Political Research Quarterly
Volume 63 Number 1
March 2010 129-142

10.1177/1065912908324201
http://prq.sagepub.com
hosted at
http://online.sagepub.com

Losing Fewer Votes

The Impact of Changing Voting Systems on Residual Votes

Michael J. Hanmer
University of Maryland, College Park
Won-Ho Park
University of Florida, Gainesville
Michael W. Traugott
University of Michigan, Ann Arbor
Richard G. Niemi
University of Rochester, New York
Paul S. Herrnson
University of Maryland, College Park
Benjamin B. Bederson
University of Maryland, College Park
Frederick C. Conrad
University of Michigan, Ann Arbor

Problems in the 2000 presidential election, especially in Florida, initiated a large-scale shift toward new voting technology. Using cross-sectional and longitudinal data, we report on the effects of changes in voting systems in Florida and Michigan. The variety of initial conditions and the numerous changes make these excellent case studies. We find that reforms succeeded in reducing the residual vote. Every change from old to new technology resulted in a decline in residual votes that was significantly greater than in areas that did not change voting equipment. The percentage of residual votes in the 2004 presidential race in localities that changed voting systems was well under 1 percent, representing a 90 percent reduction in error in Florida and a 35 percent reduction in Michigan. We run these analyses separately for undervotes and overvotes. Using ecological-inference techniques, we investigate the persistence of residual votes when technology changed and find very little persistence.

Keywords: *new voting technology; residual votes; turnout; voting machines*

A major concern arising out of the 2000 presidential election was the number of potential votes that were never cast and the number of votes that were not or could not be counted—what has come to be known as the "residual vote" (Caltech/MIT Voting Technology Project 2001).[1] Across the country, roughly 2 percent of all those who went to the polls in 2000 did not have a vote counted for president—some because they chose to abstain, others because they did not understand how to cast a valid vote or they had problems with the voting systems themselves.

The problems in the 2000 election prompted the states to explore the possibility of upgrading voting systems and procedures as well as to change election laws and administration. They also prompted the federal government to respond, via the Help America Vote Act (HAVA) of 2002 (for a review of reform efforts, see Palazzolo and Ceaser 2005). Attention focused especially on punch-card systems, in use at the time of the 2000 election by about one-third of U.S. voters. Residual votes were more than half again as high among voters using that method compared to those who used the best of the other methods. Based

Authors' Note: This work was supported in part by National Science Foundation grant number 0306698 and Carnegie Corporation grant number D05008. We acknowledge the research assistance of Chris Thomas and Tim Hanson from the State of Michigan Elections Division as well as Laura Lee.

on these estimates, Ansolabehere and Stewart (2005) estimated that if all jurisdictions using punch cards in 2000 had instead used optical scan ballots, a half-million more voters would have had their votes counted. Especially significant was the belief that the presidential election itself was influenced by a well-intentioned but poorly designed punch card known as the "butterfly ballot" (Wand et al. 2001).

Yet, the fact is that we are only now gaining a good understanding of the effects of voting systems on casting and counting votes. Analyses to date are necessarily exploratory, as there has only been one additional presidential election to investigate, an important factor because researchers want to hold constant the office at the top of the ballot that produces a high-stimulus campaign and associated high levels of turnout. Previous research has shown that in contrast to the results for punch-card systems, differences in residual-vote rates between paper, optical scan (OS), and direct recording electronic systems (DREs) have not been so consistent, leaving uncertainty about the relative merits of other systems for voters' abilities to cast their ballots without error. Nor do we know the full extent to which the 2000 election and subsequent reform efforts educated voters and election officials. Public education and media coverage, even in the absence of new voting technology, might make voters more sensitive to potential problems, and hence, make residual votes less likely, even in the absence of new voting technology (Stewart 2006).

The research reported here furthers our understanding of new voting technology by assessing the impact of changes in voting machinery that occurred between 2000 and 2004 in two states, Florida and Michigan, in both cross-sectional and over-time analyses. Because of the decentralized nature of their election administration, the presence of different initial conditions, and numerous kinds of changes, examination of these two states provides an excellent opportunity to test the effects of changes in voting technology.[2] We do this by examining small-unit election returns with a focus on longitudinal data from the same jurisdictions across time,[3] using descriptive statistics and reliable ecological-inference techniques in a carefully controlled sequence.

Most directly, the research reported here speaks to the potential influence of voting systems on election outcomes. We know from recent experience that in close elections, voter errors can alter the outcomes (Wand et al. 2001; Mebane 2004; Frisina et al. 2008). We find that the reforms were successful. In both states, every change from old to new technology was accompanied by a decline in the percentage of residual votes, thus making it less likely that the winner could be decided by errors of this sort. Although declines in residual votes occurred even in areas that did not change their voting equipment, the evidence shows that technological change mattered: the declines in jurisdictions that introduced new equipment were significantly greater than in jurisdictions where voting equipment did not change. In absolute terms, the net effect was substantial, reducing residual votes in the 2004 presidential race to well under 1 percent in the two states (a 90 percent reduction in Florida and a 35 percent reduction in Michigan). Since it is widely believed that problems with voting systems reduce confidence,[4] lead to questions about the legitimacy of any given election outcome and the system as a whole (Saltman 2006), and could lead people to disengage from politics (Caltech/MIT Voting Technology Project 2001), these results are rather encouraging.

Our results also highlight the need for more and better data. While political scientists are never shy to make such a claim, the need is becoming more widely recognized. In leading the charge for the creation of a Democracy Index, Gerken (forthcoming), a legal scholar, laments the dearth of data and argues that if armed with more concrete information about the performance of their local and state election-administration systems, citizens and election officials will be able to make the push for the sort of large-scale reforms necessary to improve elections.[5] It is also true that the very act of adopting new technology implies more frequent and rapid changes in such technology over time, in terms of both hardware and software. In sum, this analysis is critical for the evaluation of progress toward an electoral process that loses fewer votes and thus restores voters' confidence in the functioning of the electoral system.

Literature Review and Hypotheses

When new voting systems were introduced between 2000 and 2004, the reduction of lost votes was a primary concern. Newer technologies were adopted with the explicit intention of reducing or eliminating overvotes, spoiled ballots, and to the extent that they were unintended, undervotes (see Caltech/MIT Voting Technology Project 2001). System type has been documented as a key factor in the spate of studies showing that punch cards (especially the Votomatic variety) consistently yield higher residual-vote rates than other voting systems (Bullock and Hood 2002; Knack and

Kropf 2003; Buchler, Jarvis, and McNulty 2004; Ansolabehere and Stewart 2005). Thus, the focus is now on OS ballots and DREs as preferred systems and on paper ballots insofar as they provide useful baseline data. Information about punch cards is also useful for historical comparisons to assess improvement, even as the evidence about problems with them has been so strongly negative that they are being phased out completely. Likewise, lever machines are being phased out because they have not been manufactured since 1982, so information about their performance is primarily of historical interest.

