Ballot Formats, Touchscreens, and Undervotes:
A Study of the 2006 Midterm Elections in Florida\(^1\)

Laurin Frisina\(^2\)  Michael C. Herron\(^3\)  James Honaker\(^4\)
Jeffrey B. Lewis\(^5\)

Initial Draft: November 23, 2006
This Draft: December 3, 2006

\(^1\)The authors thank Greg Huber for comments on an earlier draft of this paper and thank election officials in Charlotte, Collier, DeSoto, Hardee, Hillsborough, Lee, Manatee, Pinellas, and Sarasota Counties for providing data and answering election-related questions. The most recent version of this paper can be found at [http://www.dartmouth.edu/~herron](http://www.dartmouth.edu/~herron).

\(^2\)B.A. in Political Science, University of Rochester, June, 2006 (laurinfrisina@hotmail.com).

\(^3\)Associate Professor, Department of Government, Dartmouth College and Visiting Associate Professor of Political Economy, Wallis Institute of Political Economy, University of Rochester. 6108 Silsby Hall, Hanover, NH 03755 (Michael.Herron@dartmouth.edu).

\(^4\)Assistant Professor, Department of Political Science, University of California at Los Angeles. 4289 Bunche Hall, Los Angeles, CA 90095-1472 (tercer@ucla.edu).

\(^5\)Assistant Professor, Department of Political Science, University of California at Los Angeles. 4289 Bunche Hall, Los Angeles, CA 90095-1472 (jblewis@ucla.edu).
Abstract

The 2006 midterm elections in Florida have focused attention on undervotes, ballots on which no vote is recorded on a particular contest. This interest was sparked by the high undervote count—more than 18,000 total undervotes out of 240,000 ballots cast—in Florida’s 13th Congressional District race, a race decided by a total of 369 votes. Using a combination of ballot-level and precinct-level voting data, we show that the high undervote rate in the 13th Congressional District race was almost certainly caused by the way that Sarasota County’s electronic touchscreen voting machines placed the 13th Congressional District race above and on the same screen as the Florida Governor 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 that race appeared below the Governor race on voting machine screens. Using a precinct-level statistical imputation model to identify and allocate excess undervotes, we find that there is a 90 percent chance that Sarasota undervotes were pivotal in the 13th Congressional District race. With more precise estimates that use ballot data, we find that there is essentially a 100 percent chance that Jennings would have won the CD 13 race had Sarasota voters voted in Charlotte County.
1 Introduction

Despite widespread concerns that electronic voting machines adopted in the wake of the 2000 presidential election would wreak havoc on the recent 2006 midterm elections, reports of election-altering ballot machine failures have thus far been relatively few. Although the midterm 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, few observers or participants have claimed that problems in election administration are to blame for the defeat of a preferred candidate.\(^1\) Indeed, compared to the controversies surrounding balloting in the 2000 Presidential race in Florida (e.g., Merzer 2001, Posner 2001), the 2004 Presidential race in Ohio (e.g., Freeman & Bleifuss 2006), and the 2004 Washington gubernatorial race (in which the final margin separating winning from losing candidate was 261 votes out of over 2.8 million total votes), the aftermath of the 2006 midterm elections has so far been relatively pacific.

Prior to the 2006 elections, concerns over electronic voting focused primarily on its susceptibility to tampering and a lack of independently verifiable audit trails (e.g., Feldman, Halderman & Felten 2006). Nonetheless, it was widely agreed that electronic voting machines were superior to the punchcard systems that in some cases they replaced insofar as electronic machines prevent overvotes (selecting more than one candidate in a given contest) and produce fewer undervotes (instances in which no candidate is selected) as well (Kimball 2003, Carrier 2005). A widely noted virtue of touchscreen machines is that they can be configured to draw a voter’s attention to races he or she initially overlooked or otherwise failed to record a preference (Celeste, Thornburgh & Lin 2005). In addition, touchscreen machines can be programmed to operate in many different languages.

\(^1\)Democrat Jim Webb defeated Republican incumbent George Allen by 9,329 votes in the Virginia Senate election; Democrat Joe Cortney defeated incumbent Republican Rob Simmons by 91 votes in Connecticut’s 2nd Congressional District; and, Republican incumbent Randy Brock is at the time of this paper’s writing leading Democrat Tom Salmon by 137 votes in the Vermont state auditor election pending a recount.
Despite these seeming virtues, the 2006 elections revealed that touchscreen machines do not uniformly produce low undervote rates. Indeed, in some instances touchscreen voting machines are associated with very high undervote rates, most notably in Florida’s 13th Congressional District (hereafter CD 13).

The election night canvass in CD 13 had Democrat Christine Jennings losing to Republican Vern Buchanan by some 377 votes; under Florida election law, this very narrow margin triggered an automatic recount.2 According to the recount, Buchanan beat Jennings by 119,142 to 118,741 votes—a margin of only 401 votes.3 The same count revealed that 21,303 ballots in CD 13, approximately 8.2 percent of those cast, recorded no candidate choice in the House race.4 These undervotes were heavily concentrated in Sarasota County, the largest of the five counties that contribute to CD 13. The other four counties are Charlotte, DeSoto, Hardee, and Manatee.5 Ignoring absentee ballots, which are optical scan, and focusing only on touchscreen ballots, the CD 13 undervote rate was 14.8 percent.

What accounts for Sarasota County’s extraordinarily high undervote rate in the CD 13 contest is the focus of this paper. Beyond what the prevalence of CD 13 undervotes might reveal about touchscreen voting in general, they also form the basis of Jennings’s legal challenge to Buchanan’s election victory. According to her legal complaint, “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

---


3The Florida Secretary of State calls this “2nd Set of Unofficial Returns,” See http://election.dos.state.fl.us/elections/resultsarchive/enight.asp. The recount vote totals cited here do not include overseas and military ballots. As of this paper’s writing, the certified margin between Buchanan and Jennings has shrunk to 369 votes, and Jennings has formally contested the election outcome.

4This count includes a handful of so-called overvotes in which a voter voted for more than one CD 13 candidate. Overvotes, which are invalid of course, can be cast by voters who use optical scan ballots but not by those using touchscreen machines. The latter are often called DRE (for Direct Recording Electronic) machines. Initial CD 13 recount results were downloaded from http://election.dos.state.fl.us/pdf/certCanvasCom.pdf.

5Pre-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.
counted due to the pervasive malfunctioning of electronic voting machines.”

CD 13 undervotes have also motivated a lawsuit by Sarasota voters who claim that they were disenfranchised by the failure of Sarasota County’s touchscreen machines. The plaintiffs in this lawsuit are requesting a re-vote.

Whether it is correct to attribute Sarasota’s high undervote rate to “machine malfunction” is in part a question of engineering and in part a question of semantics, i.e., what exactly constitutes a “malfunction?” While we cannot directly address engineering issues here and do not presume to provide a legal definition of “malfunction,” what we do provide is careful documentation of the extent of the undervote problem and its likely effect on the election outcome in CD 13. Although we cannot definitively rule out the possibility that there was some voting machine malfunction in the sense that Sarasota County’s touchscreen machines failed to record and tabulate actual screen touches for Buchanan and Jennings, we nonetheless show that such a malfunction must have been very localized and specialized if it had occurred. The touchscreen voting machines used in Sarasota did not produce unusually high undervote rates in other races. And, unusually high undervote rates did not occur in other races in which candidate buttons were located in a similar position on the screen to the candidate buttons for CD 13.

While within Sarasota County the undervote problem appears to be isolated to the CD 13 race, it does not appear that the high undervote rate in this race is attributable to features of the candidates themselves, the office they were contesting, or to features of the Buchanan-Jennings campaign. Indeed, elevated undervoting was not observed in CD 13 outside of Sarasota County nor was it a feature of absentee balloting within Sarasota. As we demonstrate below, the most likely culprit for the high CD 13 undervote rate is ballot layout. As has been noted in a variety of press accounts,

---

6The Jennings complaint was downloaded from http://www.heraldtribune.com/assets/pdf/SH81371120.PDF on November 21, 2006.
7This lawsuit is supported by various election advocacy groups. See the press release, titled “Sarasota Voters File Lawsuit for Revote in Congressional Race,” issued by People for the American Way. The release, which points out that the voter plaintiffs are both Democrats and Republicans, can be found at http://www.pfaw.org/pfaw/general/default.aspx?oid=23142.
among counties participating in the CD 13 race only in Sarasota was the CD 13 race itself placed on the same ballot screen with the Governor’s race which included six alternative candidates plus a write-in line. In other counties and on Sarasota’s absentee ballot, the CD 13 race appeared on its own ballot screen, next to (in a horizontal sense) the Florida United States Senate race, or along with a large number of other races in the case of optical scan ballots.

To buttress our conclusion that Sarasota County’s CD 13 undervote problem resulted from the ballot layout used in this county, we show very similar effects in Charlotte and Lee Counties where the race for Florida attorney general was combined with the Florida governor’s race on a single ballot screen. We find that Charlotte County’s ballot format is associated with an excess of undervotes that was roughly twice as prominent as found in Sarasota’s CD 13 ballots (and Lee County’s ballot format in the attorney general race is similar although not quite as pronounced). The focus in the public on the CD 13 race over the Attorney General contest presumably reflects the fact that the former race was exceedingly close whereas the Florida attorney general was close but not recount-close, so to speak. 8

It remains possible, of course, that a programming or design flaw in Charlotte County’s, Lee County’s, and Sarasota County’s touchscreen machines caused improper vote counts when a race (either the CD 13 rate or the race for Florida attorney general) was placed on the same page as the Florida governor’s race. Indeed, this possibility led the Florida Division of Elections, part of the Florida Department of State, to devise a testing procedure with the intention of determining if voting machine software or hardware problems were the sole cause or a contributing factor to the CD 13 undervote in Sarasota County. 9 The initial phase of the voting machine audit, a simulated election in Sarasota County, has not uncovered any significant software or hardware issues. 10

---

8According to official results from the Florida Department of State, the winning Florida Attorney General candidate, Bill McCollum, received 52.7% of the popular vote. This winning percentage is presumably not sufficiently large so as to think that voters could anticipated the results of the attorney general race prior to election day.

