Buchanan was declared the winner by 369 votes, and he was sworn into the 110th Congress in January, 2007. As of the date of this article's printing, challenges to Buchanan's election remain pending in the courts and in the United States House of Representatives.

We have offered two conclusions regarding the plethora of Sarasota County undervotes in the 13th Congressional District race, the first pertaining to cause and the second to consequences. With respect to the former, we have argued that the cause of the Sarasota undervote was a confusing ballot format. Our evidence for this claim draws on variance in ballot formats across counties and types of voters (early, election day, and absentee). In particular, we highlighted the way that the touchscreen voting machines in Sarasota County grouped races and in particular grouped a Congressional race with the Florida gubernatorial race. The problematic grouping that Sarasota employed—having a two-candidate United States House race on the same touchscreen page as the Florida gubernatorial race—was employed by other counties, albeit with different races, and in these other counties and associated races we find anomalous undervoting. Moreover, where a confusing grouping of races is absent, anomalous undervoting is absent as well.

Thus, what looks on the surface like a unique Sarasota County problem or an issue idiosyncratic to Florida's 13th Congressional District that appeared in November, 2006 is in fact more general. The ballot format effect we have identified appears generic, and this finding should contribute to research on optimal ballot design.

With respect to our second conclusion, like the butterfly ballot used in Palm Beach County in the 2000 presidential race the Sarasota ballot format in the 2006 midterm elections was pivotal to the winner of tightly-contested 13th Congressional District race. We estimate that, had Sarasota

County used a ballot format akin to those in neighboring counties, Jennings would have beaten Buchanan. Indeed, our ballot-level analysis leaves virtually no doubt that the excess Sarasota undervote would have broken in Jennings's favor in a manner that would have easily reversed the certified election outcome.

There remain two key issues over which we do not have leverage. The first is the precise reason as to why grouping races on touchscreens is a problem. Is the issue asymmetry between races, i.e., a race with two candidates grouped with a race with multiple candidates? Or, is it total number of candidates on a page? Or, is the issue vertical stacking of races versus horizontal grouping? Or, is there a tradeoff between grouping races and the number of pages in a ballot wherein greater number of pages itself leads to undervotes? We have offered a limited theoretical discussion of these issues, and work in the vein of Herrnson, Abbe, Francia, Bederson, Lee, Sherman, Conrad, Niemi & Traugott (2005), Herrnson et al. (2006), and Herron & Lewis (2006) may help sort out the missing links.

Second, and as mentioned previously on several occasions, because we have presented a statistical analysis of vote patterns and not a physical examination of voting machines, we cannot completely rule out voting machine malfunction as a source of the Sarasota undervote. Is it technically possible that software or hardware glitches are responsible for the Sarasota Congressional race undervote, the Charlotte and Lee attorney general undervotes, the apparent ballot format effects in Collier County, and so forth? Yes, we believe that this is possible. However, if the issue is a software or hardware glitch, the glitch would have had to have manifested itself among almost all voting machines in Sarasota County but affected only one race and then have manifested itself in almost all machines in Charlotte and Lee Counties yet affected a different race.

We conclude with what we believe is a clear implication of our results: touchscreen ballot formats can affect vote counts in a dramatic way and, as the Florida CD 13 contest illustrates, they can be pivotal to election outcomes. While we have identified one problematic format, we have not offered an assessment of what is the best format for a typical voter. And furthermore, we have

said nothing about problematic formats in other voting technologies, none of which, one would imagine, is neutral. The matter of what the best format might be remains a subject of ongoing research.

Appendix A Precinct-level CD13 undervote allocations

This appendix explains the procedures we used to allocate Sarasota County's election day and early voting CD 13 undervotes with precinct-level voting returns.

A.1 Compositional Transformations

For the various counties considered here we know how many votes were cast in each precinct for each ballot choice (Republican, Democratic, or Undervote) and by each voting method (early voting, election day voting, and absentee voting).[19] Let $V_{i,c}^{e,m}$ represent the vote share in precinct i, race e, by method $m \in \{\text{early}, \text{e.d.}, \text{abs.}\}$, for choice $c \in \{D, R, U\}$.

For any given race and voting method (that is, temporarily ignoring superscripts), individual vote shares are constrained to the simplex,

$$V_{i,c} \in [0, 1] \qquad \forall\, i, c, \tag{1}$$

and the set of votes in a precinct across the three choices sums to unity,

$$V_{i,R} + V_{i,D} + V_{i,U} = 1 \qquad \forall i. \tag{2}$$

The space of each vector V_i is therefore the three dimensional simplex. For compositional data in a J-dimensional simplex, the transformation of Aitchison (1986) creates a set of $J - 1$ log ratios

[19] We discard votes for the minor candidates who contested the gubernatorial race and for all write-ins, and we also discard overvotes. All of the discarded votes are negligible totals in the races studied here.

each of which compare the vote of one party to that of a baseline or reference party. Without loss of generality we use the Democratic party as our reference choice, and this yields two transformations:

$$Y_{i,RD}^{e,m} = ln\left(\frac{V_{i,R}^{e,m}}{V_{i,D}^{e,m}}\right) \quad (3)$$

$$Y_{i,UD}^{e,m} = ln\left(\frac{V_{i,U}^{e,m}}{V_{i,D}^{e,m}}\right) \quad (4)$$

The set of log vote ratios Y are now individually and collectively unconstrained. Examination of such ratios in other contexts has found them to be well fitted by a multivariate t or multivariate normal distribution (Katz & King 1999, Jackson 2002, Tomz, Tucker & Wittenberg 2002). Thus. our key modeling assumption is that collectively Y_i are joint multivariate normal across all relevant c, e, and m. The reverse transformation from Y to V implies that the vote shares themselves are distributed additive-logistic-normal.

A.2 Imputation

Define S as a dichotomous indicator that is one in Sarasota county and zero elsewhere; let V and Y be observed votes shares and transformations as set out above; and, let $\overset{*}{V}$ and $\overset{*}{Y}$ be the latent vote shares and transformations that would have been observed if there were no ballot flaw in that race. Clearly, $\overset{*}{Y}_i^{CD13} = Y_i^{CD13}$, $\forall s_i = 0$, but elsewhere, $\overset{*}{Y}_i^{CD13}$ is unobserved. However, many other races and election methods are observed in Sarasota County and its neighbors.

