Residual Votes Attributable to Technology

Stephen Ansolabehere
Massachusetts Institute of Technology
Charles Stewart III
Massachusetts Institute of Technology

We examine the relative performance of voting technologies by studying presidential, gubernatorial, and senatorial election returns across hundreds of counties in the United States from 1988 to 2000. Relying on a fixed-effects regression applied to an unbalanced panel of counties, we find that in presidential elections, traditional paper ballots produce the lowest rates of uncounted votes (i.e., "residual votes"), followed by optically scanned ballots, mechanical lever machines, direct register electronic machines (DREs), and punch cards. In gubernatorial and senatorial races, paper, optical scan ballots, and DREs are significantly better in minimizing the residual vote rate than mechanical lever machines and punch cards. If all jurisdictions in the United States that used punch cards in 2000 had used optically scanned ballots instead, we estimate that approximately 500,000 more votes would have been attributed to presidential candidates nationwide.

The election of the president of the United States in 2000 hinged on an aspect of the election system that had received scant attention from political scientists and political practitioners over the preceding century—the functioning of voting equipment. The most dramatic manifestation occurred in Palm Beach County, Florida, where two major problems cast doubt over the integrity of the election. Poor ballot design confused a significant number of voters about how to cast a vote. In addition, poor vote tabulator design made it difficult to determine intentions of voters. The "chads" from some punch cards had partially dislodged, making it impossible for the vote tabulator to count the ballots.

The methods used to cast and count ballots is surely one of the most mundane aspects of elections. But legal and political battles over the performance of voting technologies and the certification of the election results in Florida raised fundamental concerns about the fairness of the electoral process in the United States. Disputed elections can lower the perceived legitimacy of democratic elections, and some technologies might make it more likely to have disputed elections. In addition, voting technologies might violate equal protection of voters. People in counties with worse technologies may have a lower chance that their votes are counted. Finally, as a matter of election reform, improved voting technologies may increase the number of votes actually counted and, therefore, the effective

THE JOURNAL OF POLITICS, Vol. 67, No. 2, May 2005, Pp. 365–389

turnout. The magnitude of the problems exposed in the Florida recount surprised many political scientists, because they exceeded the estimated effects of many election administration reforms (Traugott 2003).

Concern over voting equipment in the wake of the 2000 election has given rise to a host of political and official studies into the effectiveness of the voting process.[1] Critical to all these assessments is a clear understanding of how, and to what degree, technologies used to cast and record ballots might interfere with all legally cast ballots being counted. Anecdotes from Florida and elsewhere illustrated that voting technologies might not function as designed, but these anecdotes are not generally informative about the extent to which technology interferes with the ability of people to register their preferences.

This paper provides an extensive, nationwide analysis of the degree to which the number of ballots counted depends on the voting technology used. It is the most expansive analysis that we know of, simultaneously covering all counties in the United States where the appropriate data can be amassed, the years 1988 to 2000, and elections for president, U.S. Senate, and governor. We examine a dependent variable that has become widely used in studying voting technology performance—"residual vote," which is the difference between the number of voters appearing on Election Day and the number of ballots actually counted in a given race. We exploit the panel structure of the data to hold constant a wide variety of county-level factors that affect voting patterns, such as demographics, political culture, and administrative practices. We measure the effect of changes in technology on changes in residual vote within counties over time and differences across counties.

The central finding of this investigation is that voting equipment has strong and substantial effects on residual votes. The difference between the best performing and worst performing technologies is as much as 2% of ballots cast. Surprisingly, paper ballots—the oldest technology—show the best performance. Paper ballots that are either hand counted or optically scanned have the lowest average incidence of residual votes in presidential elections and, down the ballot, in Senate and gubernatorial elections. These technologies perform consistently better than lever machines and punch cards. Electronic voting machines (aka DREs) have a statistically higher residual vote rate than hand-counted paper and optically scanned ballots.

A somewhat different question is what explains most of the variation in residual votes. Nearly 60% of the variation is accounted for by the county, rather than

[1] Several federal commissions issued substantial reports on the election process, most notably National Commission on Federal Election Reform (2001). At the state level see Florida Governor's Select Task Force on Election Procedures, Standards, and Technology (2001), Georgia Secretary of State (2001), Iowa Secretary of State (2001), Maryland Special Committee on Voting Systems and Election Procedures in Maryland (2001), Michigan Secretary of State (2001), Missouri Secretary of State (2001). The Help America Vote Act (HAVA) required all states to review their voting technology practices. Access to all of the resulting "state plans" is available through the Election Reform Information Project at the following URL: http://www.electionline.org/site/docs/pdf/hava_information_central.pdf.

by demographics or technology. Demographics and technology combined explain only about 15% of the variation in residual vote rates. Including indicators of county increases the percent explained to 70%. This finding suggests an institutional account of the incidence of uncounted votes. We suspect that the importance of county reflects the importance of local election administration.

Little scientific research exists into the performance of voting technologies. A handful of papers on this topic were published in the 1950s and 1960s, as manual lever machines moved from cities to the hinterlands (Mather 1964; White 1960). Academic interest in the topic was renewed in the 1980s with the adoption of punch cards and optical scan ballots. All of this research looks at a limited number of locales or only exploits cross-sectional variation. Mather (1964) established that turnout in Iowa counties that used lever machines was less than counties that used traditional paper ballots. White (1960) found that towns and counties in Michigan that used lever machines experienced greater "roll-off" or "voter fatigue" in referenda voting than did towns and counties that used paper ballots. Asher (1982) found that Ohio counties that used paper ballots had the least "fall-off," followed by punch cards ballots, and finally lever machines.[2] Studying the 1986 Oklahoma general election, Darcy and Schneider (1989) found a consistently positive correlation between the percentage of a precinct's population that was Black and roll-off, but their findings concerning the interaction between race and voting technology (i.e., optical scan vs. paper ballots) were inconclusive. Using an experimental design, Shocket, Heighberger, and Brown (1992) found that punch card ballots induced voters both to produce more overvotes (i.e., an excess of legal votes) *and* more undervotes (i.e., fewer votes than allowed under the rules), compared to other technologies. Nichols and Strizek (1995) reported roll-off was generally lower in the precincts of the city of Columbus that used electronic voting machines in 1992 on an experimental basis.