The characteristics of various voting systems provide insights into differences in their residual-vote rates and suggest a number of specific hypotheses. Tomz and Van Houweling (2003, 48, Table 1) provide the basis for this reasoning, and we adopted their typology. The impact of system type is clearest with respect to overvotes. Two types of systems—DREs and lever machines—make it impossible to overvote, as they are programmed or mechanically designed so that one cannot vote for more than the number of candidates to be elected to a given office. OS systems are likely to reduce overvotes as well, but voters often fail to check their ballots at the precinct level, where mistakes could be corrected, and ballots may be transported to a central location for counting. Even if they are counted locally, OS equipment may be programmed to warn voters about overvotes, but it does not prevent them from casting ballots that contain them. Punch-card systems provide no check, and therefore, not even a warning about overvotes.

The potential for unintended undervotes also varies by system type. Most DREs require page-by-page and sometimes office-by-office movement down the ballot and use color coding to distinguish parts of the ballot, making it harder to overlook an office. And just before casting the ballot, review screens also call attention to uncast votes. Lever machines, OS equipment, and paper ballots all allow voters to scan the ballot quickly and to vote for offices in random order, possibly missing some offices in the process. Paper systems that are not checked at the precinct provide no means of calling the voter's attention to missed opportunities. Lever machines may suffer from how they are laid out; research shows that voters sometimes fail to see ballot propositions, which tend to be located off to the side of the machine or above the most prominent offices (Roth 1994, 1998). Punch cards often do not permit a voter to see the entire ballot at once and can also be subject to confusing layouts, as in the case of the "butterfly ballot."

Finally, correcting mistakes (and therefore, avoiding a spoiled ballot) is simple with DREs and lever machines. With the former, one either simply presses an alternative candidate or "unvotes" one's previous choice and then selects a new one. With the latter, one unvotes by lifting the lever and then depressing a new lever. With OS and paper ballots, correcting a mistake may "only" involve an erasure, or it may mean getting a new ballot and starting over. A mispunched card clearly requires starting over. These comparisons suggest that residual votes (overvotes, undervotes, and spoiled ballots) should be least frequent when voters use DREs.

Empirical evidence about these propositions—aside from the consensus about punch cards—is mixed and not entirely supportive. Ansolabehere and Stewart (2005, 378, 380) find relatively high rates of residual votes associated with DREs, at least for president. In contrast, Brady et al. (2001, 29, 34) find no significant difference in residual-vote rates between OS and DRE systems, suggesting that newer DREs used in 2000 may be the explanation. Likewise, Tomz and Van Houweling (2003, 55) find that DREs are associated with equivalent or lower residual-vote rates. Evidence about lever machines suggests that they may result in higher rather than lower residual-vote rates than OS systems, especially for offices below the presidency (Ansolabehere and Stewart 2005; Brady et al. 2001, 34-37). Kimball and Kropf (2005) add that residual-vote rates within a single method (paper ballots) vary because of ballot design.

An additional consideration is that the number of residual votes may decline because of increased awareness of the problem of lost votes as well as from new technology (Stewart 2006). As a consequence of controversy and subsequent media visibility, voters may pay more attention to the task at hand, thereby reducing overvotes where they are still possible as well as voting more often for lower level offices and for ballot propositions. Likewise, the training of both poll workers and citizens that goes along with the introduction of new technology not only may affect the use of that technology but may spill over into other areas, reducing the number of residual votes even where no system changes have occurred (Singley and Anderson 1988).

On the basis of the extant literature, we formulated a number of hypotheses that can be appropriately tested with data from Florida and Michigan. These hypotheses, and the specific reasoning behind each one, are as follows:

Hypothesis 1: Overall, the residual-vote rate will decline between 2000 and 2004 even where technology was unchanged; it will decline more in Florida than in Michigan.

Rationale. Changes in voting technology have occurred largely in a context of extensive media coverage and high levels of information as well as heightened concern about voting procedures and outcomes, especially in Florida.[6] The organization VotersUnite! assembled information on media reports of voting problems across the United States, and it showed many more reports in Florida than in Michigan in 2004 and 2006 (eighty-one compared to seven in 2004 and sixteen compared to two in 2006). The high level of concern in Florida is reflected in the 2004 exit poll that asked how confident voters were that the votes in their state would be counted accurately. Across the nation, 91 percent indicated that they were "very" or "somewhat" confident that the count would be accurate, compared to 95 percent among those interviewed in Michigan. But in Florida, there were only 74 percent who felt this way, the lowest in the nation.[7]

Hypothesis 2: Residual-vote rates will be lower for DREs than for OS systems.

Rationale. Through computerization, DREs can prevent some types of errors and reduce others, leading to a reduction in residual votes. With respect to prevention, DREs are programmed with built-in features that simply do not allow overvotes. OS systems can provide warnings about overvotes, but voters can still cast a ballot with overvotes on an OS system. With respect to the reduction of errors, DREs are programmed to lead the voter through the ballot a race or a page at a time, which should reduce the number of undervotes. They also have review screens that shade in different colors races in which a vote was not cast, signaling to voters when they did not cast a vote in a race. While OS ballots can be reviewed easily, more undervotes are likely on these systems because (1) there is not a formal stage at which the voter is prompted to review the ballot and (2) the optical scanners are not likely to be programmed to provide feedback about undervotes, meaning that races without any marks or races with marks that are not dark enough to be read might not be noticed by the voter.[8]

Hypothesis 3: The consistency in residual voting rates from 2000 to 2004 will be low, but it will be higher when an "older" voting system (e.g., lever system) was not changed than when a switch was made to "new" technology (e.g., DRE or OS), and it will be higher in Michigan than in Florida.

Rationale. Residual voting for president typically occurs because of carelessness or a misunderstanding of a particular ballot format. Heightened awareness of voting problems was especially prevalent in Florida in 2004 as a result of the problems four years earlier and the continuing media coverage of attempts to resolve them.

Research Methods and Data

Testing these hypotheses requires more than a simple comparison of jurisdictions in which various kinds of voting technologies are used, although such comparisons are important. Previous research tells us that a good deal of the variation in outcomes such as overvoting and undervoting is explained by factors peculiar to the individual jurisdiction. Demographic characteristics and their interaction with voting technologies are important insofar as racial, educational, and other differences account for significant variations in rates of residual voting (Darcy and Schneider 1989; Brady et al. 2001; Tomz and Van Houweling 2003; Buchler et al. 2004). Ballot features and ballot length, the salience of individual races, and the quality of election administration also play a role (Nichols 1998a, 1998b; Bullock and Dunn 1996; Kimball, Owens, and Keeney 2004). Ansolabehere and Stewart (2005) convincingly demonstrate the significance of local variations, specifically at the county level, as well as one important way to control for such factors using a fixed-effect regression that includes dummy variables for the nearly 2,000 counties from which their data are drawn.