Table 5 lists the twelve election methods we include in our imputation model. The vote shares from each method are transformed into two log vote ratios, the ratio of the Republican to Democratic vote and the ratio of the undervote to the Democratic party vote share.[20] We omit the gu-

[20] In the judicial retention and amendment returns, the variables are the log ratio of no to yes votes, and the log ratio of undervotes to yes votes. Early returns in CD 13 are predicted with early returns from the other elections, and the two Absentee variables. Election day returns are predicted with election day and absentee variables.

Race	Methods
Congressional District 13	Absentee
United States Senate	Early, Election Day, Absentee
Agricultural Commissioner	Early, Election Day
Chief Financial Officer	Early, Election Day
Lewis Supreme Court Retention	Early, Election Day
Amendment 8	Early, Election Day

Table 5: Races and Voting Methods used in Imputation Model

bernatorial race because it was the race that shared a ballot screen with the CD 13 race in Sarasota County. Absentee votes for the CD 13 race are included as forecasting variables as these ballots are paper. We include all precincts in CD 13 from Charlotte, Hardee, Manatee, and Sarasota Counties except the small number of precincts in Manatee that were split between the CD 11 and CD 13 races. As a point of notation, from this point onward we refer to the log vote ratios in the early and election day voting in the CD 13 race as Y's and the log ratios of all these other races in Table 5 as X's although the latter are still constructed by equation 3.

If we consider $\overset{*}{Y} = Y$ for all $s_i = 0$ and all observations of $\overset{*}{Y} = Y$ within Sarasota as completely missing data (with observed covariates X), then our model has the same architecture as any conventional multivariate normal imputation model (Rubin 1987, Schafer 1997, King, Honaker, Joseph & Scheve 2001). We can estimate the posterior distribution of the missing values and draw imputations from this distribution to create fully observed datasets from which it is straightforward to create quantities of interest such as the vote totals of each candidate.

However, as there are only two patterns of missingness in our data, the critical complication of imputation algorithms, running large numbers of simultaneous equations, can be avoided.[21] Ignoring covariance between early and election day returns, within the CD 13 race the imputation

[21] All observations are either fully observed or are missing the four log ratios of early and election day voting in the CD 13 race.

model is simply two sets of bivariate normal regressions:

$$\begin{aligned}(\overset{*}{Y}_{RD}, \overset{*}{Y}_{UD}) &\sim f_{\text{bivariate normal}}(\mu_{RD}, \mu_{UD}, \Sigma) \\ \mu_{RD} &= \mathbf{X}\beta \\ \mu_{UD} &= \mathbf{X}\gamma, \end{aligned} \tag{5}$$

where $\mathbf{X}$ consists of 14 log vote ratios from the elections in Table 5 plus a constant vector. Imputations of $\overset{*}{Y}$ from this model yields completely observed data. However, given the simplicity of the patterns of missingness in our model, we can elaborate on the conventional multivariate normal model to include vote shares within Sarasota as censored observations.

A.3 Constraints

Observed CD 13 vote shares in Sarasota County contain some information about latent values. If the Sarasota CD 13 ballot format had been equivalent to the ballot format used in other CD 13 counties, then it is reasonable to assume that some Sarasota ballots which contain CD 13 undervotes would have registered a vote for a candidate while no votes successfully cast for either candidate would change. Thus, in Sarasota County the vote shares for both Buchanan and Jennings must be strictly increasing in undervotes and the share of undervotes correspondingly decreasing. Moreover, there are upper bounds for how much Buchanan and Jennings vote shares could change if all Sarasota CD 13 undervotes were allocated to these two candidates. Thus, the observed vote shares in Sarasota give us a series of bounds on the latent vote shares:

$$\overset{*}{V}_{i,U} \leq V_{i,U} \qquad \forall i : s_i = 1 \tag{6}$$

$$(V_{i,R} + V_{i,U}) \geq \overset{*}{V}_{i,R} \geq V_{i,R} \qquad \forall i : s_i = 1 \tag{7}$$

$$(V_{i,R} + V_{i,U}) \geq \overset{*}{V}_{i,D} \geq V_{i,D} \qquad \forall i : s_i = 1 \tag{8}$$

These bounds on $\overset{*}{V}$ imply a set of bounds on $\overset{*}{Y}$, the most straightforward of which is:

$$\overset{*}{Y}_{i,UD} \leq Y_{i,UD} \tag{9}$$

Additionally, if we knew the true undervote, $\overset{*}{V}_{i,U}$, we could define the functions:

$$Y^{+}_{i,RD}(\overset{*}{V}_{i,U}) = ln\left(\frac{V_{i,R} + (V_{i,U} - \overset{*}{V}_{i,U})}{V_{i,D}}\right) \tag{10}$$

$$Y^{-}_{i,RD}(\overset{*}{V}_{i,U}) = ln\left(\frac{V_{i,R}}{V_{i,D} + (V_{i,U} - \overset{*}{V}_{i,U})}\right), \tag{11}$$

which provide bounds on

$$Y^{+}_{i,RD}(\overset{*}{V}_{i,U}) \geq \overset{*}{Y}_{i,RD} \geq Y^{-}_{i,RD}(\overset{*}{V}_{i,U}). \tag{12}$$

We simplify these functions to their limiting values as $\overset{*}{V}_{i,U} \rightarrow 0$

$$Y^{+}_{i,RD} = ln\left(\frac{V_{i,R} + V_{i,U}}{V_{i,D}}\right) \tag{13}$$

$$Y^{-}_{i,RD} = ln\left(\frac{V_{i,R}}{V_{i,D} + V_{i,U}}\right) \tag{14}$$

Using equation 9 and the simplified form of equation 12 we can set the limits of integration for:

$$L(\beta, \gamma, \Sigma | S_i = 1) = \int_{Y^{-}_{i,RD}}^{Y^{+}_{i,RD}} \int_{-\infty}^{Y_{i,UD}} p_{\text{bvn}}(r, s | \mathbf{X}_i\beta, \mathbf{X}_i\gamma, \Sigma) \delta s \; \delta r \tag{15}$$

while the precincts outside Sarasota more straightforwardly contribute:

$$L(\beta, \gamma, \Sigma | S_i = 0) = p_{\text{bvn}}(Y_{i,RD}, Y_{i,UD} | \mathbf{X}_i\beta, \mathbf{X}_i\gamma, \Sigma) \tag{16}$$

A.4 Rejection Sampling

We parametrically bootstrap the parameters from our imputation model. From each bootstrapped set of parameters we create one imputed dataset where all the election outcomes are the same as observed values except that early and election day CD 13 vote shares in Sarasota County precincts are draws from their posterior distributions, conditional on other observed elections in those precincts. Although five or ten imputed datasets is sufficient in most analyses, we want to create confidence intervals of some quantities and so impute 1,000 datasets.