Following the 2000 election, there have been at least three cross-sectional studies of a national scope—Knack and Kropf (2003) study the 1996 election and discover performance differences across different types of machines (lever machines are best and punch cards are worst) and that "voided ballots" are highest in counties with high African-American populations. Brady et al. (2001) and Kimball, Owens, and Keeney (forthcoming) study the 2000 election, also finding similar performance differences across machine types.

This paper advances the methodology of past research by expanding the data across time and space. All of the previous research has been cross-sectional, typically for a small range of political jurisdictions. We examine a long time frame, from 1988 to 2000, and the entire nation.

Expanding the analysis across time and space allows us to exploit the panel structure of electoral data. Use of voting technologies varies considerably across

[2] Asher's "fall-off" rate is the total number of electors voting in a county minus the total number of ballots cast for a gubernatorial candidate, divided by total number of electors voting. This is identical to our "residual vote" measure used later in the paper.

counties, but also within counties over time. Specifically, we can estimate the effect of changing technology *within each county* on changes in the incidence of ballots with no vote counted. Only in Asher's (1982) study of Ohio do we find an explicit examination of the effects of switching technology within counties.[3] In the current paper we extend the logic of Asher's design into a multivariate setting, by using fixed-effects regression to examine a pooled time-series data set. Reliance on cross-sections risks confounding effects of technology with differences in other factors across counties and states. As we show below, most of the variation in the residual vote rate is attributable to county characteristics. Neither voting technologies nor demographics capture these factors, and there is considerable risk of omitted variable bias in small scale and cross-sectional analyses.

The remainder of this paper is organized as follows. The next section describes different voting technologies used in the United States and the factors that may cause each of them to fail. Then, we discuss our measure of uncounted votes. The following section describes our data and methods. We then report the results of a series of panel regressions that assess the relative performance of voting technologies. The final section concludes with a discussion of the policy implications of our findings and directions for further research.

Variability in Voting Technologies

Five types of technologies are used to cast and count votes in the United States today. In chronological order, they are hand-counted paper ballots, mechanical lever machine, punch cards, optically scanned paper ballots, and electronic voting machines (direct recording electronic machines, or DREs). Extended discussions of these machine types may be found at Fischer (2001). A detailed accounting of the features of particular machines used in the United States may be found at Center for American Politics and Citizenship (2003). Tomz and Van Houweling (2003) also contains a useful summary of the advantages and disadvantages of different voting technologies.

The oldest technology is the *paper ballot*, in which a voter makes a mark next to the name of the preferred candidates. Paper ballots are counted manually. *Mechanical lever machines*, introduced in the 1890s, are steel booths that voters step into. A card in the booth lists the names of the candidates; the names are next to levers, which are moved to record a vote. Ballots are counted by reading a series of counters inside the machine.

There are two variants of *punch card* ballots, which were introduced in the 1960s. In the first, the Votomatic, the ballot card must be inserted in a device, which contains the ballot. A vote is recorded when a voter punches out a pre-

[3] Mather (1964) gathered data across a long series of state elections, but the analysis proceeded one election at a time. Therefore, the effect of changing from paper to voting machines within a county in Iowa was left unexplored.

scored chad on the ballot card associated with a candidate. In the second variant, the Datavote, the ballot is actually printed on the card; the vote is recorded when a voter punches a hole next to a candidate's name, using a mechanical punch, much like a train ticket. Both types of punch card ballots are counted by running the cards through a card reader at the end of Election Day.

Optically scanned ballots, which began appearing in the 1970s, allow automated counting of paper ballots, using the same technology made familiar in standardized testing. Optically scanned ballots are counted on scanners, which may reside in a central office or in the precincts themselves. Finally, *direct recording electronic* (DRE) devices are electronic versions of the lever machines, and were introduced in the 1980s. There are two main variants of DREs. One type can be thought of as a lever machine with push buttons. A second variant presents voters with a touch screen computer monitor. The voter touches the name of the candidate on the screen and pages through the ballot electronically, like using an automatic teller machine at a bank. Votes are then counted in a variety of ways, but all rely on the electronically stored votes being transferred—electronically or on paper—from the machines to the central office.

Whenever a voter goes into a voting booth on Election Day intending to vote, she may leave a vote that will be counted for a particular office—or not.[4] Most trivially, the voter may intentionally abstain in a particular race. Beyond abstentions, the technology a voter uses to cast a ballot may fail to register a voter's intent, for a variety of reasons. Some of those reasons have nothing to do with the voting technology, per se, but rather, with the voter. Some voters may have greater difficulty completing a ballot than others. Literacy and language are common explanations for such problems (Posner 2001). Quite apart from what machine is being used, a county may have higher residual votes because it has more voters with low literacy. Physical impairments may affect the ability of voters casting their ballots as they intended.

The probability that a voter's ballot will be counted is also a function of how local elections are administered, which may also be independent of the voting technologies used. Analysis of the public finances of county election offices suggests that there are strong returns to scale, so county population likely affects the capacity of the election administration office (Caltech/MIT 2001). County administrators also have considerable discretion over how ballots are counted and over the certification of the vote.[5] Residual votes, then, will likely vary systematically

[4] It should be noted that there are other parts of the voting process that make it difficult to vote or even prevent some people from voting, including voter registration and polling place accessibility. Recent research suggests that the problems voters encounter before they get into the booth may be an even bigger barrier than voting equipment failures (Caltech/MIT Voting Technology Project 2001). These are subjects for further research, but not the focus of the current paper.

[5] In addition, states and localities differ in subtle and myriad ways in how votes are counted—ways that are not always apparent to the researcher. Some jurisdictions, for instance, may decide not to count write-in votes unless there are a "significant number." Other states may certify the total number of voters voting in a preliminary count, but then release a detailed accounting of all ballots cast later

from county to county. Some of this is predictable on the basis of county population.

County wealth will also affect administrative capacity. Local election offices typically have very limited resources, and resource constraints vary across jurisdictions in ways that likely affect the ability to record all the votes. Too few poll workers or inappropriate polling locations (e.g., poor lighting) may lead to higher errors. After the election, insufficient or poorly trained staffing in the election office may lead to errors in the recording of the vote, especially in checking for and resolving discrepancies. The varying resources of the counties alone should lead us to believe that the residual vote level across jurisdictions will vary.