9For details of the Division of Election’s testing procedure see the “Parallel Test” design at http://election.dos.state.fl.us/pdf/parallelTestPlan.pdf.

10See, for example, “State says all discrepancies in test election due to ‘human error’,” Herald Tribune, November 30, 2006.
The question of disparate impact of an election administration practice—in the case at hand, a ballot format—is central to any analysis of electoral institutions and is obviously of paramount interest in CD 13 where the recorded electoral margin between Buchanan and Jennings is razor thin. Whatever doubt remains about the fundamental cause of the large number of CD 13 undervotes in Sarasota County, we can say with very high confidence that the undervote patterns observed in the CD 13 race would not have obtained had Sarasota voters been presented with the same voting machinery and ballot layouts used elsewhere in CD 13. Overall, as we explain in detail later, our precinct-level statistical models suggest that there is a 90 percent chance that Jennings would have won the CD 13 election had voting in Sarasota employed the machinery and ballots layouts used elsewhere in the district. We obtain even stronger results by analyzing the individual ballots cast in Sarasota and Charlotte counties. The ballot level analysis leaves little doubt that the election outcome would been reversed in the absence of the excess Sarasota undervotes. Indeed, we find that Jennings would have received a net advantage of between 3,068 and 3,359 votes if Sarasota voters had voted in Charlotte precincts—far more than the roughly 400 required to reverse the election.

In Section 2 we provide some background on the CD 13 race and comment on three explanations purporting to explain the high undervote rate in this contest. In Section 3 we present evidence that the undervote patterns in CD 13 are consistent with election administration problems and not with other claims about their causes, and we show by analysis of other races beyond CD 13 that the mechanism behind the Sarasota undervote rate appears to be a ballot format problem. In Section 4 we show that the CD 13 outcome at the time of this paper’s writing—Buchanan over Jennings by approximately 369 votes—could well have been reversed in the absence of said problems. Section 5 concludes.
2 Background

Florida has 25 Congressional Districts and 67 counties. Midway up the Florida’s Gulf Coast sits Sarasota County. Sarasota and its 156 election day precincts lie completely within the 13th Congressional District which includes parts of Charlotte County (eight election day precincts) and Manatee County (134 election day precincts, some of which are split between CD 13 and CD 11, the 11th Congressional District in Florida) and all of DeSoto and Hardee Counties (15 and 12 election day precincts, respectively). The five CD 13 counties vary in their use of voting technologies: Charlotte and Sarasota both use iVotronic touchscreens, manufactured by Election Systems and Software (known popularly as ES&S), for their election day and early voters whereas DeSoto, Hardee, and Manatee use optical scan voting. All five counties use optical scan ballots for absentee voting.

The fact that some CD 13 ballots do not contain a vote for a given office does not imply a failure of the machinery used to record votes nor does it imply the existence of any sort of election administration deficiency. Intentional undervotes (also commonly known as roll-offs and drop-offs) are a feature of every election. Indeed, the state of Nevada offers voters the explicit option of “None of the Above Candidates” when voting in state-wide races, and this option was exercised by an average of 4 percent of voters across the seven state-wide contests held in the 2006 midterm election (Nevada turnout was 585,986 voters). Beyond discontent with available candidates, deliberately choosing not to vote in a particular contest may reflect indifference among possible candidates or uncertainty about how or why to vote.

11What 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. Some counties (e.g., Hardee in 2004) have a single absentee precinct which includes all of the county’s absentee votes; others have multiple absentee precincts (e.g., Sarasota in 2004), and others (e.g., Sarasota in 2006) associate every absentee voter with a single election day precinct at which she would have voted had she voted on election day.

12Data on vote choices are from the Nevada Secretary of State’s Official Vote Summary (http://sos.state.nv.us/nvelection/2006StateWideGeneral/ElectionSummary.htm). The turnout figure is from http://www.secretaryofstate.biz/nvelection/2006StateWideGeneral/VoterTurnout.htm.
Nonetheless, the 8.2 percent undervote rate in Florida’s CD 13, in conjunction with the 13 percent such rate in Sarasota County, is clearly extraordinary. In modern presidential races, the intentional undervote rates is generally thought to be around 2 percent. In the 2004 presidential election, the overall Florida undervote rate was approximately 0.36 percent, this rate was 0.38 percent in Sarasota County, and it was 0.43 percent in counties that used touchscreens manufactured by ES&S (Florida Department of State 2005). In absolute terms presidential undervote rates are not comparable to Congressional district undervote rates produced in a midterm election. Nonetheless, the numbers cited above suggest that there is no reason to think a priori that Sarasota County is prone to high undervote rates.

In the sense of documenting what appears to be a disparate impact of an election administration practice, this paper is similar to previous studies of undervoting and residual votes (undervotes plus overvotes) more generally, many of which have focused on punchcard voting technology. Literature on residual vote rates has demonstrated that high residual vote rates are associated with particular kinds of voters and are often concentrated in low education, minority, and (consequently) more Democratic precincts (Brady, Buchler, Jarvis & McNulty 2001, Tomz & van Houweling 2003, Mebane 2004, Ansolabehere & Stewart III 2005).

With respect to the matter of ballot format itself, paper-based optical scan ballot format issues, like the way that optical scan ballots are filled out (either with circles or arrows to be connected), shading, and the presence of instructions, influence undervote and overvote rates (Kimball & Kropf 2005). And, Wand, Shotts, Sekhon, Mebane, Jr., Herron & Brady (2001) show that the butterfly ballot in Palm Beach County, Florida, altered the winner of the 2000 presidential election. Another election administration practice known to have effects on vote totals, and hence on very close election outcomes, is the order in which names appear on ballots (Koppell & Steen 2004, Ho & Imai 2006). Finally, Herrnson, Bederson, Niemi, Conrad, Hanmer & Traugott (2006) show in a sequence of experiments that different touchscreen formats affect the error rates that voters make when selecting candidates.
2.1 Three Explanations for the Sarasota Undervote Rate

A number of competing explanations have been offered for the seemingly aberrant undervote rate in CD 13 and in particular for Sarasota County’s undervote rate in the CD 13 race. As of this paper’s writing, outright election fraud or deliberate manipulation of voting machines with an intent to cause undervotes is not among the common conjectures. Rather, attention has focused on “protest undervotes,” machine malfunction, and voter confusion arising from ballot formatting.

2.1.1 Protest Undervotes

One explanation is that of the “protest undervote” wherein voters abstain from casting a vote in a particular race to express distaste for any and all candidates running in it. Vern Buchanan, the Republican candidate in CD 13, has speculated that, as characterized by the Herald Tribune, “Voters’ disgust with the barrage of often negative campaign ads and mailers could have contributed to the undervote.” Indeed, there is some evidence from U.S. Senate races that the degree of negativity in a campaign is associated with relatively high undervote rates (Ansolabehere & Iyengar 1995).

2.1.2 Ballot Formatting

Nonetheless, hundreds of citizens who voted in Sarasota have complained about the ballot format they faced, and that “The touch-screen [sic] ballot design concealed the candidates for the 13th District.” In what sense might this race have been “concealed”? The most commonly cited mechanism that may have led to voter confusion in the Sarasota part of CD 13 revolves around the vertical stacking of multiple races on the same touchscreen page.

Recall that Charlotte and Sarasota Counties both use iVotronic touchscreens for their election day and early voters. When a voter confronts an iVotronic, her “ballot” consists of a sequence of pages of races and corresponding candidate choices. In both Charlotte and Sarasota Counties, the

---

first ballot page contained only one race, the Florida United States Senate race. Beyond their first pages, though, Charlotte and Sarasota used different ballot layouts.

In Charlotte, the second ballot page contained the CD 13 race only, and the Florida gubernatorial race appeared on the third page along with the Florida attorney general race under it. Figure 1 displays the first four pages of the Charlotte County 2006 ballot. As detailed, note that the second page of the ballot—Figure 1(b)—contains a single race and that the third page—Figure 1(c)—combines governor and attorney general races.

In contrast, the second page of the Sarasota ballot contained both the CD 13 race and the Florida governor race. The third page of the Sarasota ballot included the attorney general, the chief financial officer, and the commissioner of agriculture races. Figure 2 displays the first two pages of the Sarasota County ballot. Note the Florida United States Senate race—Figure 2(a)—is on its own page but that the CD 13 race is combined with the governor race—Figure 2(b).

In DeSoto, Hardee, and Manatee Counties, election day and early voters voted with two-sided, optical scan ballots. There is no sense in which optical scan voting involves repeated page-turning, so to speak, as is common with touchscreens. The Manatee ballot is pictured in Figure 3. Absentee voters in all five CD 13 counties used optical scan ballots roughly similar to Manatee’s election day ballot.

It has been conjectured that the placement of the CD 13 contest (two candidates) next to the gubernatorial contest (six candidates plus a write-in) led voters to miss the former. Why precisely this might be true is beyond the scope of this paper. Presumably the questions, “How many races on a page is too many?” and “How many candidate choices on a page is too many?” are amenable to experimental research on human-computer interactions just as the butterfly ballot format was.

15The source for the Charlotte ballot pictures is email received on November 15, 2006 from the office of the Charlotte County Supervisor of Elections.
16Source for the Sarasota County ballot picture is personal communication between the authors and Dan Wallach on November 27, 2006.
17Source for the Manatee County ballot is http://www.votemanatee.com/PollLocationPub.asp?PollingId=856.
Figure 1: Pages One through Four of Charlotte County’s 2006 Midterm Election Ballot studied experimentally by Sinclair, Mark, Moore, Lavis & Soldat (2000).
2.1.3 Voting Machine Malfunctions

A third explanation for the high Sarasota undervote rate in CD 13 is machine or engineering failure.\textsuperscript{19} There are numerous possibilities in this area: software counting bugs that affected Sarasota’s iVotronics in a way that caused undervotes; touchscreen misalignment problems that caused voters to “miss” 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 and their associated consequences is long and beyond the scope of this paper. See Wallach (2006) for a discussion.