Our approach relies on both cross-sectional and longitudinal data. Using cross-sectional analysis within a given state permits us to observe residual votes for a variety of offices and ballot propositions while, as much as possible, holding the ballot constant. We are not, of course, holding the ballot exactly constant, because of differences from place to place in lower level offices and ballot issues. But we are holding constant the presidency, prominent (U.S. Senate, governor) and not-so-prominent (judicial offices especially) statewide offices, and statewide ballot propositions such as referenda and bond issues. For these offices

and issues, both the presence of an election and the specific candidates are held constant. Since individual ballots are rarely available for analysis and county-level data provide too few cases to study system transitions, we use the smallest-level aggregate data possible. In Florida, this meant precincts because the redistricting changes were relatively easy to deal with. In Michigan, we used townships and cities to facilitate comparisons over time, because there was a sufficient number to study system transitions and the boundaries were stable.

Longitudinal data represent "natural experiments" that occurred when jurisdictions introduced new voting systems, creating a situation akin to a laboratory experiment in which prior behavior can be observed (voting under the old system), an intervention is implemented (the new system is put in use), and new behavior is observed (voting under the new system). The use of the same (or sometimes similar, though not identical) units over time is a way of approximating the condition of random assignment in a laboratory experiment that ensures that "all other things are equal." It serves much the same function as controlling statistically for county and other neighborhood characteristics. To be sure, it is not exactly the same as random assignment, because factors such as the local tax base or public interest in solving administrative problems can affect who adopts change and when. But in this approach, observations are made over a short period of time—four years, or one presidential election, in our case—to ensure that relatively few changes have occurred other than the technology used for voting.

Data

Our analysis focuses on Florida and Michigan, two states that use a wide range of voting technology. We pursue the analysis both cross-sectionally and over time in those available units where boundaries were unchanged between the 2000 and 2004 elections or where the boundary changes could be tracked such that the precincts could be merged to ensure that we were comparing the same place over time. Some of these requirements were complicated by the intervening decennial census. Florida has the additional advantage that some counties made separate undervote and overvote data available at the precinct level. Florida has a decentralized system of election administration with considerable variation in the types of voting systems used; in the past, this meant that punch cards and OS ballots were used heavily, and a few counties relied on paper ballots and lever machines. By 2004, all counties with punch cards, paper ballots, or lever machines had changed to either OS systems or DREs. A number of counties stayed with OS systems for the entire time period. We selected counties with the hope of capturing variation in system transitions.

The decentralization of election administration meant that we had to approach each county individually to obtain precinct-level data. Data could be obtained in electronic format in eleven counties and matched over time in seven, including Miami–Dade and Palm Beach, the two largest counties. For 2004, these data cover 2,559 precincts, or roughly 38 percent of the state total.[9]

Michigan has the most decentralized system of election administration in the United States in terms of the number of election administrators; townships and cities have historically been responsible for evaluating and selecting their own voting technology. In response to HAVA, however, Michigan developed a plan whereby every jurisdiction would adopt some form of optical character recognition (OCR) or OS system by 2006. In 2000 and 2004, most Michigan jurisdictions used OS systems or punch cards, though quite a few still used lever machines and paper ballots; by 2004, the movement toward OS systems was well under way, with considerable reduction in the number of jurisdictions using other technologies. For Michigan, we were able to collect data for all of the 1,522 townships and cities. Because some jurisdictions use different types of voting systems for polling place and absentee voting but do not report the totals separately, however, we analyze 1,341 (or 88 percent) of these units.[10]

Several points need to be made about the data. First, comparisons in Florida were complicated by the intervention between the 2000 and 2004 elections of the decennial census and the resulting redistricting that led to the redefinition of precinct boundaries. While this is not an issue for cross-sectional analyses, it is a significant complication for longitudinal analyses. When possible, we collected information about precinct changes in the counties we analyzed and adjusted units to account for such changes. When a 2000 precinct was subdivided for the 2004 election, we recombined the parts. When combinations of precincts were formed for the 2004 election, we created an artificial, combined precinct for 2000 that matched.[11] For Michigan, we use data at the town and city level, the common unit of election administration in the state, as opposed to the counties in Florida; thus, changes in precinct boundaries did not introduce complexities.[12] In the case of both Florida and Michigan, the effective

Table 1
Residual Votes in the 2000 and 2004 Elections in Florida and Michigan, by Type of Voting Equipment Used (Percent)

	President			U.S. Senate			Number of Cases	
Equipment	2000	2004	Change[a]	2000	2004	Change[a]	2000	2004
Florida								
Lever	0.3			16.7			39	
PC	6.3			7.2			1,508	
OS (precinct)	0.8	0.3	–0.5	2.2	1.6	–0.6	305	971
OS (central)	4.6			4.1			147	
DRE		0.6			4.1			1,588
Total[b]	5.2	0.5	–4.8	6.4	3.2	–3.3	1,999	2,559
Michigan								
Paper	2.2	1.6	–0.6	6.5	[c]		79	19
Lever	1.7	1.0	–0.6	6.0			210	116
PC	2.0	2.0	0.0	3.3			342	230
OS (precinct)	0.9	0.6	–0.3	2.3			671	961
OS (central)	1.7	0.6	–1.1	4.4			39	15
Total[b]	1.4	0.9	–0.5	3.5			1,341	1,341

Note: Selected precincts (Florida) or townships (Michigan). PC = punch card; OS = optical scan; DRE = direct recording electronic system.
a. Change calculated before rounding.
b. Weighted by the number of precincts or townships using each type of equipment.
c. There was no U.S. Senate election in Michigan in 2004.

N for analysis was appropriate for detecting the small magnitude and shifts over time that we expected to observe (see the exchange between Miller 2005a and 2005b and Alvarez, Ansolabehere, and Stewart 2005). Second, while there has been recent consolidation in the voting-system industry, there are still more than one hundred different models of voting systems currently available, counting physical devices and the available software options used to operate them. This has meant that, like other analysts, we group voting systems by type of system rather than by make and specific model (for a similar approach, see Ansolabehere and Stewart 2005).

Results

Residual-Vote Rates by System Type in 2000 and 2004

We begin by looking at the percentage of residual votes in 2000 and 2004 by voting system type in a cross-sectional analysis that shows that both Florida and Michigan experienced results that looked like other jurisdictions that have been analyzed for these same elections. In Florida, we report the rates for both president and U.S. Senate.[13] As expected, given the shift away from punch cards, the residual-vote rate was considerably lower in 2004 for both races (Table 1).[14] Also interesting is that the rate of residual votes declined even if analysis is limited to the comparison of precincts using OS systems with precinct counting in each year. In 2000, the residual-vote rate for president was slightly lower than the county average observed for OS systems across the country; in the Senate race, it was half a percentage point lower (Brady et al. 2001, 29, 36). Despite these lower starting points, the rate dropped by about 60 percent and 30 percent, respectively. This is a first indication that changes over time may have occurred because of higher levels of information and concern (Hypothesis 1). As expected, residual-vote rates were higher when optically scanned ballots were counted at a central location rather than at the precinct, where voters slide the ballot in themselves and can be warned about overvotes.