Turnout is another potentially important factor that affects administration of elections in ways that lead to higher residual votes. If a county experiences an unusually high turnout rate, then there may be longer lines. This can interfere with voting several ways. Voters may feel rushed to complete the voting process, and in fact they may not be allowed to stay in the voting booth as long as they would like. Also, high turnout indicates many new voters, who may be unfamiliar with voting procedures. When turnout is high, not only will more voters need instruction, but poll workers will have less time to instruct voters on the way to use the voting equipment.

Of greatest interest to this investigation, however, is how votes might not be counted as a consequence of the technology used to cast ballots. Voting machines occasionally malfunction. Machine types vary in the frequency of mechanical (or other) failures, in how obvious the failures are, and in how easily failures can be remedied. One obvious advantage of traditional paper ballots is that they are fairly robust in the face of mechanical failures. The primary failure associated with paper ballots is simply running out of ballots. If an optical scanning machine breaks, optical scan forms can always be hand-counted (assuming the breakdown of the scanner is caught). On the other hand, machines of both the mechanical and electronic variety are notorious for hidden failures. For instance, if a counter that records the votes cast on a mechanical lever machine fails on Election Day, the malfunction may never be caught; if it is, there is no backup remedy to handle the failure. Likewise, if the internal logic unit of a DRE fails on Election Day, there might be no way to recover the affected ballots.

One of the failures of Votomatic punch cards in Palm Beach County was mechanical. In that case, controversies over "dimpled, pregnant, and hanging chads" were really about the failure of the punch cards to perform as designed.[6] When a voting machine fails mechanically, election officials will record a voter *who intends to vote* as having received a ballot, but when the ballot is counted, the vote will not register.

on, producing the appearance in some cases that more ballots had been counted than actually cast. See Brace and Chapin (1993) for the classic discussion of problems in election data reporting across states.

[6] For a broad discussion of the history of punch cards and the mechanical property of punch cards used in election devices see "Doug Jones's punched card index," http://www.cs.uiowa.edu/~jones/cards.

Machines can fail in another, subtler way that the Palm Beach County case also illustrates: machines can be poorly designed from the perspective of human usability. In Palm Beach County, the flaw was the infamous "butterfly ballot" that apparently confused voters in the presidential election (see Bullock and Dunn 1996; Darcy 1986; Darcy and Schneider 1989; Herron and Sekhon 2001; Wand et al. 2001; Wattenberg, McAllister, and Salvanto 2000). Looking beyond butterfly ballots, failures of ballot design more generally may make voting sufficiently confusing or inconvenient that some voters may become frustrated outright and leave without casting a ballot; others may be sufficiently misled that they may not complete the ballot and not even know about it. Lever machines, for example, present voters with an undifferentiated row of steel switches. It is hard to tell where one office ends and another begins. One common brand of lever machines (Automatic) places ballot questions and propositions above the eye level of many voters.

In general, different machine types present different challenges to voters and election administrators. This variation may very well affect how thoroughly voters complete the ballots they are faced with and how thoroughly the votes are tallied.

Distribution of Machine Types

Our focus is on the five main types of machines, without making distinctions within types.[7] In almost all states county election officials decide which machinery to use, so counties are, almost everywhere, the appropriate unit of analysis. Some counties do not have uniform voting technologies. In these situations, municipalities and, sometimes, individual precincts use different methods. These counties are called mixed systems. They appear most commonly in Massachusetts, Michigan, Maine, New Hampshire, and Vermont, where town governments usually administer elections.

The pattern of voting equipment usage across the United States looks like a crazy quilt. Americans vote with a tremendous array of equipment types. In the 2000 presidential election, one in five voters used the "old" technologies of paper

[7] In principle we would like to be able to study differences within different types of voting machines, especially the DREs, which have become so controversial. For instance, do "full face" DREs (like the Shouptronic) produce more residual votes than "single page" DREs (like the Accuvote-TS)? Unfortunately, there is not enough make and model information available for a sufficient number of counties to allow us to make precise estimates of differences across machine types. Failure to account for within-category heterogeneity will increase measurement error, when it exists. Research that attempts to document the heterogeneous features of existing election technologies is now beginning to appear. See Center for American Politics and Citizenship (2003).

The data also do not distinguish the equipment used to count absentee ballots when the jurisdiction's in-precinct method of voting cannot be used by mail. Tomz and Van Houweling (2003) discuss the problems presented by conflating absentee and in-person voting. In 1972, 96% of ballots were cast on Election Day in traditional precincts, compared with 79% of ballots in 2000 (Census Bureau, *Current Population Survey, Voter Supplement*, 1972 and 2000). We tested for correlation between the percent of absentee ballots cast and the county residual vote rate. It is statistically insignificant.

and levers—1.3% paper and 17.8% levers. Punch cards were used by just over one-third of voters (34.4%). Over one in four used optically scanned ballots. One in ten used electronic devices. The remaining 8.1% were in counties that used a mix of systems.

Within states there is typically little uniformity. This is illustrated in Table 1, which reports the percent of the population in each state that used the various types of voting technologies in 2000. Some states used only one method of voting, such as those with only mechanical lever machines (Connecticut and New York), DREs (Delaware), punch cards (D.C. and Illinois), and optical scanning equip-

TABLE 1

State Population using Types of Voting Technologies, 2000

	Voting Technology					
State	Punch	Lever	Paper	Scan	Electronic	Mixed
Alaska				76.6%		23.4%
Alabama		3.1%	.2%	81.2%	15.5%	
Arkansas	20.6%	12.7%	6.7%	56.8%	1.2%	2.0%
Arizona	20.1%			79.9%		
California	80.2%			15.3%	4.6%	
Colorado	45.7%		.5%	29.5%	24.3%	
Connecticut		100.0%				
D.C.	100.0%					
Delaware					100.0%	
Florida	64.5%	.1%	.1%	35.4%		
Georgia	43.8%	18.1%	.1%	38.0%		
Hawaii				100.0%		
Iowa		10.5%	.9%	82.1%	6.5%	
Idaho	58.7%		8.0%	33.3%		
Illinois	99.9%			.1%		
Indiana	36.2%	23.4%		6.9%	33.5%	
Kansas			6.3%	59.9%	33.8%	
Kentucky		9.3%		18.5%	72.2%	
Louisiana		50.9%			49.1%	
Massachusetts[a]	1.2%	15.7%	5.6%	77.6%		
Maryland	16.5%	17.6%		53.8%	12.2%	
Maine			27.2%			72.8%
Michigan	11.4%	2.0%		5.5%	.4%	80.7%
Minnesota	2.7%		5.7%	67.1%		24.6%
Missouri	69.7%		1.3%	29.0%		
Mississippi	21.5%	17.5%		59.0%	2.0%	
Montana	17.2%		6.4%	76.4%		
North Carolina	9.6%	3.4%	.2%	51.9%	34.8%	
North Dakota	7.0%		5.2%	87.8%		
Nebraska			12.1%	87.9%		
New Hampshire[a]			23.5%	76.5%		
New Jersey	2.6%	42.6%		17.5%	37.3%	
New Mexico				10.6%	89.4%	