Broadly speaking, then, we characterize the aforementioned three explanations for the CD 13 undervote as 1) deliberate voter abstention; 2) ballot formatting; 3) engineering problems. This paper is purely a statistical exercise and as such cannot directly address the possibility that engineering lies underneath the undervote rates we study. However, to the extent that voter intentions or ballot formats appear to explain observed undervote rates, the conjecture that engineering flaws are

responsible for the CD 13 undervote rate becomes more difficult to sustain. Ultimately, no statistical analysis of observed voting data can distinguish between ballot format effects and engineering flaws that mimic ballot format effects.

What allows us to isolate the factors that are associated with the CD 13 undervote are the many
sources of within- and across-county variance in voting technologies and ballot formats used in and around CD 13. Recall, as we noted above, that Charlotte and Sarasota both used iVotronic touchscreens in the 2006 midterm elections. Thus, if there were a generic iVotronic effect in the CD 13 race, it would have affected both Charlotte and Sarasota undervote rates. And, if campaign negativity unduly affected the CD 13 race, it should have affected all voters who participated in it, not just those in one particular county. Finally, within CD 13 there is variance in ballot formats across counties and across time of voting (election day/early and absentee). Below we exploit all of these sources of variance and others as well.

3 Evidence of Anomalous Undervotes

In this section we first present a set of plots that highlight the anomalous CD 13 undervote in Sarasota County. Second, we present estimates of several statistical models applied to undervote rates in CD 11, CD 13, and CD 14. Finally, we consider a set of four non-Congressional races (Florida governor, Florida attorney general, Florida chief financial officer, and Florida commissioner of agriculture) and show how results from these races shed light on the Sarasota undervote in CD 13.

3.1 Evidence of Unusually High Undervote Rates in CD 13

A sense of the magnitude of the Sarasota undervote rate in CD 13 is apparent in Figure 4. This figure, which describes undervote rates for the five counties that contribute to CD 13, has two notable features. First, it shows that among CD 13 counties Sarasota had the largest election day minus absentee difference in undervote rates. Second, the figures shows that Sarasota County contributed the most votes of the five counties in the CD 13 race. Thus, Sarasota’s prominence in

---

20Figure 4 uses certified precinct returns except where they remain unavailable to the authors. As of this paper’s writing the pre-recount, 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.
Figure 4 cannot be attributed to sampling variance. Note that Hardee County had approximately 50 times fewer ballots than Sarasota.

Regarding the election day versus absentee gap in CD 13 undervote rates, on election day the CD 13 undervote rate was approximately 0.139 (\(\hat{\sigma} = 0.00116\)), but Sarasota absentee voters had an undervote rate of approximately 0.0250 (\(\hat{\sigma} = 0.000104\)). This translates to a difference of 0.114 (\(\hat{\sigma} = 0.00239\)). Thus, 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.

Absentee voters are a self-selected group, and perhaps Sarasota absentee voters are simply more attentive than Sarasota election day voters to undervotes. This seems implausible in light of Figure 5, which plots undervote rates in the Florida United States Senate race.\(^{21}\) As is evident in this figure, there is not much of a gap between the election day (1.3 percent) and absentee (0.9 percent) Senate undervote rates in Sarasota.

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 time of voting. This is an important point because, in theory, one could argue that Sarasota election day voters are perhaps unique in their willingness to incur high undervote rates solely in Congressional races. This argument is not very plausible in light of 2004 undervote rates.\(^{22}\)

The election day versus absentee difference among CD 13 precincts between is a key one, and one can see the importance of this in Figure 6. This figure plots for all CD 13 precincts except those in DeSoto County an election day undervote rate versus an absentee undervote rate. Note that the preponderance of Sarasota precincts (red dots plus confidence ellipses) lie above the pictured 45 degree line. This pattern implies that Sarasota precincts had election day CD 13 undervote rates greater than corresponding absentee CD 13 undervote rates. Judging by the distribution of non-Sarasota precincts (blue dots) and the way that this distribution is split by the 45 degree line, it is

\(^{21}\)The United States Senate race was the first on the ballot among all Florida counties. According to certified vote totals, Bill Nelson (Democrat, 60.3 percent) easily beat Katherine Harris (Republican, 38.1 percent) in the race.

\(^{22}\)In 2004, Katherine Harris (Republican, 55.3 percent) defeated Jan Schneider (Democrat, 44.7 percent).
Figure 4: 2006 Undervote Rates among Counties Congressional District 13.

Note: total number of ballots cast listed in parentheses.

clear that Sarasota is distinct from the rest of CD 13.

Finally, and to reinforce the point that undervoting in the CD 13 race in Sarasota County was anomalous compared to neighboring counties, consider Figure 7, which describes the location of
undervotes among election day Sarasota voters who cast one and only one undervote among races that all Sarasota voters could participate in (there were 10,428 such voters).\footnote{Data from this figure was provided by the Sarasota Supervisor of Elections and in particular was generated from the EL155.LST audit file produced by Sarasota’s touchscreen voting machines. The EL155.LST audit file lists vote profiles by both machine and precinct, and touchscreen machines which appeared in multiple precincts are considered to have been early voting machines.}

The implication of this figure is relatively straightforward: among election day Sarasota voters who cast one undervote, the modal category for this single undervote is the CD 13 race. Paradoxically, as indicated by shading in Figure 7, the CD 13 race was the closest race among the races in which all Sarasota voters could participate. Furthermore, it was very close to the top of the ballot, and the next closest race in terms of undervotes was the one-sided Sarasota Charter Review Board District 2 race. A comparable figure for Charlotte County is in Figure 8 (note the change in vertical axis scales between Figures 7 and 8). The lack of an election day CD 13 spike is evident in Charlotte, but
Figure 6: Election Day and Absentee Undervote Rates in CD 13

Note: Based on precincts Charlotte, Hardee, Manatee, Sarasota Counties. Each red dot in the figure represents a Sarasota precinct and each blue dot a non-Sarasota precinct. Ellipses around the dots represent 95 percent confidence sets.

Nonetheless the figure highlights a prominent attorney general spike. We will return to the matter of the attorney general race in Charlotte County after presenting more evidence on the CD 13 race in Sarasota.
Figure 7: Location of Undervotes among Sarasota County Election Day Voters whose Ballots Contained One Undervote

![Graph showing the location of undervotes among Sarasota County election day voters.]

Note: Based on 10,428 election day voters and races in which all Sarasota voters could participate. Darker colors signify closer races as bar darkness is inversely proportional to the margin between winning candidate (or alternative) and next closest candidate (or alternative).

3.2 Analysis of Undervote Rates in CD 11, CD 13, and CD 14

A limitation of the undervote results described above (Figures 4, 5, and 6) is that they do not control for the fact that, presumably, there is variance across CD 13 precincts in the underlying tendency to
Figure 8: Location of Undervotes among Charlotte County Election Day Voters whose Ballots Contained One Undervote

Note: Based on 727 election day voters and races in which all Charlotte voters in CD 13 could participate. Darker colors signify closer races as bar darkness is inversely proportional to the margin between winning candidate (or alternative) and next closest candidate (or alternative).

undervote. With this possibility in mind, we now consider regression analyses that seek to explain precinct-level undervote rates while controlling for precinct partisanship and undervote proclivity. We analyze three separate United States House races, those in CD 11, CD 13, and CD 14, and
this reflects the fact that it would be inappropriate for us to study CD 13 in isolation from other Congressional Districts in Florida. If what happened in the CD 13 contest is unique, i.e., if the undervote rate in this race had idiosyncratic features, then we should expect to see a lack of such features in Congressional District races that are physically close to CD 13.

As a caveat, however, we note that the CD 11 and CD 14 races were not at all close elections. In CD 11, the winning candidate received approximately 69.7% of the popular vote, and in CD 14 the corresponding winning percentage is approximately 64.4%. Non-competitive United States House races are hardly atypical in Florida and outside of it. The only Congressional race recount in Florida in the 2006 midterm elections took place in CD 13.

In the analyses that follow we divide each precinct in CD 11, CD 13, and CD 14 into election day, absentee, and early voting components. We then regress at the precinct level the appropriate Congressional race undervote rate (i.e., the total number of undervotes divided by number of ballots cast) on the United States Senate undervote rate, the fraction of U.S. Senate ballots cast for the Democratic candidate Bill Nelson, and indicator variables that control for county effects and time of voting (election day, early, and absentee) effects.\textsuperscript{24}

Before turning to regression results, though, consider Table 1, which describes election day and early voting technologies used in the CD 11, CD 13, and CD 14 counties. Note that all such counties used optical scan balloting for their absentee voters. Sarasota County, as made clear in the table, is notable insofar as it used iVotronic touchscreens and did not have its Congressional district race on a separate page. The same thing applies to Lee County, about which we have much more to say later. As an aside, the order of top races in all Florida counties was United State Senate, United States Congress, Florida governor, Florida attorney general race, Florida chief financial officer, and Florida commissioner of agriculture.

Coefficient estimates from our CD 11, CD 13, and CD 14 regressions are in Table 2.\textsuperscript{25} We see

\textsuperscript{24}The intercepts in our regressions are suppressed, and all regressions are weighted by number of ballots cast per precinct.