Hypothesis 2, on the other hand, is contradicted by these results. The percentage of residual votes with DREs was twice that of the OS systems in 2004 presidential voting (statistically significant at $p = .000$). In Senate voting in 2004, residual voting rates were greater for DREs by a margin of more than 2.5

to 1 ($p = .000$). For Senate voting, we can speculate that problems navigating the DRE interface surfaced.[15] That is, although the built-in features found on DREs should lead to a reduction in residual votes, only the overvote protection is effective for all voters. With respect to undervotes, it is possible that some voters accidentally double tapped the screen while pressing the area to advance from the presidential race to the Senate race, thus placing them beyond the race for Senate. To the extent that such voters do not pay attention to the review screen or are not sure how to go back from the review screen to races that they did not vote on, undervotes will result. Recent evidence from the 2006 race for Florida's thirteenth congressional seat (see Frisina et al. 2008) is consistent with the conclusion that a sizable proportion of the voters using the new DREs did not use the review screen or did not know how to navigate back through the ballot to correct unintentional undervotes.[16] For presidential voting, an explanation is even less clear, though it might have to do with the difficulty of formatting the large number of candidates or parties on a touch screen. Additional research is needed to isolate the mechanisms at work.

There is also evidence of the variability in the performance of lever machines similar to that observed by Brady et al. (2001). For the presidential race, the residual-vote rate for lever machines was the smallest observed. But for the Senate race, at 16.7 percent, it was more than double the already high level for punch cards (7.2 percent). Although we have no definitive evidence, it is most likely that this difference in residual votes is the result of voters' inability to find the race on the machine's interface (see Roth 1998; Caltech/MIT Voting Technology Project 2001). Alternative explanations exist, but they are not especially plausible.[17]

For Michigan, we present residual-vote rates by system type for president in 2000 and 2004 and Senate in 2000 (Table 1).[18] For both offices, overall rates in 2000 were consistently lower than those in Florida. Even punch cards had a residual-vote rate of only 2.0 percent for the president and 3.3 percent for the Senate. Nonetheless, the residual-vote rate for president in 2004 dropped from its 2000 level for four of the classifications and remained level for the fifth. The changes were generally small—unsurprising given the low starting points and the absence in that state of major voting problems. Still, the consistent declines lend support to Hypothesis 1.

Paper ballots had the highest residual-vote rate in two of the three comparisons, adding to the evidence that they are no better, and are perhaps worse, on this dimension than other voting methods (Brady et al. 2001; Ansolabehere and Stewart 2005). Furthermore, the residual-vote rate with paper ballots in 2000 was considerably greater than that for OS systems with precinct counting. The residual-vote rate for the Senate was also high among users of lever machines, though much less than the extraordinary rate observed in Florida.

Both the Florida and Michigan results are noteworthy for the extremely low absolute level of residual votes for president in 2004. In its report, the National Commission on Federal Election Reform (2001, 55) classified voting systems as "good" if the residual-vote rate for president was less than 1 percent. Florida's systems averaged half that rate, a remarkable turnaround from the fiasco in 2000.[19] Michigan, taken as a whole, was also in that range, despite the fact that one-quarter of the townships had not yet shifted to OS equipment. As noted, these numbers are partly a matter of substitution of newer, more voter-friendly equipment, but it seems likely that they are also a product of voters' becoming more equipment savvy.

The results for DREs in Florida are themselves noteworthy. Brady et al. (2001, 30, note 45) observed that two-thirds of the electronic systems in their analysis were "older, full-faced, push-button DREs" and that "touch-screens, in our opinion, are still unproven." Although definitive proof awaits information from more jurisdictions, the results from Florida in 2004 provide strong evidence of the ability of touch screens to minimize presidential residual votes.[20]

Residual-Vote Rates by System Transitions from 2000 to 2004

Although rates of residual voting by system type are a good start, residual-vote rates also need to be examined in conjunction with the change from one voting system to another across these two elections. Since one of the central purposes of electoral reform was to reduce residual votes, a direct test of the effectiveness of new technology is whether they decreased when new equipment was brought in. By observing the same jurisdictions over time, we control for factors such as differences in demographics that might contribute to observed system differences found in cross-sectional analysis. In addition, jurisdictions that used the same systems in the two elections provide a baseline for comparison of the reduction in

Table 2
Change in Residual Votes in the 2000 and 2004 Elections in Florida and Michigan, by Type of Voting Equipment Change (Percent)

Equipment Change	*N*	President				U.S. Senate			
		2000	2004	Change[a]	*t*-score	2000	2004	Change[a]	*t*-score
Florida									
OS to OS	118	1.5	0.3	–1.2	–10.93	3.0	1.7	–1.3	–13.00
OS[c] to DRE	139	4.7	0.5	–4.1	–22.56	4.0	3.0	–1.0	–5.98
PC to OS	187	9.8	0.3	–9.4	–17.87	6.5	1.8	–4.7	–22.46
PC to DRE	1,029	5.7	0.6	–5.1	–34.50	7.5	4.2	–3.3	–16.46
Michigan									
Paper to paper	19	1.5	1.6	0.0	0.09				
Lever to lever	116	1.8	1.0	–0.8	–1.79				
OS to OS[d]	670	0.9	0.7	–0.2	–2.36				
PC to PC	229	2.2	1.8	–0.3	–5.00				
Paper to OS[d]	57	2.4	0.9	–1.5	–3.15				
Lever to OS	94	1.5	0.6	–0.8	–2.20				
PC to OS	113	1.6	0.4	–1.1	–14.13				

Note: Selected precincts (Florida) or townships (Michigan) where the boundaries were unchanged between the 2000 and 2004 elections or could be merged to ensure that the same places were compared across time and absentee ballots were cast with the same technology as used at the polling place or could be excluded if they were not. PC = punch card; OS = optical scan; DRE = direct recording electronic system.

a. Change calculated before rounding.

b. One township that transitioned from OS to PC was dropped from the analysis.

c. These OS systems relied on counting at a central location rather than at the individual precincts.

d. Excludes central-count OS systems. Pooling the precinct-count and central-count townships together does not alter the substantive conclusions.

residual-vote rates. Since their equipment remains constant across the two elections, we can attribute additional amounts of residual-vote reduction in other locations to the adoption of new systems in the 2004 election.

Table 2 compares residual-vote rates for the same units across the two elections, matching precincts (Florida) or townships (Michigan) by their change, if any, in technology.[21] In both states, jurisdictions that did not change are shown first, followed by various types of changes. These comparisons leave no doubt that changes in technology resulted in lower residual-vote rates for president. For Senate voting, the story is a bit more complex.