TABLE 1 *continued*

	Voting Technology					
State	Punch	Lever	Paper	Scan	Electronic	Mixed
Nevada	82.0%			18.0%		
New York		100.0%				
Ohio	74.4%	2.3%		12.6%	10.8%	
Oklahoma				100.0%		
Oregon	47.1%		.4%	52.5%		
Pennsylvania	12.7%	62.6%	.3%	11.5%	12.9%	
Rhode Island				100.0%		
South Carolina	40.4%			15.5%	44.1%	
South Dakota	10.1%		11.3%	78.6%		
Tennessee	12.5%	23.1%		11.0%	53.4%	
Texas	30.0%	1.1%	3.3%	62.9%	2.7%	
Utah	97.6%		1.9%	.5%		
Virginia	20.4%	43.0%	.1%	16.9%	19.6%	
Vermont[a]			38.5%	61.5%		
Washington	63.7%			36.3%		
Wisconsin	2.1%		6.5%	18.2%		73.2%
West Virginia	36.5%	6.4%	11.2%	45.8%		
Wyoming	13.8%	2.8%		79.5%	2.4%	1.4%

Source: Election Data Services; state and local election officials.

[a] Measured at the town level for Massachusetts, New Hampshire, and Vermont. All other states measured at the county level.

ment (Hawaii, Oklahoma, and Rhode Island). At the other extreme, states such as Arkansas, Georgia, Indiana, Kansas, Maryland, Mississippi, North Carolina, South Carolina, and Tennessee did not have one dominant voting technology in 2000.[8]

Just as the heterogeneity of voting equipment used in the United States is impressive, changes in technology over time have also been impressive and dramatic. From 1980 to 2000, the fraction of voters using optically scanned ballots and DREs grew from 3.2% of the population covered to 38.2%. There was little change in the use of punch cards. Paper ballots fell from 9.7% of all people in 1980 to just 1.3% in 2000. Lever machines, by far the dominant mode of voting in 1980, covered 43.9% of the electorate. In 2000, only 17.8% of people resided in counties using lever machines.[9]

[8] Following the 2000 election, Georgia Secretary of State Cathy Cox declared that her state would adopt a single DRE for the 2002 midterm election, which it subsequently did. Overall, Election Data Services reports that 200 counties changed their voting technologies between 2000 and 2002, as jurisdictions with lever machines and punch cards migrated primarily to DREs. See Election Data Services (2002).

[9] There have been several studies of why counties choose particular voting technologies. See Garner and Spolaore (2001).

The trend toward electronic tabulation over the last two decades, along with the adoption of punch cards in the 1950s and 1960s, reflects the demand for the faster tabulation of ballots. Punch cards, optical scanners, and DREs use computer technology to produce a speedy and, hopefully, more reliable count.

Our analysis exploits the variation in technology usage both across counties and within counties over time. Between 1988 and 2000, nearly half of all counties (1,476 out of 3,155) adopted new technologies. And today four of the five technologies (lever machines, punch cards, optical scanning, and electronic machines) are widely used across counties.

Measuring Uncounted Ballots

The empirical analysis that follows focuses on which types of technologies produce the most complete count of votes cast. Our measure of uncounted votes is the number of blank, spoiled, or unmarked ballots, which we term the "residual vote."[10]

To clarify the statistical analysis below, we consider here residual vote as a measure of uncounted votes and the possible causes of residual vote, some of which stem from technology and some of which do not.

To calculate residual vote, we assembled data on the total number of votes cast in each county and the total number of ballots counted with a valid vote for president, U.S. Senate, and governor. Data from 1988 to 1998 were acquired from Election Data Services (EDS); data from 2000 were acquired from state and local election officials.[11]

The residual presidential vote in the average county equaled 2.3% from 1988 to 2000. Because county populations vary dramatically, this does not equal the

[10] We prefer the term "residual vote" to several other names given to this quantity for several reasons. First, this is the term used in federal legislation and court cases; see National Commission on Federal Election Reform (2001), H.R. 3295 (Ney-Hoyer Bill), S.565 (Dodd bill) of the 107th Congress and the opinion of the *Southwest Voter Education Project vs. Shelly*, Case no. 3-56498, 9th U.S. Circuit. Second, other terms that have appeared in academic and popular writing, such as "error rate," "voter fatigue," "the uncounted vote," and "spoiled ballots," suggest that the residual is pure error on the part of the machine or the voter, which it may not be. Also, residual vote is not "drop off" or "roll off" or "fatigue" because the voter may have in fact made all of the selections but the machine may have failed.

[11] Brady et al. (2001) has raised some questions about the quality of the EDS data. Although we did not thoroughly replicate the EDS data gathering operation, we independently collected from election officials as much data as possible from past election years and compared it with the EDS data. We corrected the few errors we discovered, which were minor. The major data problem we encountered did not concern the EDS data, per se, but the heterogeneity of election reporting practices across the different states, which we mention above. These practices vary more greatly than data errors. We are therefore confident that the EDS data we could not independently verify is free of errors that would systematically bias our substantive findings; our fixed-effects framework further guards us against systematic problems that may be associated with the practices of particular states and counties.

fraction of people who cast an under- or overvote for president in these years. This figure is somewhat smaller: 2.2%. Over the past decade approximately 100 million votes have been cast in each presidential election, so approximately 2.2 million ballots recorded no vote for president in each of the past four presidential elections.

There is considerable variation around this average. The standard deviation of the residual presidential vote is 2.4% weighting all counties equally and 2.0% weighting them by population. The data are also positively skewed: the first quartile of counties is 1.0%, the median is 1.8%, and the third quartile is 2.9%. The skewness statistic is 5.8.

The residual gubernatorial and senatorial vote rates are somewhat higher. The county average residual vote rate in gubernatorial and senatorial elections is 4.2%, and the fraction of all ballots cast (population weighted county average) is 4.1%. The standard deviations are 3.5% for the county average and 2.9% for the population-weighted data. The skew statistic is 2.8.