\textsuperscript{25}Undervote rates from the three Congressional district races can be modeled hierarchically (since precincts are
Table 1: Election Day Voting Technologies Across Congressional Districts 11, 13, and 14

<table>
<thead>
<tr>
<th>District</th>
<th>County</th>
<th>Technology</th>
<th>Type</th>
<th>Location of Congressional District (CD) Race</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>Hillsborough</td>
<td>Touchscreen</td>
<td>AVC Edge</td>
<td>Two column ballot, CD race right of Senate</td>
</tr>
<tr>
<td></td>
<td>Manatee</td>
<td>Optical Scan</td>
<td>Accuvote</td>
<td>Paper ballot, multiple pages</td>
</tr>
<tr>
<td></td>
<td>Pinellas</td>
<td>Touchscreen</td>
<td>AVC Edge</td>
<td>Two column ballot, CD race right of Senate</td>
</tr>
<tr>
<td>13</td>
<td>Charlotte</td>
<td>Touchscreen</td>
<td>iVotronic</td>
<td>CD race on its own page</td>
</tr>
<tr>
<td></td>
<td>Hardee</td>
<td>Optical Scan</td>
<td>Accuvote</td>
<td>Paper ballot, multiple pages</td>
</tr>
<tr>
<td></td>
<td>Manatee</td>
<td>Optical Scan</td>
<td>Accuvote</td>
<td>Paper ballot, multiple pages</td>
</tr>
<tr>
<td></td>
<td>Sarasota</td>
<td>Touchscreen</td>
<td>iVotronic</td>
<td>CD race on same page as governor race</td>
</tr>
<tr>
<td>14</td>
<td>Charlotte</td>
<td>Touchscreen</td>
<td>iVotronic</td>
<td>CD race on its own page</td>
</tr>
<tr>
<td></td>
<td>Collier</td>
<td>Touchscreen</td>
<td>iVotronic</td>
<td>CD race race on its own page</td>
</tr>
<tr>
<td></td>
<td>Lee</td>
<td>Touchscreen</td>
<td>iVotronic</td>
<td>CD, governor, and atty. gen. races on same page</td>
</tr>
</tbody>
</table>

from this table that precincts with high Senate undervote rates tend to have high Congressional undervote rates and that the precinct-level relationships between these two variables are statistically significant at conventional confidence levels (the three p-values for Senate Undervote Rate are all less than $1 \times 10^{-6}$). This is consistent with Herron & Sekhon (2003), who show using ballot data that voters who undervote on one top race tend to undervote on others. Of course, it does not follow from Table 2 that the same Florida voters who undervoted on the Senate race also undervoted on the Congressional race. Nonetheless, we also see from the table that Democratically-leaning precincts tend to produce disproportionately many undervotes, and this too is consistent with existing literature.

Table 2’s many indicator variables for county and time of voting effects are summarized in Figure 9. This figure has one bar for each county and Congressional district, and the value of a bar is the difference between an election day effect of undervoting minus the corresponding absentee effect. For instance, the first bar in Figure 9 is valued at approximately 0.01 because, in the CD 11 section in Table 2, the Hillsborough election day effect on undervoting is -0.00088, the absentee effect is -0.01092, and the difference between these two effects is approximately 0.01.26

---

26Standard errors for bar values depend on estimated covariances that are not depicted in Table 2. These covariances nested within counties), and future versions of the paper will take advantage of this structure and also add additional Congressional districts to the three considered here.
Table 2: Results for Congressional District Regressions

<table>
<thead>
<tr>
<th>District</th>
<th>County</th>
<th>Type</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>Summary Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>Hillsborough</td>
<td>Election day</td>
<td>-0.00088</td>
<td>0.00482</td>
<td>n = 659</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Absentee</td>
<td>-0.01092</td>
<td>0.00456</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Early</td>
<td>-0.00528</td>
<td>0.00517</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Manatee</td>
<td>Election day</td>
<td>-0.00513</td>
<td>0.01074</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Absentee</td>
<td>-0.00365</td>
<td>0.02168</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Early</td>
<td>0.02339</td>
<td>0.03194</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pinellas</td>
<td>Election day</td>
<td>0.00945</td>
<td>0.00625</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Absentee</td>
<td>0.00093</td>
<td>0.00790</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Early</td>
<td>0.00894</td>
<td>0.00893</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Senate Undervote Rate</td>
<td>1.01316</td>
<td>0.05720</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Senate Democratic Fraction</td>
<td>0.03806</td>
<td>0.00719</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Charlotte</td>
<td>Election day</td>
<td>-0.0421</td>
<td>0.0131</td>
<td>n = 891</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Absentee</td>
<td>-0.0388</td>
<td>0.0169</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Early</td>
<td>-0.0432</td>
<td>0.0144</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hardee</td>
<td>Election day</td>
<td>0.0016</td>
<td>0.0088</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Absentee</td>
<td>-0.0057</td>
<td>0.0185</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Early</td>
<td>-0.0025</td>
<td>0.0126</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Manatee</td>
<td>Election day</td>
<td>-0.0372</td>
<td>0.0096</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Absentee</td>
<td>-0.0339</td>
<td>0.0093</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Early</td>
<td>-0.0377</td>
<td>0.0119</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sarasota</td>
<td>Election day</td>
<td>0.0737</td>
<td>0.0098</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Absentee</td>
<td>-0.0334</td>
<td>0.0096</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Early</td>
<td>0.1111</td>
<td>0.0107</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Senate Undervote Rate</td>
<td>1.1951</td>
<td>0.1849</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Senate Democratic Fraction</td>
<td>0.0904</td>
<td>0.0165</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Charlotte</td>
<td>Election day</td>
<td>0.00107</td>
<td>0.00309</td>
<td>n = 767</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Absentee</td>
<td>-0.00220</td>
<td>0.00423</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Early</td>
<td>0.00064</td>
<td>0.00372</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Collier</td>
<td>Election day</td>
<td>0.00061</td>
<td>0.00218</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Absentee</td>
<td>0.00105</td>
<td>0.00225</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Early</td>
<td>-0.00133</td>
<td>0.00228</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lee</td>
<td>Election day</td>
<td>-0.00455</td>
<td>0.00233</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Absentee</td>
<td>0.01469</td>
<td>0.00232</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Early</td>
<td>-0.00501</td>
<td>0.00251</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Senate Undervote</td>
<td>0.50015</td>
<td>0.03548</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Senate Democratic Fraction</td>
<td>0.02388</td>
<td>0.00419</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Since Figure 9 shows differences in effects, it implicitly controls for county effects. That is, if all Hillsborough voters had, say, relatively low undervote rates on account of some factor unique to election administration in Hillsborough County, then the effect of this factor would be present in all three Hillsborough estimates in Table 2 and hence not present in the differences in Figure 9.

Positive values in the Figure 9 denote contexts where, controlling for precinct characteristics, there was more election day undervoting than absentee undervoting. Clearly, the most dramatic value in Figure 9 is that for Sarasota County in CD 13: the election day minus absentee difference is on the order of 0.10, meaning that the election day undervote rate among Sarasota County precincts was about ten percent higher than the absentee undervote rate, all things equal.\footnote{This 10 percent result is robust to inclusion of two additional variables in the CD 13 regression model: the fraction of precinct voters who supported the first constitutional amendment vote on the 2006 Florida ballot (the amendment passed 59.8\% to 40.2\% and dealt with budgeting matters) and the fraction of undervotes cast in the first amendment race.} Notably, of the four CD 13 counties studied here, only Sarasota had a significant difference between election day and absentee voters. Furthermore, the Sarasota election day minus absentee difference is easily statistically significantly different than the other three CD 13 differences (calculations available on request).

Figure 9’s other statistically significant election day versus absentee differences are in the Hillsborough part of CD 11 and the Lee County part of CD 14. For the latter, we suspect that the explanation lies in the partisanship of Lee County election day voters compared to absentee voters. Namely, the former appear to be much more Democratic than the latter: approximately 53\% of them voted for the Democratic Florida United States Senate candidate versus approximately 43\% of absentee voters. If strong Democratic partisans simply dropped out of the CD 14 race on account of its being a certain Republican victory, then one might expect to see a negative Lee County election day effect as in Figure 9. The same thing may explain the Hillsborough effect as well in that strong Democratic partisans were more likely to participate in a lopsided Democratic victory.

Regardless, in terms of substantive significance Figure 9’s Sarasota effect is over five times as large

\footnote{These values are available on request from the authors.}
Figure 9: Effect Across Three Congressional Districts of Election Day versus Absentee Voting on Undervote Rates Controlling for County, Precinct Partisanship, and Precinct Tendency to Under-vote

Note: Values of bars are based on differences between election day and absentee indicator variables from Table 2. Bold county names denote bar values that are significantly different from zero at the 0.05 level.

What might explain Sarasota’s prominence in Figure 9? In light of our three conjectures noted
earlier, it is hard to imagine that the Sarasota result reflects deliberate voter choices. If, say, CD
13 election day voters were driven away from participating in their Congressional race by a blitz of last-minute negativity, then this should have affected all four counties in the CD 13 race and not just Sarasota. As to the possibility that Sarasota voters are simply different than other CD 13 votes insofar as a willingness to undervote, this claim cannot explain the Sarasota effect in Figure 9 because the effect is calculated controlling for county effects and and precinct willingness to undervote.

Rather, recall from Table 1 that among counties studied here Sarasota County was one of two in its use of iVotronic machines in conjunction with a Congressional district race that did not appear on its own ballot page. If, for reasons that transcend this paper, voters are more likely to undervote in a two-candidate race that appears immediately above or conceivably immediately below a six-candidate race that includes a write-in option, then we would expect to see a large Sarasota CD 13 effect, precisely what we observe in Figure 9. We might also expect to see a large Lee County CD 14 effect, which we do not see. However, the CD 14 effect, we suspect, reflects partisanship based undervoting in a rather lopsided race.

Could the Sarasota effect in Figure 9 reflect engineering problems as discussed earlier? There are several ways to think about this question. First, if voting hardware or software were the problem, then there must be something unique to Sarasota’s use of iVotronics that in November, 2006 did not dramatically affect, say, Charlotte County’s or Lee County’s machines. Whether iVotronic software is tailored for individual counties is not something we know. If, though, it is not, then there must be some sort of an interaction with a generic iVotronic software problem and Sarasota’s use of it. Perhaps Sarasota’s iVotronic maintenance program is different than the maintenance program in surrounding counties.

Second, if there were hardware or software problems among Sarasota’s iVotronics, then they affected practically all such machines. This is an important point in light of our regression results in Table 2. One can always ask, that is, if a small group of aberrant precincts affected a set
of regression. To see that the Sarasota undervote problem was not restricted to a few precincts, consider Figure 10, which displays box plots of differences between election day and absentee undervote rates among precincts in CD 13.