Our imputation model as previously discussed is multivariate normal in the space of the Ys. The quantities of interest, however, are the vote totals for each candidate. This requires transforming imputed log vote ratios back to vote shares and then multiplying these vote shares by the total turnout in each precinct. The reverse transformations are:

$$\widetilde{V}_{i,U} = \exp(\widetilde{Y}_{i,UD})/W_i, \quad \widetilde{V}_{i,D} = 1/W_i, \quad \widetilde{V}_{i,R} = \exp(\widetilde{Y}_{i,RD})/W_i, \tag{17}$$

$$\text{where} \quad W_i = 1 + \exp(\widetilde{Y}_{i,RD}) + \exp(\widetilde{Y}_{i,RD}). \tag{18}$$

Our imputed values are drawn from an untruncated conditional posterior yet we want values conditional on V. Therefore we need draws from a truncated distribution that obeys $\widetilde{V}_{i,U} \leq V_{i,U}$, $\widetilde{V}_{i,D} \geq V_{i,D}$, and $\widetilde{V}_{i,R} \geq V_{i,R}$. Following Honaker, Katz & King (2002), we rejection sample each vector $\widetilde{V}_i$ until every observation passes all constraints. This rejection sampling needs to be done on the imputations regardless of whether the CD 13 returns in Sarasota County are treated as entirely missing or censored values.

Appendix B Non-parametric CD 13 Undervote Allocations from Ballot Records

In Section 4 we provide a ballot-by-ballot allocation of the Sarasota CD 13 undervote based on a flexible logit model that conditions on votes cast for five other offices. An alternative to this approach is a fully non-parametric model in which undervoters in Sarasota CD 13 are matched to counterparts in Charlotte who cast the same (or nearly the same) pattern of votes across a given set of other races.

Such an approach has both advantages and disadvantages. The principal advantage is that it allows us to base our allocations on only those Charlotte voters who otherwise voted in exactly or almost exactly the same way as a given Sarasota CD 13 undervoter. However, as the number of races on which voters are matched increases, the number of times that any given pattern of votes is observed declines and it becomes more difficult to predict how a voter casting any given pattern of other votes would vote in the CD 13 race. In our application, this "curse of dimensionality" creates not only uncertainty but potentially bias.

Notwithstanding this potential bias, in Table 6 we present allocations of the Sarasota CD 13 undervote based on matching ballots of undervoters in Sarasota to ballots from Charlotte County CD 13 precincts. We provide three separate allocations. The first matches on the same offices considered in the logit model presented in the text—see Table 3; the second considers only the top four offices; and, the third matches on all 18 common offices. Bootstrapped confidence intervals are provided for each allocation. Where possible, exact matches for each of the 17,825 Sarasota CD 13 (early and election day) undervotes are found. Where exact matches are not found, best matches are established via a simple scoring algorithm that assigns precedence to ballots that match undervote choices and which prefers matching votes for different minor-party candidates to matching votes for different major-party candidates.[22] We find that the parameters of the matching algorithm

[22]The computer code that implements the matching and allocation is available from the authors by request.

Non-parametric Reallocation of Sarasota CD 13 Undervote

Offices matched	Undervotes	Buchanan	Jennings
{U.S. Sen., AG, CFO, Agric, Amd1, JSCLewis}			
Exact matches	1963	5525	9015
	(1916, 2007)	(5459, 5605)	(8920, 9106)
Approximate matches	273	344	447
	(251, 297)	(320, 370)	(419, 477)
Total	2235	5869	9463
	(2190, 2284)	(5784, 5956)	(9369, 9551)
{U.S. Sen., AG, CFO, Agric}			
Exact matches	2440	5677	9141
	(2396, 2479)	(5599, 5743)	(9058,9236)
Approximate matches	100	63	86
	(86, 112)	(53, 74)	(75, 98)
Total	2540	5739	9228
	(2498, 2581)	(5665, 5805)	(9144, 9316)
{All 18 common offices}			
Exact matches	477	3208	5508
Approximate matches	607	3298	4617
Total	1084	6506	10126

Table 6: *Estimates based on CD 13 vote choices of Charlotte voters casting the same pattern of votes across the specified races as Sarasota voters who undervoted in CD 13. Ninety-percent confidence intervals, calculated using a non-parametric bootstrap, are given in parentheses. Adjustments to the estimated undervote probability within each vote pattern are required and are made as described in the text. Amd1 refers to the first Amendment on the November, 2006 ballot, and JSCLewis refers to the retention vote for Florida Supreme Court Justice R. Fred Lewis.*

had little effect on the allocation. The approximate matches are enumerated separately from the exact matches in Table 6. Except in the case in which matching is attempted across all 18 common contests, the number of ballots for which approximate matching is applied is relatively small.

Overall, the non-parametric allocations are very similar to what we found with our semi-parametric logit estimator. Nonetheless, all of the non-parametric allocations find more votes for both Buchanan and for Jennings than the model presented in the text and find also a larger net advantage for Jennings. Differences are on the order of a few hundred votes.