For the purpose of measuring the effects of technology, residual votes are an appropriate indicator. First, intentional abstention is a small fraction of the residual vote rate. Exit polls and post election surveys indicate that from 1988 to 2000 approximately one-half of 1% of voters intentionally abstain from voting for president.[12] The residual vote rate is 2.2% of total ballots cast. That leaves approximately 1.7 million votes (1.7% of total ballots cast) "lost" because of technological problems, administrative deficiencies, and voter confusion.

Second, the residual vote is the dependent variable; random noise in that measure due to variation in abstention rates will not produce bias. It lowers efficiency, making it less likely to find statistically significant differences across technologies. The overall level of abstentions is captured in the intercept term of the regressions below. If abstention varies across counties, across years, or across particular races, the panel structure of our data allows for abstention levels to be accounted for in the dummy variables that account for fixed county, year, and race effects.

The proof of the usefulness of the residual vote measure is in the pudding. If this measure is largely intentional abstention that is not itself due to technology, then we expect there to be no effects of technology on residual votes. In fact, there are substantial effects, which we show in the next section.[13]

[12] Responses to the American National Election Study help to provide a rough estimate of the frequency of conscious abstention. Among respondents who reported having voted, .3% reported not voting for president in 1988, .7% in 1992, 1.0% in 1996, and .3% in 2000. Therefore, the rate of actual abstention in presidential elections is roughly .5%.

[13] Many reviewers and commentators on this research have suggested that we conduct the following analysis on the two components of residual vote, overvotes, and undervotes. The reason that cannot be done is practical. As far as we know, Florida is the only state that requires all its counties to report over- and undervotes, a requirement that was instituted in light of the 2000 debacle.

Data and Methods

The lack of uniformity of voting technologies was cause for concern among many reformers in the aftermath of the 2000 election. However, to social scientists this heterogeneity is an opportunity. The wide range of different voting machinery employed in the United States, temporally and geographically, allows us to gauge the reliability of existing voting technologies.

For the remainder of this paper, we examine the relative reliability of different methods of casting and counting votes two ways. First, we contrast the incidence of residual votes across counties using different sorts of technologies. Second, we examine how changes in technologies within localities over time explain changes in residual vote rates. If existing technology does not affect the ability or willingness of voters to register preferences, then the incidence of over- and undervotes will be unrelated to what sort of machine is used in a county.

We have acquired or collected data on election returns and machine types from approximately two-thirds of the 3,155 counties in the United States over four presidential elections, 1988, 1992, 1996, and 2000. We have also acquired or collected election returns from gubernatorial and senatorial elections from 1988 through 1998. Eleven states do not ask or require counties to report the total number of voters who go to the polls, and therefore such states must be excluded. The data cover approximately 2800 counties and municipalities, though not for all years. Viewed as a percentage of all votes cast for president in each year in our analysis, we cover 56% of all votes in 1988, 65% in 1992, 68% in 1996, and 78% in 2000.

In almost all states, voting equipment is uniform within each county. Six states administer elections at the town level. For two of these (Massachusetts and Vermont) we were able to collect the requisite data for this analysis, and we have included their town-level data.[14] Otherwise, we exclude the mixed-technologies counties from the analysis. There are over 20,000 total observations in the data set—9,000 in presidential years and 11,000 in midterm years. The large number of observations produces high levels of precision in estimating average residual vote rates associated with each machine type. In the appendix we report which states fall within our sample during the elections for this time period.[15]

[14] Although we have used town-level data for New England, for simplicity's sake we will refer to counties in the paper.

[15] Counties fall out of our sample in almost all cases because the state did not require counties to report the total number of voters who took a ballot on Election Day. The fact that we cannot include all U.S. counties in this analysis raises concerns about the effects of this implicit sampling scheme on our results. The reduction of observations of course reduces efficiency. However, we have so many observations remaining that efficiency loss is the least of our concerns. The biggest concern arises if the correlation between voting technologies and residual vote is different in the counties we have data for, compared to counties that have missing data. Absent the turnout data from the missing jurisdictions, we of course cannot directly address this concern. Theoretically we know of no reason why the *relative* performance of voting machines should vary as a function of whether a state requires its counties to report total turnout. Alternatively, as one reviewer of this paper noted, switching from one

Beyond equipment, many other factors may explain rates of uncounted votes and abstentions. As discussed earlier, turnout, county wealth, and various population demographics likely affect the residual vote rate. In addition, election laws and electoral competition probably affect residual votes. Other prominent offices on the ballot, such as senator or governor, might attract people to the polls who have no intention to vote for president.

To hold constant the many factors that operate at the county level, we exploit the natural experiment that occurs when locales change machinery. We measure how much change in the residual vote occurs when a county changes from one technology to another. The average of such changes for each technology type provides a fairly accurate estimate of the effect of the technology on residual votes, because the many other factors operating at the county level (such as demographic characteristics) change relatively slowly over the brief time span of this study.

Operationally, we do this comparison by doing fixed-effects regressions on an unbalanced panel, in which the observation is a county year. Fixed-effects regressions have become very common within economics and are gaining wider currency within political science (see Greene (2000, 567–84) for a comprehensive overview of it and related techniques). A dummy variable for each county is included to measure the fixed effect associated with unmeasured local factors. To guard against other confounding factors, we also control for contemporaneous senatorial and gubernatorial races on the ballot, the state, and year of the election through another set of dummy variables. Finally, we also include the log of turnout as an independent variable.

Because all previous research into this topic has used cross-sectional analysis, we also contrast our fixed-effects analysis with regressions that instead use measures of county-level demographic factors—most notably age, race, and income—as controls.

Results

Basic descriptive statistics about residual votes for various technologies capture many of the principal results of this investigation. Table 2 presents the average residual vote rate for each type of voting equipment in presidential, gubernatorial, and senatorial elections from 1988 to 2000. The first three columns report average residual vote rates by counties. The last three columns report the residual vote rates, weighting each county by its turnout.

voting technology to another is generally not random. (However, developments in states like Georgia, where the state imposes a uniform system on all localities, may closely approximate randomness.) There may be private information available to counties about the "best" equipment to use that we cannot observe, and therefore bias may creep into our results. Our experience talking with election officials in our larger project leads us to conclude that improving the accuracy of vote counts is a second-order concern, at best, when counties change voting machines. Voting equipment changes almost always occur because of a desire to speed up the count. Therefore, we do not believe there are serious endogeneity problems to be concerned with here.