In Florida, there was a substantial decline in the percentage of residual votes for president for all three types of changes (Table 2). It is not surprising that the movement from punch cards to OS and DRE systems resulted in the biggest reductions, though it is reassuring to have this confirmed when the comparison is made in the same set of jurisdictions in the two elections. More interesting is that the change from one "new" technology to another (OS to DRE) also led to a reduction in residual votes, probably because of the fact that the DRE systems made it impossible to overvote. All of these changes were substantially greater than the change in precincts that retained their OS technology, which provides a base against which voter awareness and information levels might be assessed. This supports the conclusion that switches in technology brought about change over and above what seems to have occurred as a result of greater voter knowledge and awareness. For the Senate in Florida, the change from punch cards clearly led to reductions in residual votes. But for changes from OS to DRE systems, the impact is less clear inasmuch as the change was small in absolute terms and slightly less than for precincts that retained their OS systems. Again, the results suggest that problems navigating the DRE interface lead to higher rates of residual votes down the ballot than might be expected if one considers the built-in protections only and not the ability of voters to use the interface properly.

In Michigan, there was also a decline in the residual-vote percentage for president for all three types of changes (Table 2). As was the case with Florida, the

residual-vote rates decreased across the board in all the equipment transitions except for a few townships that maintained their use of paper ballots, where the rate was essentially unchanged. As expected, the highest level of decrease in residual-vote rates occurred in places where new voting systems were implemented. The reduction was the smallest in townships or cities that did not change their equipment, especially in the places that stayed with precinct-count OS systems. Overall, in places that used the same voting equipment across the two presidential elections in Michigan, the residual-vote rates were reduced by 0.35 percentage points. On the other hand, a 0.97 percentage-point decrease occurred in townships and cities that switched from other systems to OS ($p = .001$). Thus, results from Michigan as well as Florida indicate that switches in technology brought about change beyond that produced by greater voter knowledge and awareness.

The results continue to contradict Hypothesis 2. The results for Florida precincts changing from punch cards are especially telling. Those that switched to OS systems began (in 2000) with extremely high residual-vote rates, four percentage points higher in the presidential race than those that switched to DREs. Yet in 2004, it was the OS precincts that had fewer residual votes. In the Senate election, those that switched to OS systems began one percentage point lower but ended up nearly two and a half points lower. Contrary to the hypotheses, OS systems seem to have lower residual-vote rates for president and Senate. Additional insight into the comparison of OS and DRE systems is provided when the residual vote is disaggregated.

Disaggregating the Residual Vote in Florida in 2000 and 2004

These aggregated vote measures conceal important differences in how the residual votes were generated. That is, residual votes are made up of undervotes and overvotes, which not only result from distinct individual behaviors but can also be protected against in different ways on different system types. This is most clear for DRE systems that simply do not allow voters to record multiple votes; furthermore, they guide voters sequentially through the ballot, suggesting that they consider every race. Compare this to punch-card systems that do not prevent voters from punching out multiple chads or to OS systems that do not prevent voters from filling in multiple bubbles or connecting multiple arrows. Although some OS systems provide a warning about overvotes, the warning is not always in an obvious location or easy to read, and in any event, ballots with overvotes can still be cast. With respect to the behavioral dimension, overvotes occur among voters who do not understand the voting rules that apply to a given race or the voting process as a whole or perhaps are unaware of the proper way to change a vote. Undervotes, on the other hand, occur among those who do not notice a particular race, consciously decide not to vote for some races, become fatigued, or do not vote properly (White 1960; Mather 1964; Bullock and Dunn 1996; Wattenberg, McAllister, and Salvanto 2000).

Typically, of course, separate classifications of overvotes and undervotes are not available at the precinct level and then not below the race for president. Currently, such distinctions are not recorded and preserved by local election officials in Michigan. In Florida, while we are not able to examine these separate measures for the full variety of system types, we can for several counties.[22] Beginning with the 2000 election, there was considerable variation in the recording of both kinds of votes (Table 3). This again illustrates the importance of county-specific factors and underscores the need to examine the same places over time. Overall, the highest overvote rate for president was 9.1 percent in Duval County, where, as noted earlier, the ballot had separate pages for the presidential candidates; it appears that many Duval County voters did not notice that the second page was a continuation of the presidential race and voted for presidential pairings on both pages. For the Senate, overvotes were much lower.[23]

The most striking result from Table 3 is the sharp decline in both overvotes and undervotes across the races for president and Senate from 2000 to 2004. By virtue of built-in technological protections, overvotes were eliminated in Miami–Dade and Palm Beach counties, the only counties in this group to move to DRE systems. But even in the counties using OS devices, overvotes never topped 0.2 percent. Presidential undervotes also dropped sharply, from an average of 1.8 percent in 2000 to 0.6 percent in the two DRE counties and to 0.3 percent for the three OS counties. Senate undervotes almost always went down as well (they increased slightly in Highlands County), though they remained relatively high in Miami–Dade (4.8 percent) and Palm Beach (3.6 percent).

Separating undervotes and overvotes allows us to observe that OS and DRE systems differ the most with respect to the rate of undervotes, and in particular, undervotes for Senate. This supports our earlier suggestion that problems navigating the interface

Table 3
Overvotes and Undervotes in the 2000 and 2004 Elections in Florida, by Machine Type (Percent)

	2000					2004				
		Undervotes		Overvotes			Undervotes		Overvotes	
County	Equipment	President	Senate	President	Senate	Equipment	President	Senate	President	Senate
Duval	PC	9.1	0.8	1.8	6.0	OS	0.0	0.0	0.3	1.6
Highlands	PC	1.8	0.0	1.6	2.2	OS	0.2	[a]	0.7	2.9
Marion	PC	1.0	0.7	2.0	2.6	OS	0.1	[a]	0.3	1.4
Miami–Dade	PC	3.0	1.5	1.6	7.6	DRE	0.0	0.0	0.7	4.8
Palm Beach	PC	4.9	1.8	2.0	7.4	DRE	0.0	0.0	0.5	3.6
	Average[b]	4.6	1.4	1.8	6.9	OS Average[b]	0.1	[a]	0.3	1.6
						DRE Average[b]	0.0	0.0	0.5	4.2

Note: The general analysis strategy was to use precinct-level data. For 2000, when precinct-level data were not available, individual ballot images were used to construct precinct totals. PC = punch card; OS = optical scan; DRE = direct recording electronic system.
a. Less than .05 percent.
b. Average calculated before rounding.

beyond the race for president might be more severe when using DREs than OS systems. Still, we hesitate to generalize too much with respect to this apparent difference because of the small number of counties included in this analysis. As the importance of the separate components of the residual vote becomes more apparent from these kinds of analyses, we hope that more jurisdictions will provide these data in disaggregated form for extended analyses.

Ecological Inference: Casting Invalid Votes across the 2000 and 2004 Elections

In this section, we use an ecological-inference technique as another means of testing the effectiveness of the transition from "older" to "new" voting technologies. The measure of interest here is the rate of residual voting for president in 2004 among those who cast a residual vote for president in 2000. This is a typical ecological-inference problem since we are interested in individual behavior across two elections, here with a focus on repeated mistakes, but only have information aggregated at the precinct level. To estimate this rate of repeated mistakes, we use a technique first suggested by Thomsen (1987) and later expanded on by Achen and Shively (1995). When applied to the problem at hand, the method essentially incorporates an unobserved dimension that links the aggregate residual-vote rates in two elections and provides individual-level estimates from the aggregate data.[24] Tests of this estimator have shown it to be successful (Thomsen, Berglund, and Worlund 1991; Achen 2000; Hanmer and Traugott 2004).