Figure 10: Differences in Congressional District 13 Election Day and Absentee Undervote Rates among Four Counties
Figure 10 shows that the median election day minus absentee CD 13 undervote rate difference in Sarasota County was much greater than the maximum such values in Charlotte, Hardee, and Manatee counties. The box plot also shows the first quartile of differences, the third quartile, and the minimum and maximum differences. The implication of Figure 10 is that the CD 13 regression results in Table 2 do not reflect a few precincts with enormous undervote rate problems but rather a consistent pattern among Sarasota precincts. Thus, if there were hardware and/or software issues at play in Sarasota, they were reasonably pervasive and not restricted to a few problematic precincts or even voting machines within precincts.

3.3 More Evidence of Ballot Format-Induced Undervotes

If Sarasota’s grouping of races on its iVotronic touchscreen machines drove the CD 13 undervote in that county, then we should expect to see similar evidence of unusually high undervote rates when we observe other groupings of races in counties that use iVotronics. To this end we turn to the top four Florida races among the eight counties that comprise CD 11, CD 13, and CD 14. As before we consider all election day, early voting, and absentee precincts in these counties and calculate undervote rates for the following Florida races: governor, attorney general, chief financial officer, and commissioner of agriculture. The latter three races are called cabinet races.

Among the eight counties we study four use iVotronics: Charlotte, Collier, Lee, and Sarasota. For these four counties Table 3 describes notable groupings for the governor and cabinet races akin to the way that the CD 13 race was grouped with the Florida gubernatorial race in Sarasota County. As noted earlier, the governor’s race in Florida had six candidates plus a write-in slot whereas the cabinet officers all had two candidates and no write-in possibilities.

In light of Table 3 and our CD 13 findings, we should expect to see large election day versus absentee differences for the attorney general race in Charlotte and Lee Counties. These reflect the most asymmetric groupings in the table just as Sarasota grouped a small CD 13 race with a multi-candidate gubernatorial contest. We also might expect to see grouping effects in Collier County
Table 3: Groupings among Governor and Cabinet Races in iVotronic Counties

<table>
<thead>
<tr>
<th>County</th>
<th>Groupings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Charlotte</td>
<td>Governor paired with Attorney General</td>
</tr>
<tr>
<td>Collier</td>
<td>Governor alone, then three cabinet races on next page</td>
</tr>
<tr>
<td>Lee</td>
<td>Governor paired with Attorney General</td>
</tr>
<tr>
<td>Sarasota</td>
<td>Governor with CD 13, then cabinet races together</td>
</tr>
</tbody>
</table>

insofar as this county combined all cabinet races on a single page.

Before turning to precinct regression results we first examine some ballot data from Charlotte County with an eye on the cabinet races. Such data were already displayed in Figure 8 and an attorney general effect was noted there. Figure 11 shows similar effects.

The figure shows that Charlotte voters behaved very differently in the attorney general race than in other top races. Namely, conditional on undervoting in attorney general, many Charlotte voters had zero undervotes among a large set of races; the same thing is true conditional on not undervoting in the attorney general race. However, conditional on undervoting in the Senate race (or in the other races in Figure 11) the probability of zero undervotes among other races is small. The point here is simple: attorney general undervoters in Charlotte are not very different than attorney non-undervoters, but Senate undervoters are very different than Senate non-undervoters.

We now estimate four weighted least squares regressions, one for each top Florida race. Each regression produces 26 slope estimates: a Senate undervote estimate, a Senate Democratic share estimate, and 24 (eight by three) estimates for eight counties cross three different times of voting (complete regression results are available from the authors). Figure 12 summarizes election day minus absentee differences, controlling for precinct undervote tendencies and partisanship using, as we did before, the fraction of Senate undervotes by precinct and the fraction of Democratic Senate votes.

There are two notable features of Figure 12. First, with respect to the attorney general race, Charlotte and Lee Counties had enormous gaps between election day and absentee voting propen-
Figure 11: Distribution of Undervote Counts by Ballot for Election Day Voters in Charlotte County

Note: This figure depicts the distribution of undervote counts by ballot and across 18 common races in Charlotte County. For voters who had a U.S. Senate race undervote, the 19 rectangles in the left stack of the U.S. Senate column describe the fraction of ballots that had zero undervotes ignoring the Senate race (the lowest rectangle, i.e., approximately 15 percent), the fraction with one undervote ignoring the Senate race (the second lowest rectangle, i.e., another 10 percent), and so forth. The right hand stack of 19 rectangles in the U.S. Senate column describes the distribution of undervotes by ballot (zero, one, two, and so forth through 18) for voters who did not have a Senate race undervote.
Figure 12: Effect on Top Florida Races of Election Day versus Absentee Voting on Undervote Rates Controlling for County, Precinct Partisanship, and Precinct Tendency to Undervote
sities to undervote. The difference here is on the order of 20 percent, all things equal, which is about twice as large as the CD 13 effect identified in Sarasota County. These gaps are consistent with the race grouping conjectures that follow from Table 3.

Second, Collier County shows a consistent effect of election day voting although nothing nearly as dramatic as the Charlotte and Lee effects noted above. These Collier effects seem roughly consistent with the consequences of grouping three races on a page, i.e., too many similarly sized races on a single page leads to undervotes. However, this conclusion should be considered highly speculative,

The reason that Figure 12 is so important is because it supports the reasoning offered here for the CD 13 undervote. If one were to take seriously the notion that voter indifference, potentially induced by campaign negativity, drove the CD 13 undervote rate in Sarasota, one would have to argue that voter indifference over Florida attorney general candidates affected voters in Charlotte and Lee Counties. How this might happen is not clear. If Florida voters were in general indifferent between attorney general candidates, why would they act on their indifference only in two counties and why on election day but not during absentee voting, controlling for precinct tendencies?

Similarly, the case for hardware and software problems in Sarasota becomes more difficult to make given Figure 12. If Sarasota experienced a generic voting machine malfunction, then this same malfunction affected two other neighboring counties but only in one particular race (and not in the same race that the malfunction appeared in Sarasota). Of course this could have happened: flaws in software code can interact in ways that are hard to predict, and we can never rule out the possibility that the culprit here is engineering. Nonetheless, without laboratory evidence this seems to be a hard case to make. At this point the argument that ballot formatting drove the CD 13 undervote in Sarasota and the attorney general undervote in Lee and Collier is more compelling.
4 Allocating CD 13 Undervotes to Buchanan and Jennings

We now consider the question, what would have happened in CD 13 if the undervotes cast in Sarasota followed patterns similar to those cast elsewhere in the district? We answer this question in two ways, first with precinct returns from eight counties and second with ballot data from two.

In Sarasota County, a total of 120,686 touch screen ballots were cast in either early or election day voting. Of these, 17,811, or 14.8 percent, did not vote in the CD 13 race while only 2.6 percent undervoted across other Charlotte, Hardee, Manatee Counties in this same race. If we assume that a similar 2.6 percent of voters in Sarasota preferred to have undervoted in the CD 13 race, then approximately 14,750 voters in Sarasota County would have cast a vote in this race if the Sarasota touchscreen ballot had been no different than the ballots in Charlotte, Hardee, and Manatee.

4.1 Two Stylized Stories about the Source of Sarasota Undervotes

There are two stylized stories we might tell to explain why these 14,750 voters did not cast a vote in the CD 13 race. One, the surplus level of undervoting in Sarasota might have been driven entirely by some voters accidently not seeing the CD 13 candidates; or, two, it might have been driven by indifference on the part of certain types of voters. That is, some Sarasota undervotes were caused by individuals who saw the CD 13 race and chose not to vote in it while the rest were caused by voters who did not initially see the CD 13 race because of the Sarasota ballot format but would have seen it if given a better format. Call these suppressed ballots and the votes which would have been cast from the suppressed ballots the intended votes.28

---

28Certainly, some suppressed voters who did not see the Congressional race on the ballot would have gone on to undervote deliberately in CD 13 if they had seen it. Thus, we assume that some fraction of intended votes would have been cast as undervotes.
4.1.1 Suppression by Random Accident

At one extreme, we might suppose that the rate of suppressed votes is explained entirely by ballot mechanics, i.e., the touchscreen ballot format in Sarasota County led to a certain probability of “accidents” where voters did not see the CD 13 race and did not ever face the choice they would otherwise have made between voting for a Republican or Democratic candidate or indeed choosing not to cast a CD 13 vote at all. If surplus undervoting in Sarasota County were driven by completely random accidents, then the suppressed voters are a completely random sample of the voters in each precinct.

From the assumption that ballot mechanics explain everything, two points can be made. First, the distribution of intended votes should follow (within sampling variance) the distribution of votes cast by voters who were not suppressed. If this model were true, then the roughly 15,000 intended votes would have fallen in about the same split as the unsuppressed votes in Sarasota County and thus split roughly 7,500 for Jennings and 6,500 for Buchanan. This would result in Jennings picking up 1000 votes over Buchanan and overcoming the 369 vote deficit on the day.

Second, both candidates received fewer votes than they would have received if the intended votes had been cast, as suppression transferred votes from each of their columns to the undervote column. This second point can be clearly demonstrated with precinct data. Suppose that we calculate vote shares as the percent of the votes cast for the Democratic candidate out of all ballots cast including undervotes, such as:

\[
\text{Democratic Voteshare} = \frac{\text{Democratic Votes}}{\text{Dem. Votes} + \text{Rep. Votes} + \text{Undervotes}}. \quad (1)
\]

Then, under the story of suppression caused by accident, we should expect some voters who intended to vote Democratic to instead undervote, thus the numerator above gets smaller while the denominator stays the same size as votes are simply transferred from Democratic votes to undervotes. Thus the Democratic vote share, as a share of all votes cast, should be decreased by
Figure 13: Democratic Vote Shares in Senate and CD 13 Races for Early and Election Day Voters

Note: Each red dot in the figure represents a Sarasota precinct and each blue dot a non-Sarasota precinct.

accidents. The Republican vote share would similarly be reduced.