TABLE 2

Residual Vote in Presidential Elections, by Machine Type, U.S Counties, 1988–2000. (Numbers in parentheses are standard deviations.)

	Counties			Voters		
Machine Type	Pres.	Gov.	Sen.	Pres.	Gov.	Sen.
Paper ballot	1.8%	3.3%	3.6%	1.9%	3.2%	3.8%
	(2.1%)	(2.0%)	(3.2%)	(2.0%)	(2.2%)	(2.7%)
Lever machine	1.9%	5.1%	9.5%	1.8%	4.2%	7.0%
	(1.8%)	(3.1%)	(5.3%)	(1.8%)	(2.6%)	(3.5%)
Punch card	2.9%	3.3%	4.7%	2.5%	3.3%	4.4%
	(1.1%)	(2.1%)	(3.1%)	(1.5%)	(1.6%)	(2.7%)
Optically scanned	2.1%	3.1%	3.4%	1.6%	2.1%	3.0%
	(2.7%)	(1.9%)	(3.8)%	(2.4%)	(1.7%)	(3.1%)
Electronic (DRE)	3.0%	4.3%	8.2%	2.5%	3.7%	5.4%
	(3.0%)	(1.2%)	(4.0%)	(3.6%)	(1.9%)	(3.4%)
Mixed	2.0%	5.0%	6.1%	1.5%	3.0%	3.6%
	(1.7%)	(2.8%)	(3.9%)	(1.3%)	(1.5%)	(2.1%)
Overall	2.2%	3.6%	5.5%	2.1%	3.2%	4.7%
	(2.3%)	(2.3%)	(4.5%)	(1.9%)	(1.9%)	(3.2%)

Examining this table reveals a fairly consistent pattern of machine performance. Optically scanned ballots show the lowest average residual vote rate across almost all of the offices examined. In the presidential elections under study, voters in counties using optically scanned paper ballots averaged a residual vote rate of 1.6%. In gubernatorial and senatorial elections, those voters average a residual vote rate of 2.1% and 3.0%, respectively.

Hand-counted paper does remarkably well. Voters in counties using paper ballots have an average residual vote of 1.9% in the four presidential elections studied, and they have average residual vote rates of 3.2% and 3.8% in the gubernatorial and senatorial elections. Punch cards show the worst performance among the paper-based systems.

Voters in counties using lever machines have a very low residual vote rate in presidential elections (1.8%), but those same voters have the highest residual vote rates in senatorial elections (7.0%) and the second highest in gubernatorial elections (4.2%). Finally, electronic voting machines produce a moderate level of residual votes in presidential elections (2.5%), a much higher rate in senatorial elections (3.7%), and a high rate in gubernatorial elections (5.4%).

These summary statistics show differences in performance between presidential and "down-ballot" races, depending on the voting technology used. Which machine is "best" appears to depend, in part, on which race you are studying. Examining *why* these differences exist is beyond the scope of this paper. There are obvious speculations that can help guide future research. Above we noted, for instance, that mechanical lever machines do very well in identifying the presidential race, since it appears in the upper left-hand portion of the display. However, finding the next race (governor or senator, for instance) may be problematic, depending on things like the number of candidates on the ballot and the type face of the ballot cards (which is often six-point type). Furthermore, election workers program the inside of the machine, where various "lock-outs" have to be inserted, and count votes from *behind* the machine, without a direct reference to the ballot on the front. Thus, programming and counting errors become more likely as workers move through thousands of moving parts at the back of the machine.

To perform this analysis more generally, we estimated the fixed-effects regression we previously described. The first two columns of Table 3 report the results of these regressions; the first column is the effect of changing voting equipment on the residual vote rate in presidential elections while the second column reports the effect of changing equipment in senatorial and gubernatorial races.[16] We will return to the results in columns three and four, which do not use fixed effects, below. In all regressions, counties are weighted by overall turnout, so the interpretation of the dependent variable is the percent of ballots cast nationwide.

Because the five technology categories are linearly dependent with the constant in the regression, we exclude one of the categories. We chose the oldest "modern" technology, lever machines, as the excluded technology category. Therefore, the equipment coefficients measure how much higher or lower is the average residual vote of that equipment type, compared to lever machines.

Paper ballots are the champions in presidential ballots in this fixed-effects analysis, producing 1.4% fewer residual vote than mechanical lever machines. Next in efficacy are optical scan technologies, which produced .45% fewer residual votes. Although the coefficient for DREs is positive, the standard error is sufficiently large that we do not conclude that they produce worse residual vote rates than lever machines; we simply conclude that the performance of DREs cannot be distinguished from that of the older lever machines. Bringing up the rear, by a significant amount, are punch cards, which produced .82% more residual votes than lever machines.

Are these differences substantively "large" or "small"? One way of answering this question is to consider a thought experiment that corresponds with a common policy choice facing election officials throughout the United States: What would happen if all counties that used punch cards in 2000 had used the best computer-based system in this analysis, optical scanners? Estimating the answer to this question is fairly straightforward. The difference in coefficients suggests that a jurisdiction moving from punch cards to optical scanners should expect its residual vote rate to decline by (.0082 + .0045 =) 1.27% points. In 2000 roughly 34 million voters cast votes on punch cards. Had they cast their ballots on optically scanned ballots, approximately 431,800 more votes

[16] We used the STATA command *areg* to perform these regressions. We combined the gubernatorial and senatorial analysis in columns 2 and 4 for the sake of simplicity.