Table 4
Rates of Casting a Residual Vote for President in 2004 among 2000 Residual Voters

Equipment Change	Repeat Residual-Voting Rate (%)
Florida	
Manatee (OS to OS)	0.3
Charlotte (OS[a] to DRE)	0.5
Lake (OS[a] to DRE)	0.5
Duval (PC to OS)	0.5
Highlands (PC to OS)	1.9
Miami-Dade (PC to DRE)	1.1
Palm Beach (PC to DRE)	1.0
Michigan	
Lever to Lever	1.9
Lever to OS	0.8
Paper to OS	3.1
OS to OS	1.7
PC to PC	8.6
PC to OS	0.9

Note: Entries are ecological-inference estimates based on precinct data for Florida and township data for Michigan. For Michigan, because of the small number of observations, the central-count OS townships were dropped from the analysis. Pooling the precinct-count and central-count townships together does not alter the substantive conclusions. PC = punch card; OS = optical scan; DRE = direct recording electronic system.
a. The OS systems here relied on counting at a central location rather than at the individual precincts.

Table 4 provides estimated rates of 2000 residual voters who cast a residual vote again in 2004, by the change in voting system. The entries are ecological estimates retrieved after dividing the sample precincts

by the type of voting-system change and applying the Thomsen estimator. In Florida, between 0.3 percent and 1.9 percent of the voters who failed to have their vote for president recorded in 2000 also failed to cast a valid vote in 2004. Overall, residual-vote rates from 2000 to 2004 appear higher in Highlands, Miami–Dade, and Palm Beach counties, but none of the differences between counties is statistically significant. In short, our estimates show that the likelihood that residual voters in the 2000 presidential race again cast an invalid vote for president when they went to the polls in 2004 and voted on new voting systems is very low.

The results from Michigan, where townships and cities were grouped by their respective voting-system transitions, show that rates of repeat residual voting were higher there (Hypothesis 3). The rates are especially high in the three sets of places that did not change voting systems. The estimate for the townships that used punch-card systems in both elections stands out with an estimate that 8.6 percent of 2000 residual voters repeated their intentional or unintentional mistake in the 2004 presidential election. A surprising number of those counted out of the electoral process in these jurisdictions in 2000 were again counted out in 2004, almost one in twelve.

Consider a comparison of this category with places that used punch cards in the 2000 election but switched to OS systems in 2004. In these townships and cities, the estimated rate of repeated invalid votes is 0.9 percent, essentially indicating that there is no over-time correlation between the residual votes. Thus, compared with the places that did not replace their punch-card system, we observe the effectiveness of the new voting systems.

Conclusion

The introduction of new voting technology after the disastrous 2000 election was intended primarily to reduce the number of voters who unintentionally failed to vote for a given office or who voted but failed to have their vote counted because they did not do so properly. Comparing results for the 2004 and 2000 elections, the evidence from Florida and Michigan shows that the reforms succeeded. In both states, every change from old to new technology was accompanied by a decline in the percentage of residual votes. Where the residual vote was especially high, the decline was extraordinary—as high as a 96 percent reduction. But even where the rate was low to begin with, change brought about a lower level of uncounted votes.

Significantly, improvements occurred not only because of the introduction of new technology. Declines in the rate of residual votes occurred even in areas that did not change their voting equipment. The attention given to voting procedures evidently led to better ballot design, improved voter education, and more alert voters. Nevertheless, technological change mattered; the declines in jurisdictions that introduced new technology were significantly greater than in areas that did not change.

Change was successful not only relative to 2000 but in the small absolute level of residual votes in the 2004 presidential vote. In those areas using OS systems or DREs, the percentage of residual votes was just over 0.5 percent in Florida and just under 0.7 percent in Michigan. Overvoting, in particular, was nearly eliminated. This was true for the presidential race, in which overvoting was the predominant form of residual votes in 2000, and for the U.S. Senate race in Florida, in which overvotes were few to begin with.

HAVA was primarily a response to what happened in a presidential election, in which the percentage of residual voters was on the order of 2 percent. But the introduction of new technology has a much greater potential effect on lower level offices and ballot propositions, in which roll-off routinely reduces the percentage voting by 5 percent to 20 percent of those who go to the polls (Magleby 1984; Cronin 1989; Nichols and Strizek 1995; Wattenberg, McAllister, and Salvanto 2000). If new technology significantly reduces roll-off for these offices and propositions, the consequences could be more far-reaching than at the top of the ticket. This is a relationship that requires additional research.

Recent research on the impact of voting systems on voting behavior has significantly improved our understanding of how technology affects voter behavior. Nevertheless, the analytical possibilities have been limited by the small number of behavioral measures that are currently feasible. Future studies will need to move beyond gross categorizations of voting systems, ballots, and even the aggregated concept of residual votes. Just as there are important differences between models of automobiles, differences among individual voting-system designs are considerable as well.

The implementation of HAVA not only has introduced new technology, it has also ensured that there will be a continuous future stream of new technology in the American electoral system. Voters and election administrators will have to cope with changes in that technology at a frequent rate. Hardware will change,

and the software that guides it will as well. But there will be other pressures in the political system for change as well, such as the movement for verified voting-audit trails. For example, some jurisdictions that used their HAVA funds to purchase DREs before the 2004 election are currently considering replacing them with OS systems. This study has raised the possibility of new forms of analysis of the impact of voting technology on behavior, but new forms of data are required to push such analyses to the next level of detailed understanding and for the continual monitoring of the impact of this change.

Notes

1. The residual vote combines into a single measure overvotes (a voter selects more candidates for an office than is allowed), undervotes (a voter does not select a candidate for an office), and uncounted ballots.

2. In Florida, the jurisdictions we focus on used punch-card, lever, and OS systems in 2000 and OS or DRE systems in 2004; some jurisdictions stayed with OS for both elections. The decisions to adopt new technology were made at the county level. In Michigan, punch-card, lever, paper, and OS systems were used in both 2000 and 2004, with most jurisdictions using the same systems in both elections but many transitioning to OS systems, the eventual statewide system adopted in conjunction with the HAVA-mandated state plan.

3. It is common practice to limit analysis to units with unchanged boundaries (e.g., Ansolabehere and Stewart 2005; Mebane and Herron 2005) as this reduces the likelihood that any observed changes could be attributed to a significant change in either the population's demographic characteristics or voting behavior in those units. This is especially true when two elections that occur across a relatively short period of time are analyzed. As we discuss later, we did include jurisdictions with boundary changes but only when there was enough information to permit us to combine units such that we were sure we were comparing the same geographic areas across time.

4. For example, a CBS poll of registered voters taken before the 2004 election found that only 35 percent of the respondents reported had "a lot" of confidence that their votes would be properly counted (http://www.cbsnews.com/stories/2004/10/31/opinion/ polls/main652496.shtml?CMP=ILC).