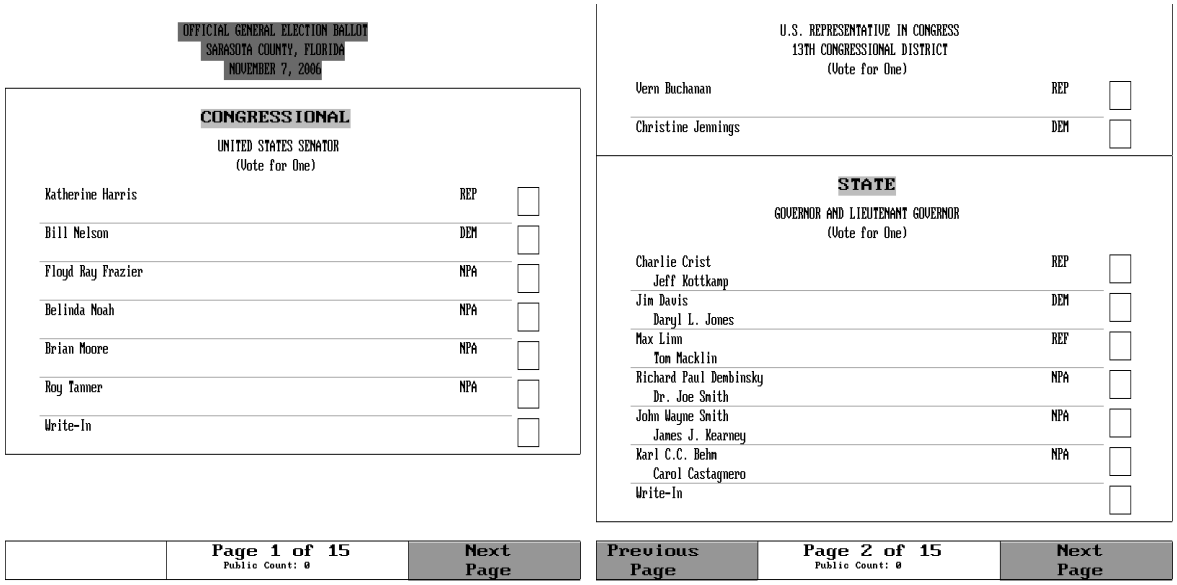
The "curse of dimensionality" problem associated with our non-parametric estimates is revealed in the results by the decreasing number of Sarasota CD 13 undervote ballots that are predicted to remain undervotes as the number of races on which voters are matched is increased. Matching on only four offices yields a prediction of 2,540 undervotes, whereas matching on 18 offices yields 1,084 undervotes. The reason for this is that, as of the number of offices matched on increases, each voter shares her vote pattern with fewer and fewer other voters. For example, when matching on 18 races there are well over 400 million possible voting patterns, and 5,146 of the 17,825 ballots to be allocated were unique within Sarasota county.

We estimate the probability of undervoting using Charlotte County ballots and machines given an undervote in Sarasota as the Charlotte undervote probability divided by the Sarasota undervote probability (for a given vote pattern). For any CD 13 undervoter who cast a unique vote pattern within Sarasota, the non-parametrically estimated *ex ante* probability of undervoting in Sarasota is one. In the limit, if we could continue to increase the number of offices matched on, all voters would cast unique ballots and we would estimate that every Sarasota undervoter had an *ex ante* probability of one of doing so. This small sample bias exists to a lesser extent in all of our non-parametric allocations, and it accounts for the overall increase in the number of votes allocated to both Jennings and Buchanan.

The increase in net votes for Jennings may arise because we used Sarasota voters to allocate votes between Jennings and Buchanan in our logit model and use Charlotte voters here. If Jennings supporters where more likely to undervote than Buchanan supporters *ceteris paribus*, then using observed Jennings and Buchanan choices to estimate the probability of voting for one over the other would be biased against Jennings. On the other hand, the greater margin found here could also reflect greater *ceteris paribus* support for Jennings in Charlotte County.

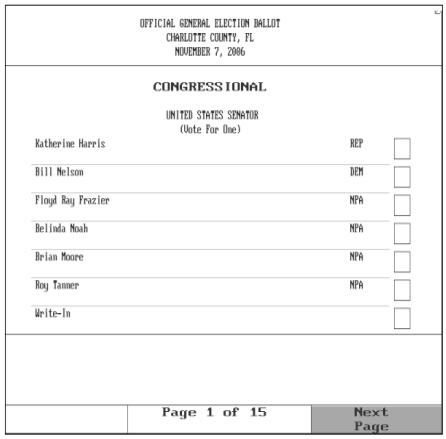
Whether one prefers the more general analysis presented here or the more conservative analysis presented in the body of the article, the conclusions are the same. The overvote in Sarasota was excessive even controlling for other voting behavior and the CD 13 election's result would have

surely been reversed in the absence of the excess undervote.

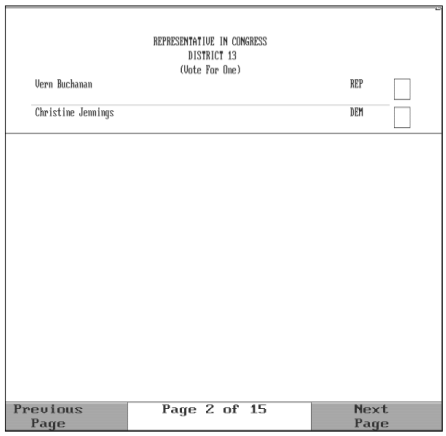


(a) Page 1 (b) Page 2

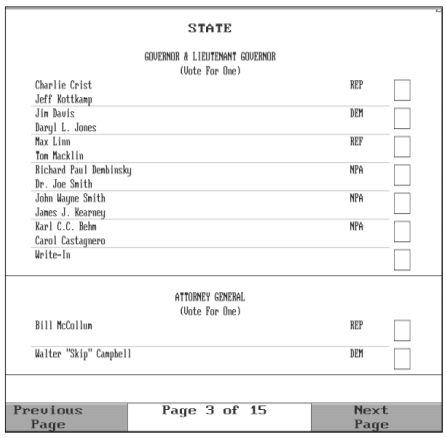
Figure 1: Pages One and Two of Sarasota County's 2006 Midterm Election Ballot