TABLE 3

Residual Vote Multivariate Analysis, Presidential, Gubernatorial, and Senatorial Elections, 1988–2000

	Fixed-effects estimates		No fixed effects, explicit controls	
	President	Gov. & senator	President	Gov. & senator
Equipment effects:				
Punch card	.0082 (.0015)	–.0030 (.0018)	.0077 (.0005)	–.021 (.001)
Lever machine	Excluded equip. category	Excluded equip. category	Excluded equip. category	Excluded equip. category
Paper	–.014 (.002)	–.014 (.003)	–.0012 (.0014)	–.022 (.002)
Optical scan	–.0045 (.0014)	–.014 (.002)	.00071 (.00070)	–.032 (.001)
Electronic (DRE)	.0022 (.0015)	–.012 (.002)	.0080 (.0010)	–.0097 (.0013)
Shift in tech.	.0010 (.0007)	–.0004 (.0010)	.00005 (.00067)	–.0021 (.0013)
Log (turnout)	.0095 (.0026)	.031 (.003)	–.0004 (.0001)	.0005 (.0002)
Gov. or Sen. on ballot	–.0011 (.0007)	–.004 (.001)	–.003 (.001)	.005 (.001)
Senator	—	.008 (.001)	—	.009 (.001)
Percent Over 65	—	—	.047 (.008)	.104 (.009)
Percent 18–24	—	—	–.012 (.009)	.027 (.010)
Percent White	—	—	–.030 (.002)	–.045 (.003)
Percent Hispanic	—	—	.011 (.004)	.005 (.005)
Median Income (10,000s)	—	—	–.002 (.001)	–.001 (.001)
Constant	–.11 (.03)	–.29 (.03)	.025 (.002)	.027 (.016)
N	8,982	11,625	8,982	11,625
R^2	.79	.74	.14	.43
Fixed effect: (not shown)	Year × State County	Year × State County	Year × State	Year × State
Number of categories	3,346	2,245	—	—
F test	F(3345,5572) = 2.971 ($p < .0001$)	F(2244,9318) = 3.705 ($p < .0001$)	—	—

would have been included in the presidential tally. A similar calculation suggests that had all voters who used lever machines cast ballots using optically scanned ballots, approximately another 80,000 ballots would have been included in the tally. Taken together, this represents roughly one-half of 1% of presidential turnout. Since 1972, the standard deviation in national presidential turnout has been 2.2%. Therefore, if these "lost votes" were "recovered" through the adoption of better voting technologies, the increase in presidential turnout would be notable.

Turning to gubernatorial and senatorial elections, the regression results are quite different. In presidential voting, lever machines are in the middle of the pack in terms of reliability. In gubernatorial and senatorial voting, they are at the back. As well, the performance of DREs is much better in these races than they were for president. With a residual vote rate 1.2% lower than lever machines, DREs perform comparably to paper (1.4% lower) and optical scanning (1.4% lower).

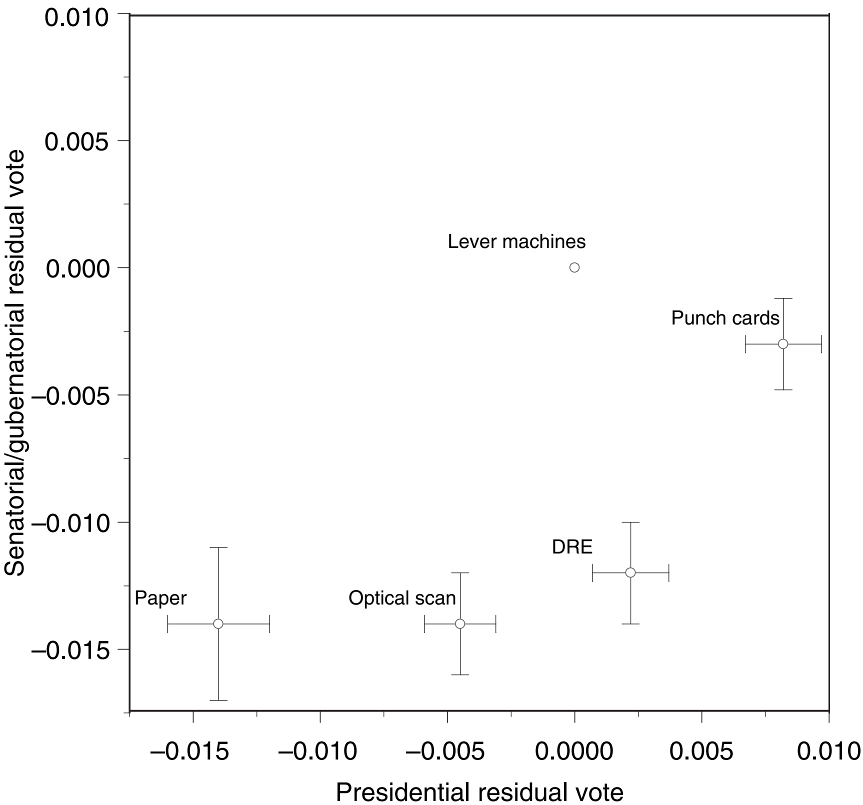
This difference in performance across the two types of elections is illustrated in Figure 1, which presents in graphical form the fixed-effects voting equipment coefficients from Table 3. (The crosshairs indicate the standard errors associated with the coefficients on each dimension. The coefficient is at the intersection of the crosshairs.) While there is a continuum of performance along the residual vote rate for president, there is a clear distinction between two groups of equipment in terms of senatorial/gubernatorial residual vote. For political jurisdictions considering making a switch in voting equipment, this graph illustrates one clear choice: a movement from either lever machines or punch cards to paper, optical scanning, or DREs should increase reliability along at least one dimension, if not both. Changing within the three dominant technologies (e.g., from optical scanning to electronics or from paper to optical scanning) provides ambiguous gains, if any at all.

Returning to Table 3, we also see some interesting, more subtle results that also pertain to voting technologies. First, it would seem intuitively obvious that when a jurisdiction switches its voting technology, voters unfamiliar with the new technology would be more likely to make mistakes, and therefore residual vote should go up. However, the sign of the "technology shift" dummy variable is effectively zero in both analyses.[17] Before dismissing the importance of a shift in technology, we should note one major problem with this variable that may attenuate the estimated effect: the technology shift variable is probably endogenous. Election administrators who are rolling out a new voting technology are usually worried about local voters using it incorrectly. Therefore, it is quite possible that these officials step up voter education efforts whenever new technologies are implemented. If so, then this coefficient only measures the *net effect* of errors due to the introduction of new technologies.

[17] The "shift in technology" dummy variable is equal to 1 if the county adopted a new voting technology in year *t*, 0 otherwise.