5. The Democracy Index has already gained attention, with Senators Barack Obama and Hillary Clinton introducing separate bills to amend HAVA that were in part inspired by the idea.

6. While awareness and concern are separate mechanisms, we believe that in this context, they pull in the same direction. That is, greater awareness should lead to lower rates of residual voting as voters know better what sorts of problems might arise, and heightened concern should lead to more focus, and thus, lower rates of residual votes even if voters do not know exactly what potential problems to look for.

7. This analysis is based on exit-poll data available from the Roper Center (http://www.ropercenter.uconn.edu/elect_2004/).

8. We are not aware of any OS systems in 2000 or 2004 that were programmed to provide feedback about undervotes. However, some systems now can be programmed to provide this sort of feedback.

9. One way to assess the consequences of using this subset of the data is to compare in each election the Bush percentage of the vote in the subset of precincts to the percentage in the entire state. In 2000, Bush received 52.1 percent of the vote statewide and 50.3 percent in this subset of the precincts, while in 2004, Bush received 51.7 percent of the vote statewide and 51.8 percent in this subset of the precincts. An analysis of the distribution of the selected precincts by their Bush percentage compared to the Bush percentage in the state across both elections showed that they were typical of the state as well. Additional graphical information about these relationships is available from the corresponding author, Michael Traugott. The demographic composition of the sampled counties on race, education, and income were all similar to statewide indices as well. These all lead us to believe that we have a representative sample from Florida.

10. Absentee voters cannot use DRE or lever machines. Therefore, when absentee votes were recorded separately, we removed them from jurisdictions using those two types of machines. When absentee votes were not recorded separately, we dropped the jurisdictions. When only five DRE townships remained after these exclusions, we also dropped them given this extremely small *n*. In other words, nearly all of the cases that were dropped were dropped out of necessity as they would not allow us to examine separately the outcomes of interest by machine type. Moreover, we believe that the set of townships in our sample serves as an excellent representation of the state. For example, with respect to the comparison of Bush's support in the townships in our data set relative to the entire state, we find that support for Bush in our sample was extremely similar to support for Bush in the state. In 2000, Bush received 48 percent of the vote statewide and 46 percent in our substantial subset of townships, while in 2004, Bush received 46 percent of the vote statewide and 47 percent in our subset of townships.

11. Unfortunately, some counties made such extensive changes to their precinct boundaries that tracing the changes was impossible. These cases were dropped from the over-time comparisons.

12. While precinct-level data are available for many but not all cities and townships in Michigan, the boundary translation tables for the post-2000 census precincts that were available for Florida counties were not generally available for cities and townships in Michigan, so we stayed with the township- and city-level data because the number of units was large, the boundaries did not change, and in theory, we gain more complete state coverage.

13. Most counties in our analysis made changes in precinct boundaries from 2000 to 2004 because of redistricting, more often than not resulting in an increase in the number of precincts. In 2000, precincts in Volusia County are omitted, and in 2004, precincts in Martin County are omitted, in each case because of the lack of available precinct-level data. Note that the entries are mean scores of precincts, rather than aggregate percentages. A direct comparison shows that the discrepancy between these two different modes of calculation is negligible, while treating precincts as the unit of analysis enables us to generate test statistics across different equipment types.

14. The analyses presented here are based on unweighted precincts (Florida) and municipalities (Michigan). We made this

decision because the unit of analysis should be electoral settings and not individuals, but we also reran the analysis using turnout in the unit as a weight. When we did, there were no appreciable differences in the estimated size of the residual votes in places using different kinds of voting systems or that experienced different transitions from one kind of voting system to another.

15. In both 2000 and 2004, the Senate races in Florida involved open seats, and both were competitive and hotly contested, thus mitigating concerns that variations in the intensity of the races influence the results. The margins were relatively close; a Democratic former astronaut won in 2000 by 5 percentage points, while a Republican won in 2004 by fewer than 2 percentage points.

16. Another possibility is that there were malfunctions with the voting systems; however, there is nothing in the public record that reports on such problems.

17. One possibility is that the race for Senate was not particularly interesting to voters in the only county that used lever machines. Given the apparent high interest in the race elsewhere in the state, this seems unlikely. Another possibility is that there was a malfunction in that part of the machine; however, there is nothing in the public record that reports on such an event.

18. Michigan did not have a Senate race in 2004. Because of concerns with the comparability of a statewide race and U.S. House races held at the district level, we do not report residual-vote rates for the next item on the ballot in 2004.

19. With respect to the punch-card results, the distinctive features of the ballot that led to high rates of overvotes, such as the butterfly ballot in Palm Beach County and the two-page listing of presidential candidates in Duval County, provide a partial explanation for the larger residual-vote rate in Florida. In Duval County, five of the ten presidential and vice-presidential pairings were presented on the first page of the ballot booklet, with the remaining pairings and a write-in option appearing on the next page. Bush, Gore, Nader, and two minor-party candidates were on the first page, and five additional minor-party candidates were on the second page. When voters turned the page, some could have interpreted this as a separate race, especially since sample ballots instructed them to "vote on each page." Approximately 22,000 presidential ballots were not counted because of overvotes (see http://www.sfgate.com/cgi-bin/article.cgi?f=/c/a/2000/12/11/MN116795.DTL). Evidence consistent with this inference can be found in Table 4.

20. This does not mean that newer touch-screen systems are flawless. Note, in the Florida 2004 data, that OS systems showed considerably fewer residual votes for both president and U.S. senator. They also fail to overcome all the problems voters face (such as voting for the intended candidate, voting for multiple-candidate offices, use of straight-party features, write-ins; see Herrnson et al. 2008).

21. In Florida, when there was any change in the precinct boundaries, we aggregated multiple precincts to make the units constant over time. Because of comparability problems, four counties—Brevard, Marion, Martin, and Volusia—were not included in the analysis. For Michigan, there were relatively few transitions involving central-count OS systems. Thus, Table 2 reports the results for OS systems with precinct counting; pooling the OS precinct-count and OS central-count systems does not alter the results.

22. The results for the 2000 election for all but Duval County were generated from individual-level ballot images collected by the media consortium; the Duval results were calculated from precinct-level data provided by Lublin and Voss (2001). For 2004, precinct-level data were used.

23. The high rate of undervotes for Senate might suggest that presidential overvoters were likely to roll off the ballot and thus undervote for Senate. However, an analysis of precinct-level data using an ecological-estimation procedure demonstrates that this is not the case. Senate undervoters never composed the largest portion of presidential overvoters, and only twice, in Duval and Miami–Dade, was their percentage the second highest.

24. The Thomsen estimator is based on the premise that there is a latent partisan dimension that links behavior across time. That is, the underlying partisan dimension drives a voter's choices across two elections. By establishing the partisan micro-foundation, one can model behavior across elections to estimate the rate at which voters record residual votes in both elections, vote for the same party across the two elections, transition from one party to another, transition from a residual vote to a valid vote, or transition from a valid vote to a residual vote. We focus on the first measure. More technical details can be found in Thomsen (1987) and Park (2008).