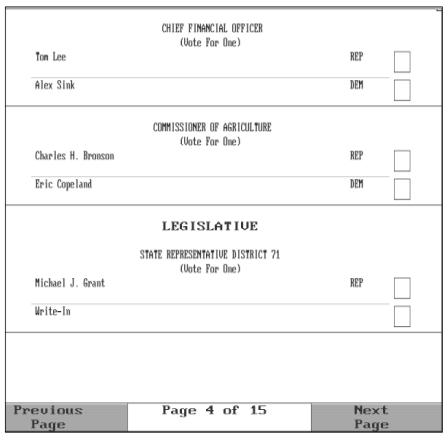
(a) Page 1



(b) Page 2



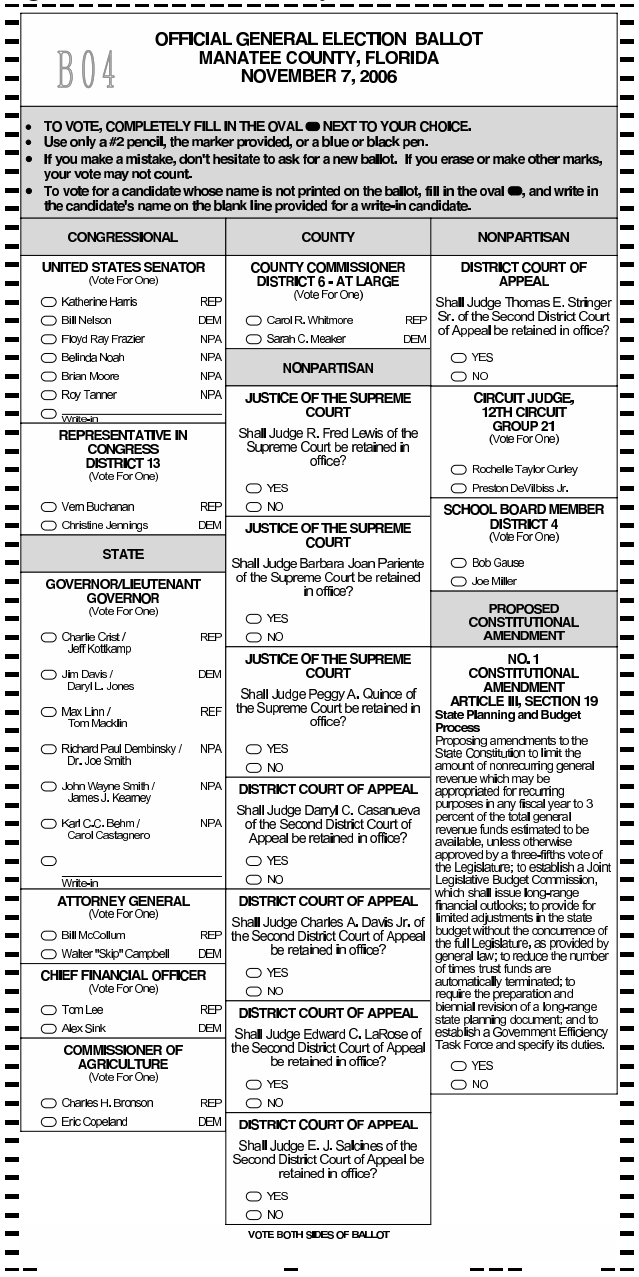
(c) Page 3



(d) Page 4

Figure 2: Pages One through Four of Charlotte County's 2006 Midterm Election Ballot

Figure 3: Picture of Page 1 of Manatee County's 2006 Midterm Election Ballot (Type B04)



B04

OFFICIAL GENERAL ELECTION BALLOT
MANATEE COUNTY, FLORIDA
NOVEMBER 7, 2006

- TO VOTE, COMPLETELY FILL IN THE OVAL ⬭ NEXT TO YOUR CHOICE.
- Use only a #2 pencil, the marker provided, or a blue or black pen.
- If you make a mistake, don't hesitate to ask for a new ballot. If you erase or make other marks, your vote may not count.
- To vote for a candidate whose name is not printed on the ballot, fill in the oval ⬭, and write in the candidate's name on the blank line provided for a write-in candidate.

CONGRESSIONAL

UNITED STATES SENATOR
(Vote For One)

- ⬭ Katherine Harris — REP
- ⬭ Bill Nelson — DEM
- ⬭ Floyd Ray Frazier — NPA
- ⬭ Belinda Noah — NPA
- ⬭ Brian Moore — NPA
- ⬭ Roy Tanner — NPA
- ⬭ Write-in

REPRESENTATIVE IN CONGRESS DISTRICT 13
(Vote For One)

- ⬭ Vern Buchanan — REP
- ⬭ Christine Jennings — DEM

STATE

GOVERNOR/LIEUTENANT GOVERNOR
(Vote For One)

- ⬭ Charlie Crist / Jeff Kottkamp — REP
- ⬭ Jim Davis / Daryl L. Jones — DEM
- ⬭ Max Linn / Tom Macklin — REF
- ⬭ Richard Paul Dembinsky / Dr. Joe Smith — NPA
- ⬭ John Wayne Smith / James J. Kearney — NPA
- ⬭ Karl C.C. Behm / Carol Castagnero — NPA
- ⬭ Write-in

ATTORNEY GENERAL
(Vote For One)

- ⬭ Bill McCollum — REP
- ⬭ Walter "Skip" Campbell — DEM

CHIEF FINANCIAL OFFICER
(Vote For One)

- ⬭ Tom Lee — REP
- ⬭ Alex Sink — DEM

COMMISSIONER OF AGRICULTURE
(Vote For One)

- ⬭ Charles H. Bronson — REP
- ⬭ Eric Copeland — DEM

COUNTY

COUNTY COMMISSIONER DISTRICT 6 - AT LARGE
(Vote For One)

- ⬭ Carol R. Whitmore — REP
- ⬭ Sarah C. Meaker — DEM

NONPARTISAN

JUSTICE OF THE SUPREME COURT
Shall Judge R. Fred Lewis of the Supreme Court be retained in office?

- ⬭ YES
- ⬭ NO

JUSTICE OF THE SUPREME COURT
Shall Judge Barbara Joan Pariente of the Supreme Court be retained in office?

- ⬭ YES
- ⬭ NO

JUSTICE OF THE SUPREME COURT
Shall Judge Peggy A. Quince of the Supreme Court be retained in office?

- ⬭ YES
- ⬭ NO

DISTRICT COURT OF APPEAL
Shall Judge Darryl C. Casanueva of the Second District Court of Appeal be retained in office?

- ⬭ YES
- ⬭ NO

DISTRICT COURT OF APPEAL
Shall Judge Charles A. Davis Jr. of the Second District Court of Appeal be retained in office?

- ⬭ YES
- ⬭ NO

DISTRICT COURT OF APPEAL
Shall Judge Edward C. LaRose of the Second District Court of Appeal be retained in office?

- ⬭ YES
- ⬭ NO

DISTRICT COURT OF APPEAL
Shall Judge E. J. Salcines of the Second District Court of Appeal be retained in office?

- ⬭ YES
- ⬭ NO

VOTE BOTH SIDES OF BALLOT

NONPARTISAN

DISTRICT COURT OF APPEAL
Shall Judge Thomas E. Stringer Sr. of the Second District Court of Appeal be retained in office?

- ⬭ YES
- ⬭ NO

CIRCUIT JUDGE, 12TH CIRCUIT GROUP 21
(Vote For One)

- ⬭ Rochelle Taylor Curley
- ⬭ Preston DeVilbiss Jr.