Is this accident theory plausible in any sense? Figure 13 plots the vote shares of Jennings vertically and for the Democratic Senate candidate horizontally. These two vote shares should predict each other: as more (or fewer) voters in a district cast ballots for the Senate Democratic candidate, we would expect more (or fewer) voters to also cast a ballot for Jennings in her House race.\(^{29}\)

What is clear from Figure 13 is that, at any level of Democratic Senate support, the number of Democratic votes for Jennings is below what is expected: red points are clustered horizontally below blue points and the red summary line dips below the blue summary line. Similarly, Buchanan vote share in Sarasota County falls below what would be expected given Senate vote shares.\(^{30}\) This

\(^{29}\)The United States Senate race is chosen for this example as of all races it has the strongest relationship with the Congressional race.

\(^{30}\)The corresponding figure is not shown, but patterns in it are as predicted by accidents that suppress Republican votes and change them into undervotes.
effect seems to be more pronounced in the early voting than in the election day voting, which is consistent with the story that some poll workers were warned of the problematic Sarasota ballot format problem and tried to pass on this warning to voters who voted on the day of the election. The slope of the Sarasota lines in Figure 13 is smaller than the non-Sarasota lines, which is what we would expect if some constant fraction of all votes were being randomly suppressed and converted to undervotes.  

4.1.2 Suppression by Indifference

Beyond this accident story, there is an alternate stylized conjecture for how suppressed intended votes would have turned out had they not been suppressed. Suppose that there are two types of voters, engaged and disengaged. Engaged voters followed the CD 13 election closely, clearly prefer one candidate over another, and care deeply about the outcome. When faced with a choice, engaged voters vote for the candidate they prefer. Disengaged voters may have no information about the CD 13 race, or are disenchanted with both candidates, or for some other reason they do not care about the outcome of the House race. When faced with the choice, Buchanan versus Jennings, disengaged voters may deliberately choose to undervote or they might choose between the candidates in a random fashion out of a perceived duty to vote. Disengaged voters may be very knowledgeable and care deeply about other races on the ballot; we simply assume that they are not interested in the CD 13 House race.

In these two hypothetical extremes, engaged voters would be expected to search the ballot for their preferred House candidate and would not likely be tripped up by a ballot design flaw. Moreover, if an engaged voter initially and accidentally misses voting for her preferred House candidate, when an electronic review screen warns her that she has not voted in the House election, she is inclined to make the added effort to go back and correct this oversight. Conversely, dis-

---

31In every other statewide race voted on in the 2006 midterm elections in Florida, the relationship in Sarasota County predicting vote share from other election results is exactly the same as the relationship outside of Sarasota County (these results to be presented in a forthcoming version of this paper).
engaged voters are not seeking out the House election on the ballot, they are consequently more likely to accidentally miss voting, and when they are warned they have made this error they are less likely to spend the effort to go back and correct an undervote. In this story, one about engaged and disengaged voters, the former are less likely to be suppressed the latter, more likely.

According to this way of thinking, at the extreme, if all suppressed votes were wholly disengaged voters, then intended votes would have been equally split between Buchanan and Jennings. Thus, if about 15,000 votes were suppressed, each candidate would have received an equal apportionment of about 7,500 each. Jennings would not pick up votes on her opponent, and Buchanan would still have won the election by 369 votes.

Let us now redefine the vote share of the Democratic candidate among votes cast for the Democratic or Republican candidates, that is excluding undervotes, as:

$$\text{Democratic Voteshare} = \frac{\text{Democratic Votes}}{\text{Dem. Votes + Rep. Votes}} \quad (2)$$

If the story of suppression owing to indifference occurs, an equal number of votes should be removed from Buchanan and Jennings. If Jennings were previously winning by some number of votes, say 100, then subtracting an equal number of votes from her and from her opponent will cause the vote share for the winning candidate to increase as a 100 vote lead, i.e., a smaller total number of votes cast will look like a greater ratio. If the earlier story about undervotes was instead correct and suppression was caused completely by accident, then votes would be removed from Jennings vote share in proportion to how votes were cast. Thus, under the “accident” story Jennings (or Buchanan) vote share, as calculated above, should not change when suppression occurs but under the “indifference” story, the winner in any precinct should seem to have a greater share of the cast ballots when indifferent voters are suppressed.

Figure 14 shows the relationship between Senate Democratic and CD 13 Democratic vote shares but now using this different measure of vote share that excludes undervotes. What we see
Figure 14: Democratic Vote Shares in Senate and CD 13 Races for Early and Election Day Voters, Normalizing by Valid Votes Cast

Note: Each red dot in the figure represents a Sarasota precinct and each blue dot a non-Sarasota precinct.
here is that the Sarasota distribution, and the red line, is now above the blue distribution and blue line. Thus the margin of victory is more dramatic in Sarasota precincts in the House race than would have been predicted given the relationship between the margin of victory in Senate and House races in other counties. This supports the theory that suppression did not occur completely randomly; instead, voters who were indifferent were more likely to be suppressed causing the margin of victory to increase.

4.1.3 Combining the Two Stylized Versions of Undervotes

The true process that occurred in Sarasota County is undoubtedly a mixture of our accident and indifference stories. Almost certainly some Sarasota voters were suppressed by the grouped race ballot design completely at random and completely by accident; these voters would be drawn randomly from the set of voters in a given precinct. However, with respect to the CD 13 race the iVotronic review screen that rescues some voters from being suppressed was probably more likely to rescue engaged voters than voters who were indifferent and did not want to spend extra effort voting in a race they had weak or no preference in. Thus some greater proportion of the suppressed vote was indifferent voters who might be expected to split their votes evenly between the candidates. This means that the observed margin of victory in suppressed precincts will be greater than would have been observed if the precinct had not been suppressed.

4.2 Allocating Suppressed Votes with Precinct Returns

Our approach to the problem of suppressed votes is to use available aggregated precinct data to estimate what the precinct vote totals in Sarasota County would have been if the ballot design and equipment there had been equivalent to those used in the other counties in CD 13. Note that in precincts in CD 13 which are outside Sarasota County we can observe all the relationships between results in other races and the result in CD 13 race. These relationships will help us predict how we would have expected the Congressional election in each Sarasota precinct to turn out, given that
we know how all the other races turned out in the same precinct.

   In precincts inside Sarasota County, we obviously do not observe the counter-factual answer we are seeking; thus, the Sarasota precincts cannot straightforwardly contribute information to our model of the relationships between the CD 13 race and other races on the ballot. However, these precincts do also contain some useful information. Although we do not know how the vote totals in such precincts would have turned out if there had not been any ballot design issues, we do know that, with respect to Buchanan and Jennings totals, these totals must be higher than the totals that resulted on the day of the election. That is, the ballot design issue we have discussed could only transfer intended votes from each candidate and turn them into undervotes. Therefore, although we can not observe intended votes, we know Buchanan and Jennings would have received *more* votes than he or she received, respectively, on election day.

   Our statistical model of Sarasota undervote reallocation makes two key assumptions:

   • *The relationships between electoral races are the same across all precincts in CD 13.* This is not an assumption that all districts appear or vote the same. Clearly some vote heavily for one candidate and some for another. Rather, we assume that the ability to predict one race given knowledge of all the other races applies inside Sarasota in the same way that it applies in non-Sarasota precincts.

   • *Both candidates, Buchanan and Jennings, would not have received fewer votes if the ballots were correctly designed than they did on the day with the flawed ballot design.* That is, the issue of race grouping (CD 13 with Florida governor) did not add votes to either candidate but only changed actual votes into undervotes.

   From these assumptions, we derive a statistical model assuming that the vote shares for candidates, as well as the proportion of voters intentionally undervoting, are additive-logistic-normally distributed as in Katz & King (1999). Full technical are in Appendix A.

   Our model gives us a probability density over all precinct results and thus cumulatively over all
election outcomes. We take one thousand random draws from this predicted density and calculate one thousand predicted election outcomes. From our random draws we can answer key questions such as, what fraction of Sarasota undervotes were suppressed votes of voters who intended to cast a ballot for a candidate, how would those votes have broken between Buchanan and Jennings, and how might that have influenced the election outcome. Additionally, and crucially, we can express our degree of confidence in each of these quantities.

Our results are summarized in Table 4. This table provides estimated election results among early and election day voters in Sarasota County. What is key 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 as, according our reallocation, votes move from the undervote to Buchanan and Jennings. The vote totals and pickups reported in Table 4 are the most likely, or expected, values. In addition, associated 90% confidence interval estimates provide a measure of uncertainty in our pickup totals.

Looking first at early voting, we can see that the confidence interval for the undervoting column is -4396 to -4685, so we can be 90 percent confident that there would have been between 4396 and 4685 fewer early voting undervotes if there had been no ballot design problem in Sarasota County. Our best guess is that 2436 suppressed voters intended to vote for Jennings and 2110 voters intended to vote for Buchanan in the early voting. Thus, Jennings would gain 326 votes over Buchanan among early suppressed voters.

In election day voting we can be 90 percent confident that Jennings would gain between 5002 and 5861 more votes in Sarasota had there been no ballot design issues. Similarly, we can be 90 percent confident that between 4244 and 5060 voters intended to vote for Buchanan were instead counted as undervotes.