References

Achen, Christopher. 2000. The Thomsen estimator for ecological inference. Unpublished manuscript, University of Michigan.

Achen, Christopher, and W. Phillips Shively. 1995. *Cross-level inference.* Chicago: University of Chicago Press.

Alvarez, R. Michael, Stephen Ansolabehere, and Charles Stewart III. 2005. Studying elections: Data quality and pitfalls in measuring the effects of voting technologies. *Policy Studies Journal* 33:15-24.

Ansolabehere, Stephen, and Charles Stewart III. 2005. Residual votes attributable to technology. *Journal of Politics* 67:366-89.

Brady, Henry E., Justin Buchler, Matt Jarvis, and John McNulty. 2001. Counting all the votes: The performance of voting technology in the United States. Mimeo. Institute of Government Studies, University of California, Berkeley.

Buchler, Justin, Matthew Jarvis, and John E. McNulty. 2004. Punch card technology and the racial gap in residual votes. *Perspectives on Politics* 2:517-24.

Bullock, Charles S., III, and R. E. Dunn. 1996. Election roll-off: A test of three explanations. *Urban Affairs Review* 32:71-86.

Bullock, Charles S., III, and M. V. Hood III. 2002. One person—No vote; one vote; two votes: Voting methods, ballot types, and undervote frequency in the 2000 presidential election. *Social Science Quarterly* 83:981-93.

Caltech/MIT Voting Technology Project. 2001. Residual votes attributable to technology: An assessment of the reliability of existing voting equipment. http://www.hss.caltech.edu/%7Evoting/CalTech_MIT_Report_ Version2.pdf.

Cronin, Thomas E. 1989. *Direct democracy: The politics of initiative, referendum, and recall.* Cambridge, MA: Harvard University Press.

Darcy, Robert, and Anne Schneider. 1989. Confusing ballots, roll-off, and the black vote. *Western Political Quarterly* 42: 347-64.

Frisina, Laurin, Michael C. Herrron, James Honaker, and Jeffrey B. Lewis. 2008. Ballot formats, touchscreens, and undervotes: A study of the 2006 midterm elections in Florida. *Election Law Journal* 7:25-47.

Gerken, Heather. Forthcoming. *The democracy index: Getting from here to there in election reform.* Princeton, NJ: Princeton University Press.

Hanmer, Michael J., and Michael W. Traugott. 2004. The impact of voting by mail on voter behavior. *American Politics Research* 32:375-405.

Herrnson, Paul, Richard G. Niemi, Michael H. Hanmer, Benjamin B. Bederson, Frederick C. Conrad, and Michael W. Traugott. 2008. *Voting technology: The not-so-simple act of casting a ballot.* Washington, DC: Brookings Institution.

Kimball, David C., and Martha Kropf. 2005. Ballot design and unrecorded votes on paper-based ballots. *Public Opinion Quarterly* 69:1-22.

Kimball, David C., Chris T. Owens, and Katherine M. Keeney. 2004. Unrecorded votes and political representation. In *Counting votes: Lessons from the 2000 presidential election in Florida*, ed. Robert P. Watson. Gainesville: University Press of Florida.

Knack, Stephen, and Martha Kropf. 2003. Voided ballots in the 1996 presidential election: A county level analysis. *Journal of Politics* 65:881-98.

Lublin, David, and D. Stephen Voss. 2001. Federal elections project. Washington, DC, and Lexington, KY: American University and the University of Kentucky.

Magleby, David B. 1984. *Direct legislation: Voting on ballot propositions in the United States.* Baltimore, MD: Johns Hopkins University Press.

Mather, George B. 1964. *Effects of the use of voting machines on total votes cast: Iowa—1920-1960.* Iowa City: Institute of Public Affairs, University of Iowa.

Mebane, Walter R., Jr. 2004. The wrong man is president! Overvotes in the 2000 presidential election in Florida. *Perspectives on Politics* 2:525-35.

Mebane, Walter R., Jr., and Michael C. Herron. 2005. Ohio 2004 election: Turnout, residual votes and votes in precincts and wards. In *Democracy at risk: The 2004 election in Ohio.* Washington, DC: Democratic National Committee.

Miller, Geralyn M. 2005a. Methodology, statistics, and voting error: An exploration of 2000 presidential election data in two states. *Policy Studies Journal* 33:1-13.

Miller, Geralyn M.. 2005b. The benefits of scholarly discourse: A response to "Studying elections: Data quality and pitfalls in measuring the effects of voting technologies." *Policy Studies Journal* 33:25-29.

National Commission on Federal Election Reform. 2001. *To assure pride and confidence in the electoral process.* Charlottesville: Miller Center of Public Affairs, University of Virginia.

Nichols, Stephen M. 1998a. State referendum voting, ballot roll-off and the new electoral technology. *State and Local Government Review* 30:106-17.

———. 1998b. Electronic voting machines and citizen participation in state judicial elections: Toward greater judicial accountability. *Southeastern Political Review* 26:699-711.

Nichols, Stephen M., and Gregory A. Strizek. 1995. Electronic voting machines and ballot roll-off. *American Politics Quarterly* 23:300-18.

Palazzolo, Daniel J., and James W. Ceaser. 2005. *Election reform: Politics and policy.* Lanham, MD: Lexington.

Park, Won-ho. 2008. Ecological inference and aggregate analysis of elections. PhD diss., University of Michigan.

Roth, Susan King. 1994. The unconsidered ballot. *Visible Language* 28:48-67.

———. 1998. Disenfranchised by design: Voting systems and the election process. *Information Design Journal* 9:1-8.

Saltman, Roy G. 2006. *The history and politics of voting technology.* New York: Palgrave Macmillan.

Singley, Mark K., and John R. Anderson. 1988. A keystroke analysis of learning and transfer in text editing. *Human-Computer Interaction* 3:223-74.

Stewart, Charles III. 2006. Residual vote in the 2004 election. *Election Law Journal* 5:158-69.

Thomsen, Søren. 1987. *Danish elections 1920–79: A logit approach to ecological analysis and inference.* Arhus: Politica.

Thomsen, Søren, S. Berglund, and I. Worlund. 1991. Assessing the validity of the logit method for ecological inference. *European Journal of Political Research* 19:441-77.

Tomz, Michael, and Robert P. Van Houweling. 2003. How does voting equipment affect the racial gap in voided ballots? *American Journal of Political Science* 47:46-60.

Wand, Jonathan N., Kenneth W. Shotts, Jasjeet S. Sekhon, Walter R. Mebane, Michael C. Herron, and Henry E. Brady. 2001. The butterfly did it: The aberrant vote for Buchanan in Palm Beach County, Florida. *American Political Science Review* 95:793-810.

Wattenberg, Martin P., Ian McAllister, and Anthony Salvanto. 2000. How voting is like taking an SAT test: An analysis of American voter roll-off. *American Politics Research* 28:234-50.

White, John P. 1960. *Voting machines and the 1958 defeat of constitutional revisions in Michigan.* Ann Arbor: Institute of Public Administration, University of Michigan.