SCHOOL BOARD MEMBER DISTRICT 4
(Vote For One)

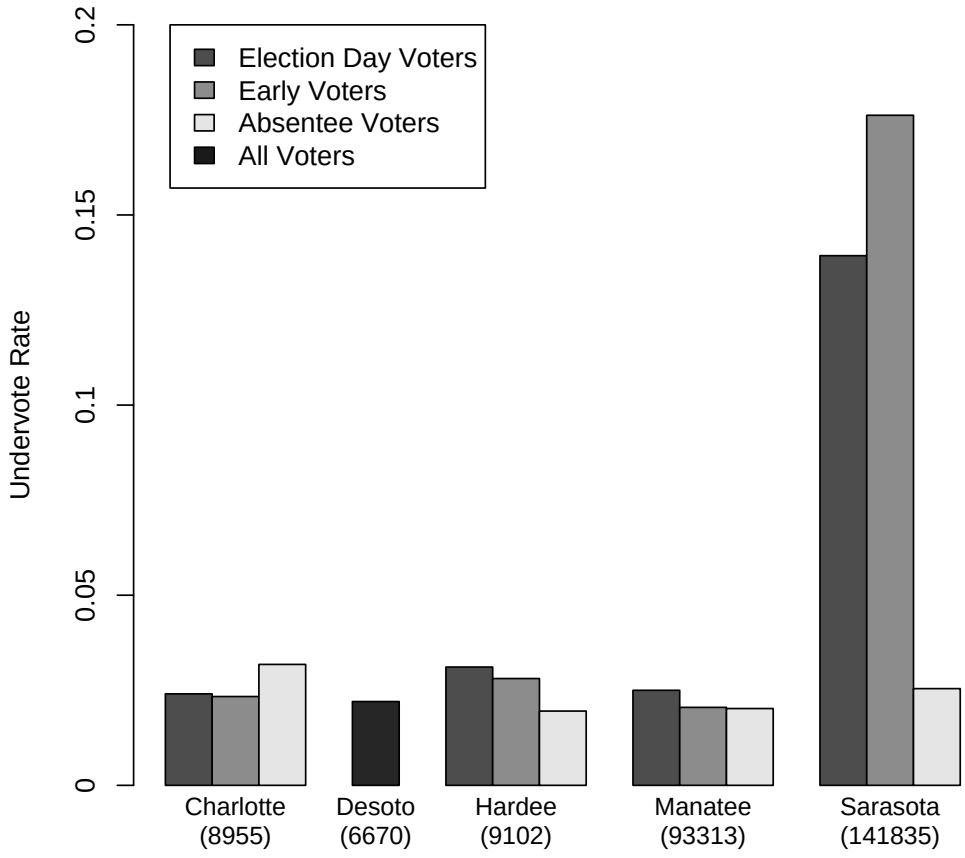
- ⬭ Bob Gause
- ⬭ Joe Miller

PROPOSED CONSTITUTIONAL AMENDMENT

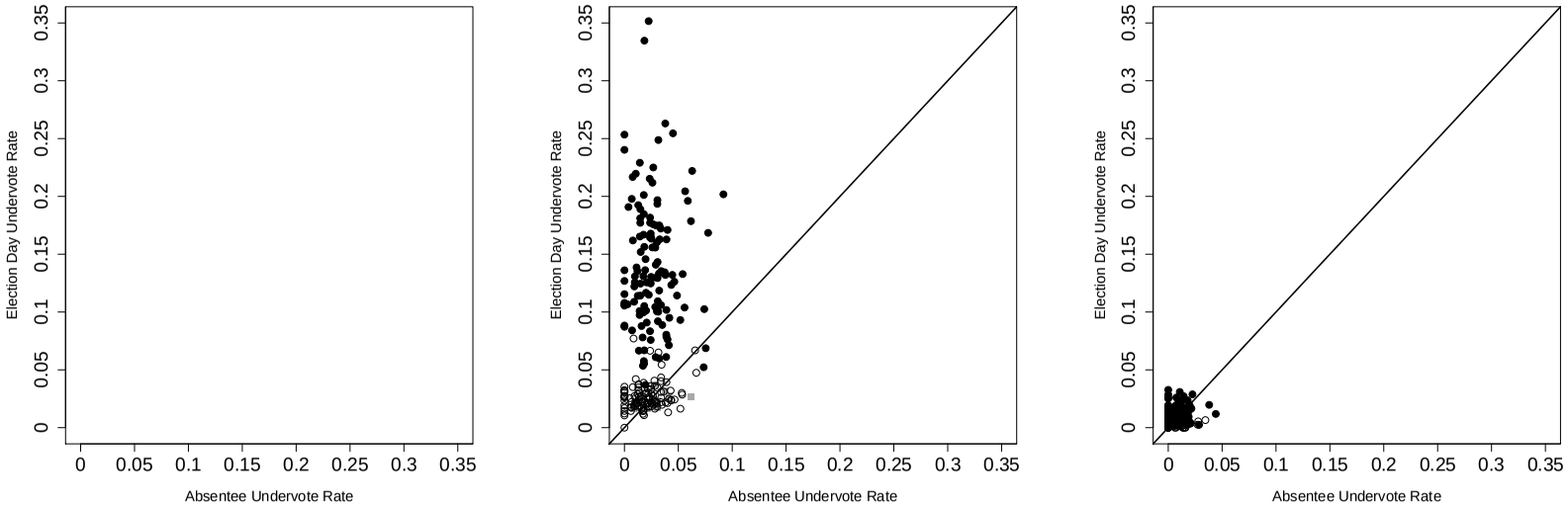
NO. 1 CONSTITUTIONAL AMENDMENT ARTICLE III, SECTION 19
State Planning and Budget Process
Proposing amendments to the State Constitution to limit the amount of nonrecurring general revenue which may be appropriated for recurring purposes in any fiscal year to 3 percent of the total general revenue funds estimated to be available, unless otherwise approved by a three-fifths vote of the Legislature; to establish a Joint Legislative Budget Commission, which shall issue long-range financial outlooks; to provide for limited adjustments in the state budget without the concurrence of the full Legislature, as provided by general law; to reduce the number of times trust funds are automatically terminated; to require the preparation and biennial revision of a long-range state planning document; and to establish a Government Efficiency Task Force and specify its duties.