Overall, what do these numbers mean? Looking to combined results, we can be 90 percent confident that *between 14,322 and 14,896 voters in Sarasota County were suppressed from voting*
Table 4: Summary of Allocation Results

<table>
<thead>
<tr>
<th></th>
<th>Jennings</th>
<th>Buchanan</th>
<th>Undervote</th>
</tr>
</thead>
<tbody>
<tr>
<td>EARLY VOTING:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estimated Vote:</td>
<td>16903</td>
<td>12943</td>
<td>866</td>
</tr>
<tr>
<td>Estimated Pickup:</td>
<td>2436</td>
<td>2110</td>
<td>-4546</td>
</tr>
<tr>
<td>Lower bound of 90% CI:</td>
<td>2176</td>
<td>1852</td>
<td>-4685</td>
</tr>
<tr>
<td>Upper bound of 90% CI:</td>
<td>2702</td>
<td>2362</td>
<td>-4396</td>
</tr>
<tr>
<td>ELECTION DAY VOTING:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estimated Vote:</td>
<td>45267</td>
<td>41123</td>
<td>2283</td>
</tr>
<tr>
<td>Estimated Pickup:</td>
<td>5421</td>
<td>4647</td>
<td>-10068</td>
</tr>
<tr>
<td>Lower bound of 90% CI:</td>
<td>5002</td>
<td>4244</td>
<td>-10310</td>
</tr>
<tr>
<td>Upper bound of 90% CI:</td>
<td>5861</td>
<td>5060</td>
<td>-9811</td>
</tr>
<tr>
<td>EARLY + ELECTION DAY VOTING:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Election Returns:</td>
<td>54313</td>
<td>47309</td>
<td>17763</td>
</tr>
<tr>
<td>Estimated Totals:</td>
<td>62171</td>
<td>54066</td>
<td>3148</td>
</tr>
<tr>
<td>Projected Pickup:</td>
<td>7858</td>
<td>6757</td>
<td>-14615</td>
</tr>
<tr>
<td>Lower bound of 90% CI:</td>
<td>7352</td>
<td>6267</td>
<td>-14896</td>
</tr>
<tr>
<td>Upper bound of 90% CI:</td>
<td>8383</td>
<td>7231</td>
<td>-14322</td>
</tr>
</tbody>
</table>

**Predicted probability of Jennings’s pickup > 369: 0.896**

*in the thirteenth district race.* Of these more than 14,000 votes, our best estimate is that Jennings would have received 7858 and Buchanan, 6757. This would cause Jennings to gain 1101 votes over Buchanan, thus overcoming her 369 vote deficit and changing the outcome of the CD 13 race.

However, it is important to realize that all of these numbers are simply best guesses. The truth in each case (early and election day, Buchanan and Jennings) might be higher or lower. From all 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 the election.
Crucially, we are interested in the fraction of the time the vote gain by Jennings is greater than 369. That is, we want to know the probability that if the suppressed votes in Sarasota had been recorded, Jennings would have won the election to the House of Representatives. This probability, according to our vote imputation model, is 89.6 percent.

### 4.3 Caveat to Precinct Allocation

While our precinct-level imputations provide reasonable results about what would have happened in Sarasota County in the absence of a ballot format problem, they are not without limitations. In particular, associated with them are several potential sources of bias, and the amount of information contained in aggregate precinct returns is not sufficient to determine definitively whether balloting problems resulted in Buchanan’s victory over Jennings by 369 votes. And, as mentioned above, our statistical model cannot pin down the exact mechanism that led to excess undervotes in the Attorney General and CD 13 elections when those elections were presented on the same ballot screens as the Florida Governor race.

Our models are predicated on two fundamental assumptions: (1) Net of any voting machine effects, precinct-level variation in early voting and election day undervote rates and support levels for candidates is similar across counties conditional on precinct-level absentee voting and voting in other races common to those counties. That is, we assume that, if elections were commonly administered across counties, knowing that a precinct was located in a given county would provide no additional predictive leverage on undervote rates or candidate support beyond what can be learned by observing other election results from that precinct. (2) There are no further substantial and systematic voting machine or ballot format effects beyond the grouping issues identified.

Violation of either assumption could bias our findings. However, in order to assess the validity of these assumptions, we have performed a number of statistical experiments in which we imputed election returns among CD 13 counties for offices in which there are no reasons to suspect ballot format problems. In those experiments, our imputed values closely approximated observed election
returns.

Assuming that our two key assumptions are not grossly violated, the precinct-level data we have described here pin down the magnitude of the undervoting problem in CD 13 to within plus or minus 300 undervotes. Nonetheless, the allocation of those undervotes to Buchanan and Jennings is not as precise as we would like given the closeness of the race, our assessment of the probability of Jennings winning the election absent balloting problems in Sarasota County is more sensitive to model assumptions and specification than we would like and, obviously, a confidence interval of plus or minus 1,000 votes net pickup for Jennings while small in percentage terms is large relative to the 369 vote difference that separates her and Buchanan.

4.4 Allocating Suppressed Votes with Ballot Data

Another way to estimate Sarasota’s excess early and election day undervotes in CD 13 to Buchanan and Jennings is to compare patterns of undervotes on Charlotte County and Sarasota County ballots (setting aside the problematic governor and attorney general races). 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, U.S. Senate, chief financial officer, commissioner of agriculture, Amendment 8, and the retention of Supreme Court Justice Lewis), we use the simple logit models in Table 5, column one.\(^{32}\) Because only 7,326 of Charlotte’s election day and early voters lived in CD 13 and only 174 of those undervoted, we cannot accurately 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 do allow for a good deal of flexibility. 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. The estimated parameters of this model are shown in the second column of

\(^{32}\)We are still trying various alternative specifications of our logit model, but thus far the results do not appear to depend importantly on the exact specification used.
Table 5: Logit Models Predicting Voting and Candidate Choice in CD 13

<table>
<thead>
<tr>
<th>Variable</th>
<th>Charlotte (vote=1)</th>
<th>Sarasota (vote=1)</th>
<th>Sarasota</th>
<th>Dem. (vote=1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>US Senate</td>
<td>1.99 (0.42)</td>
<td>0.55 (0.09)</td>
<td>2.47 (0.03)</td>
<td></td>
</tr>
<tr>
<td>CFO</td>
<td>1.24 (0.55)</td>
<td>0.26 (0.12)</td>
<td>2.10 (0.05)</td>
<td></td>
</tr>
<tr>
<td>Agriculture Commissioner</td>
<td>1.17 (0.25)</td>
<td>0.97 (0.04)</td>
<td>-1.99 (0.07)</td>
<td></td>
</tr>
<tr>
<td>Amendment 8</td>
<td>0.38 (0.33)</td>
<td>0.16 (0.04)</td>
<td>0.31 (0.04)</td>
<td></td>
</tr>
<tr>
<td>Supreme Court: Lewis</td>
<td>-0.15 (0.49)</td>
<td>-0.10 (0.05)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amendment 8 (No) × Supreme Court</td>
<td>0.22 (0.53)</td>
<td>0.27 (0.06)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>US Senate × CFO</td>
<td>0.23 (0.56)</td>
<td>0.53 (0.12)</td>
<td>-0.22 (0.05)</td>
<td></td>
</tr>
<tr>
<td>US Senate × Agric. Comm.</td>
<td></td>
<td></td>
<td>-0.46 (0.07)</td>
<td></td>
</tr>
<tr>
<td>US Senate × Amendment 8</td>
<td></td>
<td></td>
<td>-0.02 (0.05)</td>
<td></td>
</tr>
<tr>
<td>Agric. Comm. × Amendment 8</td>
<td></td>
<td></td>
<td>-0.46 (0.06)</td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-0.86 (0.41)</td>
<td>-0.70 (0.09)</td>
<td>-2.70 (0.02)</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>7,326</td>
<td>119,897</td>
<td>102,073</td>
<td></td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-704.45</td>
<td>-483,329</td>
<td>-29,857</td>
<td></td>
</tr>
<tr>
<td>Geometric Mean Probability</td>
<td>0.90</td>
<td>0.67</td>
<td>0.75</td>
<td></td>
</tr>
</tbody>
</table>

To estimate how the Sarasota CD 13 undervotes that should have been votes would be 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. That is, we assume that once we observe the other choices made by a voter, there is no additional information about that voter’s CD 13 candidate preference to be gained from knowing that she undervoted in that 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 5, 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 Sarasota voter had voted using Charlotte’s machines? We assume the model governing Charlotte voters’ undervoting would apply to Sarasota voters if they had voted in Charlotte.\textsuperscript{33} That is, if we apply the Charlotte undervote model to a Sarasota voter, we would get a reasonable estimate of the probability that this voter would have participated in the CD 13 election if she had been voting in Charlotte. Thus, for any voter $i$ we can consider two events, the event $C_i$ in which she undervotes using a Charlotte machine and the event $S_i$ in which she undervotes using a Sarasota machine. We assume that any voter who undervotes in Charlotte would undervote if voting in Sarasota, thus the event $S_i$ obtains whenever the event $C_i$ obtains; 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 Charlotte machines. By the definition of conditional probability, $P (C_i|S_i) = P (C_i \cap S_i) / P (S_i)$, which in this case is simply $P (C_i) / P (S_i)$. Estimates of $P (S_i)$ and $P (C_i)$ are obtained from the logit regressions described above. Given these estimates of $P (C_i|S_i)$ 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 yields the reallocation in Table 6.\textsuperscript{34}

\textsuperscript{33}Among other thing this assumption requires that there were no other layout or machine issues in Sarasota or Charlotte that we are not accounting for.

\textsuperscript{34}Confidence intervals are obtained through the use of a bootstrap oversampling the data used to produce the models shown in Table 5.
Table 6: Summary of Allocation Results from Ballot Data

<table>
<thead>
<tr>
<th></th>
<th>Jennings</th>
<th>Buchanan</th>
<th>Undervote</th>
</tr>
</thead>
<tbody>
<tr>
<td>Election returns</td>
<td>54313</td>
<td>47567</td>
<td>17825</td>
</tr>
<tr>
<td>Estimated Vote:</td>
<td>63327</td>
<td>53365</td>
<td>3012</td>
</tr>
<tr>
<td>Estimated Pickup:</td>
<td>9014</td>
<td>5798</td>
<td>-14813</td>
</tr>
<tr>
<td>Lower bound of 90% CI</td>
<td>8853</td>
<td>5641</td>
<td>-14529</td>
</tr>
<tr>
<td>Upper bound of 90% CI</td>
<td>9169</td>
<td>5956</td>
<td>-15101</td>
</tr>
</tbody>
</table>

**Predicted probability of Jennings’s pickup > 369: \approx 1**

As compared to our precinct-level results, here we find much 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 3068 and 3359 votes in favor of Jennings with a best guess of 3215. This is roughly two-thousand more votes than is implied by our precinct-level reallocation model. Note the precinct-level confidence interval was plus or minus roughly 975 votes. Here the confidence interval is only plus or minus 146 votes. Thus, through the use of ballot-level data, we improved the precision of our estimates nearly seven-fold.

5 Conclusion

The motivation for this paper was the unusually high undervote rate in the 13th Congressional District race in Sarasota County, Florida. This race featured Republican Vern Buchanan versus Democratic Christine Jennings, and we have explained the undervote rate by drawing on variance in ballot formats across counties and types of voters (election day, early, 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 governor’s race. The problematic grouping method 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 generally find high undervote rates. The only times we see a problematic grouping without a large undervote is when the associated race is not at all close.

Thus, what looks on the surface like a Sarasota County issue is in fact more general. Moreover, like the butterfly ballot format used in Palm Beach County in the 2000 presidential race, it appears that the Sarasota ballot format in the 2006 midterm elections was pivotal to the winner of tightly-contested 13th Congressional District race in Florida. We estimate that, had Sarasota used a ballot format akin to those in neighboring counties, with probability 0.9 Jennings would have beaten Buchanan.

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 candidate 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 cannot address these questions at this time, but work in the vein of Herrnson, Abbe, Francia, Bederson, Lee, Sherman, Conrad, Niemi & Traugott (2005), Herron & Lewis (2006), and Herrnson et al. (2006) may be applicable here.

Second, and as mentioned previously on several occasions, because this paper presents 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 undervote and the Charlotte and Lee undervotes? Yes, this is possible. However, it seems to be a significant stretch of the imagination: 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 simple and conservative implication of our main finding: iVotronic touchscreen voting systems should not combine important races on the same voting page. Regardless of why exactly combining races is a problem, this proposal seems likely to avoid it.

A Appendix

This technical appendix explains the procedures used to allocate Sarasota County’s election day and early voting undervotes using precinct-level data.

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) by each voting method (early voting, election day voting, and absentee voting). Let \( V_{i,c}^{e,m} \) represent the vote share in precinct \( i \), race \( c \), by method \( m \in \{ \text{early, e.d., abs.} \} \), 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,

\[
V_{i,c} \in [0, 1] \quad \forall \ i, c,
\]

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

\[
V_{i,R} + V_{i,D} + V_{i,U} = 1 \quad \forall \ i.
\]

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.

\[35\text{We discard votes for the minor candidates who contested the gubernatorial race and for all write-ins, and we also discard overvotes. All of the discarded votes are negligible totals in the races studied here.}\]
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: transformations are:

\[
Y^{c,m}_{i,R} = \ln \left( \frac{V^{c,m}_{i,R}}{V^{c,m}_{i,D}} \right) \\
Y^{c,m}_{i,U} = \ln \left( \frac{V^{c,m}_{i,U}}{V^{c,m}_{i,D}} \right)
\]  

(5)

(6)

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 \( V^* \) and \( Y^* \) be the latent vote shares and transformations that would have been observed if there were no ballot flaw in that race. Clearly, \( Y^*_{i,CD} = Y^*_{i,CD13}, \forall s_i = 0 \), but elsewhere, \( Y^*_{i,CD13} \) is unobserved. However, many other races and election methods are observed in Sarasota County and its neighbors.

The eight election methods that we include in our imputation model are listed in Table 7. The vote shares from each method are transformed into two log vote ratios for sixteen fully observed variables. We omit the gubernatorial 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 are included as forecasting variables as these ballots are paper and not subject to any of the proposed mechanisms for the CD 13 ballot failure. We include all precincts in CD 13 from Charlotte, Hardee, Manatee, and Sarasota.
Race | Methods
--- | ---
Congressional District 13 | Absentee
United States Senate | Early, Election Day, Absentee
Agricultural Commissioner | Early, Election Day
Chief Financial Officer | Early, Election Day

Table 7: Races and Voting Methods used in Imputation Model

Note: The Sarasota County absentee ballots in the CD 13 race can be used to forecast early and election day totals as the former did not suffer from any of the touchscreen formatting problems.

Counties except the small number of precincts in Manatee that were split between the CD 11 and CD 13 races. At the time of this writing Desoto County data disaggregated by election method is not available. 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 7 as $X$’s, although latter were constructed by equation 5.

If we consider $\hat{Y} = Y$ for all $s_i = 0$ and all observations of $\hat{Y} = Y$ within Sarasota as completely missing data (with observed covariates $X$), then our model has the same architecture as any conventional multivariate normal imputation model (Schafer 1997, Donald B 1987, 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.\textsuperscript{36} Ignoring covariance between early and election day returns, within the CD 13 race the imputation

\textsuperscript{36}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:

\[
(Y_{RD}^*, Y_{UD}^*) \sim f_{\text{bivariate normal}}(\mu_{RD}, \mu_{UD}, \Sigma)
\]

\[
\mu_{RD} = X\beta
\]

\[
\mu_{UD} = X\gamma
\] (7)

where X consists of 16 log vote ratios from the elections in Table 7, plus a constant vector. Imputations of \(Y^*\) from this model will give us 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 ballot had been equivalent to the ballots 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 allocation and the vote 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. In the limit, all under votes could break for Buchanan or Jennings. Thus, the observed vote shares in Sarasota give us a series of bounds on the latent vote shares:

\[
V_{i,U}^* \leq V_{i,U} \quad \forall i : s_i = 1
\] (8)

\[
(V_{i,R} + V_{i,U}) \geq V_{i,R}^* \geq V_{i,R} \quad \forall i : s_i = 1
\] (9)

\[
(V_{i,R} + V_{i,U}) \geq V_{i,D}^* \geq V_{i,D} \quad \forall i : s_i = 1
\] (10)
Figure 15: Ternary plot of Hypothetical Sarasota and Manatee County Precincts

Note: The two precincts are denoted $V_m$ (Manatee County) and $V_s$ (Sarasota County). In $V_s$, which suffers from the ballot flaw described in the body of this paper, the latent vote share must be closer to the Buchanan and Jennings corners and thus can only fall in the red region. The vertical dashed line defines which candidate won the plurality.

As an illustration, imagine reported results in two hypothetical precincts, $V_s$ in Sarasota County and $V_m$ in neighboring Manatee County. In Manatee, where there was no ballot format problem, $V_m^* = V_m$. In the Sarasota precinct, assume Jennings received 45 percent of the vote and Buchanan 35 percent, and assume that the remaining 20 percent of ballots were CD 13 undervotes. In the ternary plot in Figure 15, this precinct is represented by the point $V_s$. The region shaded in red represents all points that are closer to both bottom vertices than $V_s$ is, thus the set of possible election results if the intended votes were counted. This red region is the space where $V_s^*$ might be located. These bounds on $V^*$ imply a set of bounds on $Y^*$. The most straightforward is:

---

37 The vertical dashed line separates points closer to Buchanan from points closer to Jennings. Neither candidate has received a majority in this precinct, but Jennings currently has a plurality. Buchanan could still win the precinct plurality if enough undervotes fell his way, as the red region crosses this line.
\[ Y_{i,UD}^* \leq Y_{i,UD} \] (11)

Additionally, if we knew the true undervote, \( V_{i,U}^* \), we could define the functions:

\[
Y_{i,RD}^+(V_{i,U}) = \ln \left( \frac{V_{i,R} + (V_{i,U} - V_{i,U}^*)}{V_{i,D}} \right) 
\] (12)

\[
Y_{i,RD}^-(V_{i,U}) = \ln \left( \frac{V_{i,R}}{V_{i,D} + (V_{i,U} - V_{i,U}^*)} \right) ,
\] (13)

which provide bounds on

\[
Y_{i,RD}^+(V_{i,U}) \geq Y_{i,RD}^* \geq Y_{i,RD}^-(V_{i,U})
\] (14)

We simplify these functions to their limiting values as \( V_{i,U} \to 0 \)

\[
Y_{i,RD}^+ = \ln \left( \frac{V_{i,R} + V_{i,U}}{V_{i,D}} \right) 
\] (15)

\[
Y_{i,RD}^- = \ln \left( \frac{V_{i,R}}{V_{qq_{i,D}} + V_{i,U}} \right) 
\] (16)

Using equation 11 and the simplified form of equation 14 we can set the limits of integration for:

\[
L(\beta, \gamma, \Sigma|S_i = 1) = \int_{-\infty}^{Y_{i,RD}^+} \int_{Y_{i,RD}^-}^{Y_{i,UD}} p_{bvn}(r, s|X_i\beta, X_i\gamma, \Sigma) \delta s \delta r
\] (17)

while the precincts outside Sarasota more straightforwardly contribute:

\[
L(\beta, \gamma, \Sigma|S_i = 0) = p_{bvn}(Y_{i,RD}, Y_{i,UD}|X_i\beta, X_i\gamma, \Sigma)
\] (18)
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 the 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 1000 datasets.

Our imputation model as previously discussed is multivariate normal in the space of the $Y$s. The quantities of interest to us, 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 transformation is:

$$
\tilde{V}_{i,U} = \frac{\exp(\tilde{Y}_{i,UD})}{W_i} \quad (19)
$$
$$
\tilde{V}_{i,D} = 1/W_i \quad (20)
$$
$$
\tilde{V}_{i,R} = \frac{\exp(\tilde{Y}_{i,RD})}{W_i} \quad (21)
$$

where

$$
W_i = 1 + \exp(\tilde{Y}_{i,UD}) + \exp(\tilde{Y}_{i,RD}) \quad (22)
$$

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 $\tilde{V}_{i,U} \leq V_{i,U}$, $\tilde{V}_{i,D} \geq V_{i,D}$, and $\tilde{V}_{i,R} \geq V_{i,R}$. Following Honaker, Katz & King (2002), we rejection sample each vector $\tilde{V}_i$ 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 as censored values.
References


Freeman, Steven F. & Joel Bleifuss. 2006. *was the 2004 presidential election stolen?* New York: Seven Stories Press.