- ⬭ YES
- ⬭ NO

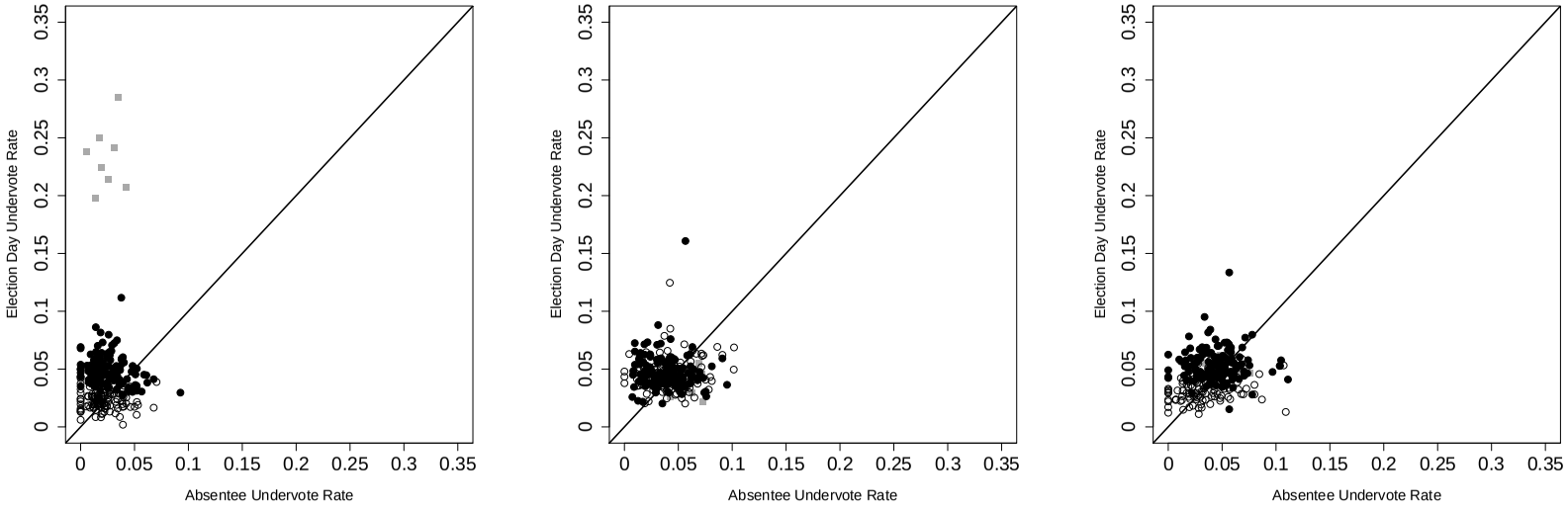
Figure 4: 2006 Undervote Rates among Counties Congressional District 13.



Note: total number of ballots cast listed in parentheses.



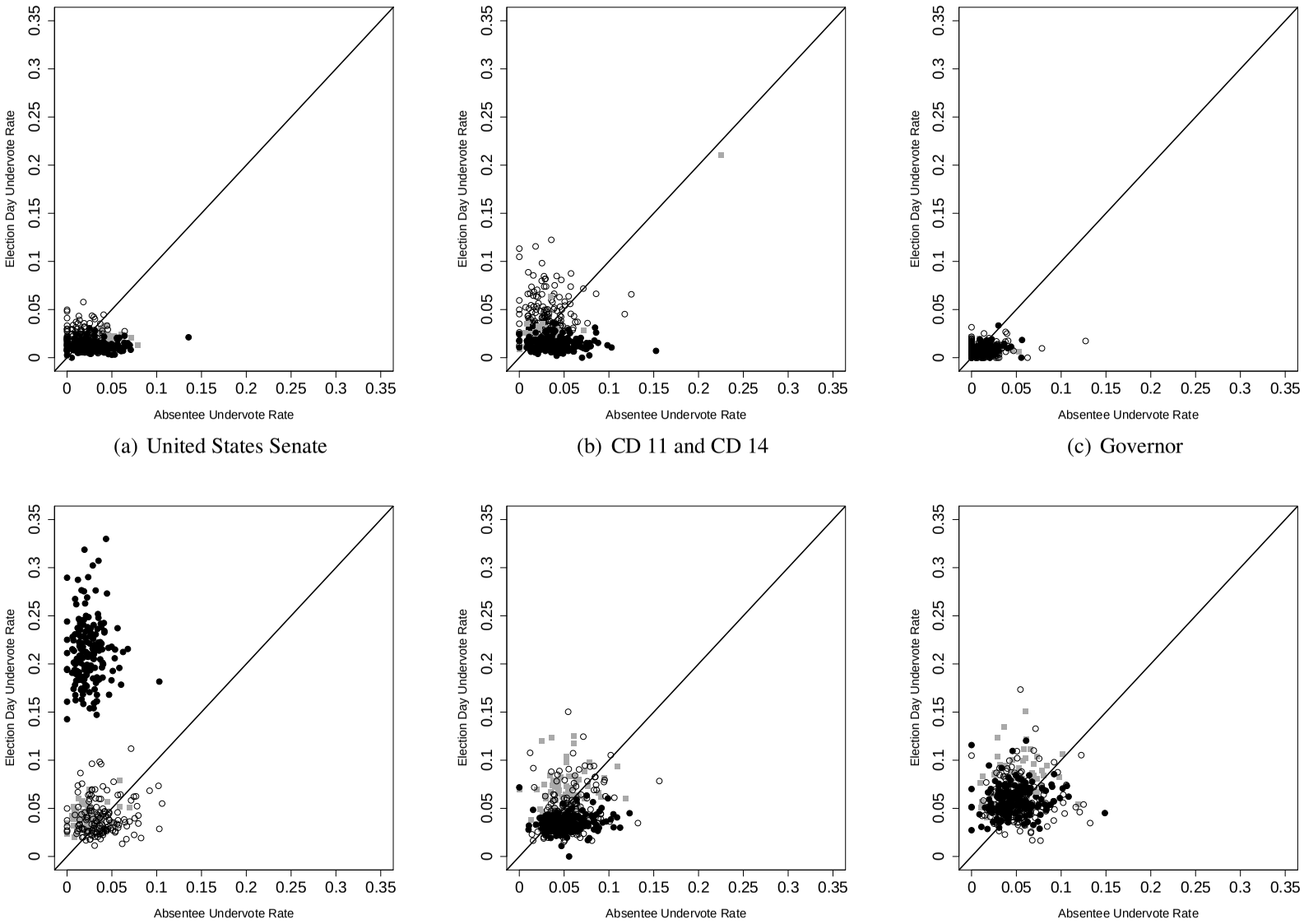
(a) United States Senate (b) CD 13 (c) Governor



(d) Attorney General (e) Chief Financial Officer (f) Comm. of Agriculture

Figure 5: Undervote Rates in Top Races for CD 13 Precincts

Note: Each solid dot represents a Sarasota County precinct; each grey square a Charlotte County precinct; and each open dot a Hardee or Manatee precinct.



(a) United States Senate (b) CD 11 and CD 14 (c) Governor

(d) Attorney General (e) Chief Financial Officer (f) Comm. of Agriculture

Figure 6: Undervote Rates in Top Races among CD 11 and CD 14 Precincts

Note: Each solid dot represents a Charlotte or Lee County precinct; each open dot a Hillsborough, Manatee, or Pinellas County precinct, and each grey square a Collier County precinct.

References

Aitchison, John. 1986. *The Statistical Analysis of Composition Data*. London: Chapman & Hall.

Ansolabehere, Stephen & Charles Stewart III. 2005. "Residual Votes Attributable to Technology." *Journal of Politics* 67(2):365–389.

Brady, Henry E., Justin Buchler, Matt Jarvis & John McNulty. 2001. "Counting All the Votes: The Performance of Voting Technology in the United States." Report published by the Survey Research Center and Institute of Governmental Studies, University of California, Berkeley.

Carrier, Michael A. 2005. "Vote Counting, Technology, and Unintended Consequences." *St. John's Law Review* 79(3):646–687.

Celeste, Richard F., Dick Thornburgh & Herbert Lin. 2005. *Asking the Right Questions About Electronic Voting*. Washington, DC: National Academies Press.

Dill, David L. & Dan S. Wallach. 2007. "Stones Unturned: Gaps in the Investigation of Sarasota's Disputed Congressional Election." Unpublished manuscript available at `http://www.cs.rice.edu/~dwallach/pub/sarasota07.pdf`.

Feldman, Ariel J., J. Alex Halderman & Edward W. Felten. 2006. "Security Analysis of the Diebold AccuVote-TS Voting Machines." Working paper, Center for Information Technology Policy and Department of Computer Science, Princeton University.

Herrnson, Paul S., Benjamin B. Bederson, Richard G. Niemi, Frederick Conrad, Michael J. Hanmer & Michael Traugott. 2006. "The Not So Simple Act of Voting: An Examination of Voter Errors with Electronic Voting." Unpublished working paper, University of Maryland, College Park.

Herrnson, Paul S., Owen G. Abbe, Peter L. Francia, Benjamin B. Bederson, Bongshin Lee, Robert M. Sherman, Fred Conrad, Richard G. Niemi & Michael Traugott. 2005. "Early Appraisals of Electronic Voting." Forthcoming, *Social Science Computing Review*.

Herron, Michael C. & Jasjeet S. Sekhon. 2005. "Black Candidates and Black Voters: Assessing the Impact of Candidate Race on Uncounted Vote Rates." *Journal of Politics* 67(1):154–177.

Herron, Michael C. & Jeffrey B. Lewis. 2006. "From Punchcards to Touchscreens: Some Evidence from Pasco County, Florida on the Effects of Changing Voting Technology." Unpublished working paper, Dartmouth College.

Ho, Daniel E. & Kosuke Imai. 2006. "Randomization Inference With Natural Experiments: An Analysis of Ballot Effects in the 2003 California Recall Election." *Journal of the American Statistical Association* 101(475):888–900.

Honaker, James, Jonathan Katz & Gary King. 2002. "A Fast, Easy, and Efficient Estimator for Multiparty Electoral Data with Improved Lemon Scent." *Political Analysis* 10(1):84–100.

Jackson, John E. 2002. "A Seemingly Unrelated Regression Model for Analyzing Multiparty Elections." *Political Analysis* 10(1):49–65.

Janiszewski, Chris. 1998. "The Influence of Display Characteristics on Visual Exploratory Search Behavior." *The Journal of Consumer Research* 25(3):290–301.

Katz, Jonathan & Gary King. 1999. "A Statistical Model for Multiparty Electoral Data." *American Political Science Review* 93(1):15–32.

Kimball, David. 2007. "Voting Equipment and Residual Votes on Ballot Initiatives: The 2006 Election in New Jersey." Unpublished manuscript available at `http://www.umsl.edu/~kimballd/NJ06resid.pdf`.

Kimball, David C. 2003. "Voting Methods Two Years After Florida." Paper Presented at the Annual Meeting of the Midwest Political Science Association.

Kimball, David C. & Martha Kropf. 2005. "Ballot Design and Unrecorded Votes on Paper-Based Ballots." *Public Opinion Quarterly* 69(4):508–529.

King, Gary, James Honaker, Anne Joseph & Kenneth Scheve. 2001. "Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation." *American Political Science Review* 95(1):49–69.

Knack, Stephen & Martha Kropf. 2003. "Roll-Off at the Top of the Ballot: International Undervoting in American Presidential Elections." *Politics & Policy* 31(4):575–595.

Koppell, Jonathan GS & Jennifer A. Steen. 2004. "The Effects of Ballot Position on Election Outcomes." *Journal of Politics* 66(1):267–281.

Mebane, Jr., Walter R. 2004. "The Wrong Man is President! Overvotes in the 2000 Presidential Election in Florida." *Perspectives on Politics* 2(3):525–535.

Mebane, Jr., Walter R. & David L. Dill. 2007. "Factors Associated with the Excessive CD-13 Undervote in the 2006 General Election in Sarasota County, Florida." Unpublished manuscript available at http://macht.arts.cornell.edu/wrm1/smachines1.pdf.

Rubin, Donald B. 1987. *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley & Sons.

Schafer, Joseph L. 1997. *Analysis of Incomplete Multivariate Data*. New York: Chapman and Hall.

Tomz, Michael, Joshua A. Tucker & Jason Wittenberg. 2002. "An Easy and Accurate Regression Model for Multiparty Electoral Data." *Political Analysis* 10(1):66–83.

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

Vanderleeuw, James M. & Glenn H. Utter. 1993. "Voter Roll-Off and the Electoral Context: A Test of Two Theses." *Social Science Quarterly* 74(3):664–73.

Wallach, Dan S. 2006. "Declaration of Dan S. Wallach." Unpublished legal declaration pursuant to Christine Jennings's election contest.

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