FIGURE 1

Comparison of Voting Technology Performance Coefficients for President and Senator/Governor

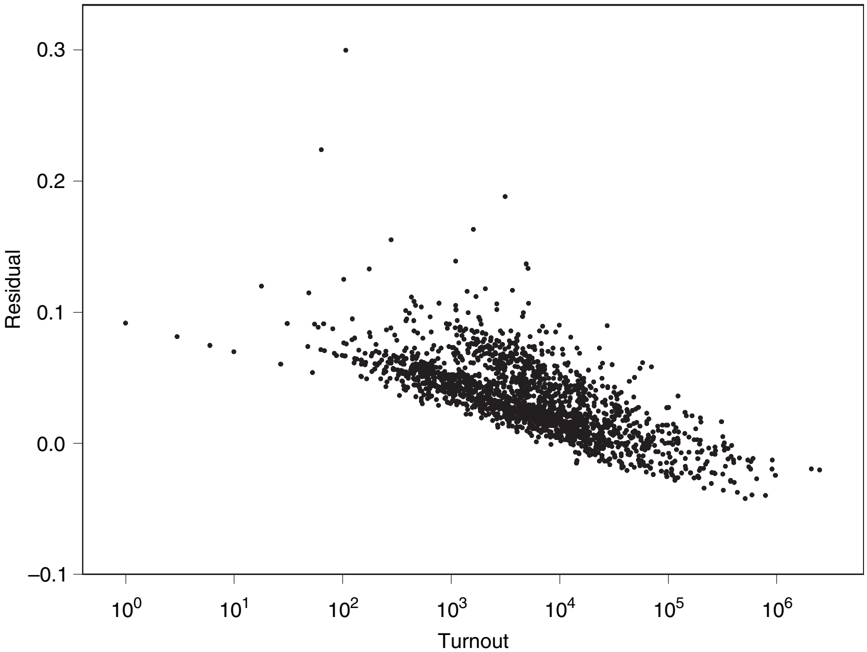


NB: The *x*-axis records the indicated technology-related coefficient from column 1 of Table 3. The *y*-axis records the indicated technology-related coefficient from column 2. The cross-hairs indicate ±1 standard error of the indicated coefficient. Lever machines, the omitted technology category, are indicated with the point at (0,0).

Finally, the behavior of the *turnout variable* reveals an important subtle effect that turnout and size of the electorate have on voter errors. The turnout coefficient is positive in columns 1 and 2. Because the county fixed effects control for the overall size of each county's electorate (or average turnout), the turnout variable measures the effects of short-term turnout fluctuations. Within a county, a big turnout surge results in more "lost votes." This is consistent with the commonly expressed view that exceptionally high-turnout races induce more inexperienced voters to come to the polls, a higher proportion of whom may fail to cast a countable ballot for a variety of reasons.

FIGURE 2

County-specific Effects Against Turnout in County/town. (Source: Residuals from Regression Reported in Column 1, Table 3)



The importance of using county-specific dummy variables in a fixed-effects framework to control for a laundry list of local factors that influence residual vote rates can be illustrated by examining the values of these county-specific dummy variables. These dummy variables can be thought of as the "baseline" residual vote rate for each county. In Figure 2 we have graphed the values of these 1,954 separate dummy variables against logged turnout for each presidential election. Note the strong negative correlation between the value of the fixed-effect coefficients and turnout. This is further evidence that cross-sectional factors that are correlated with size of the jurisdiction have a strong influence on the level of residual vote in a jurisdiction.

Just how important these myriad local factors are can be seen by comparing the R^2 statistics of the OLS estimates (columns 3 and 4) with the fixed-effects estimates (columns 1 and 2) in Table 3. From a variance-explained perspective, it appears that most of what influences whether votes get counted is due to population-dependent factors that are distinct from the type of voting technology used.

The performance of the county-specific coefficients in this analysis provides a cautionary note concerning other research that is currently emerging on the performance of voting technology. Knack and Kropf (2003), Brady et al. (2001), and others have recently performed cross-sectional analysis similar to what is presented in this paper. A comparison of the coefficients in columns 1 and 2 with the coefficients in columns 3 and 4 suggests that unless researchers are lucky enough to control for the relevant nontechnological, jurisdiction-specific factors affecting residual vote rates, the risk of encountering omitted variables bias is high. In the case of the technology effects, the paper ballot and optical scan coefficients are a full order of magnitude larger when we use the fixed-effects approach than when we enter a list of controls directly into the regressions. Not only does omitted variable bias affect the estimated size of the technology effect in the cross-sectional analysis, but it can affect other variables, too, as the sign change on the turnout variable attests to.[18]

Finally, we checked the robustness of our results in a variety of ways. We tried various transformations of the dependent variable, and we split the data into counties of different sizes (under 5000 votes, 5000 to 100,000 votes, and over 100,000 votes). We added a dynamic, element, by including a lagged dependent variable. The pattern of results in the fixed-effects analysis was always the same.

Conclusions

The primary empirical finding of this paper is that voting technologies are not neutral with respect to recording votes cast by voters on Election Day. The overall residual vote rate is greater than the proportion of voters who report abstaining by a factor of five. In presidential races, punch cards perform the worst and optical scanners perform the best. In gubernatorial and senatorial races, mechanical lever machines are worst, followed by punch cards; three technologies—paper, optical scanning, and DREs—tie for best. Voting technologies also vary in how well they capture votes as one goes down the ballot.

One innovation in this paper is exploiting the panel structure of election data, which gives us some leverage over the omitted variables bias problem that attends cross-sectional studies of this topic. A comparison of our findings with the most recently published cross-sectional study of residual vote rates illustrates that panel analysis produces substantively different results. Knack and Kropf (2003) study the 1996 presidential election.[19] While they agree with our finding that punch

[18] Knack and Kropf (2003, 888–89) find a negative effect of turnout surges on the residual vote rate and admit puzzlement. Our findings here confirm the correctness of their conjecture about the solution to their puzzle—that it is due to an omitted variable bias in cross-sectional analysis.

[19] Comparison of results across the two papers requires a small application of arithmetic, since Knack and Kropf's omitted voting technology category is Votomatic punch cards, whereas ours is mechanical lever machines. The discussion here is based on converting the coefficients to the same residual category. In addition, we combine Votomatic and DataVote counties into a single "punch card" category, whereas Knack and Kropf continue the distinction. Our (unreported) preliminary analysis revealed that the residual vote effects of the two types of punch cards could not be distinguished. Knack and Kropf also discover that the residual vote rate of DataVote counties in 1996 was not statistically different from Votomatic counties. Finally, Knack and Kropf do not report results for residual votes in senatorial and gubernatorial races, so we cannot compare those results.

cards are by far the inferior technology, our results rank the remaining technologies differently. We find that in presidential elections hand counted paper ballots are the least prone to losing votes, whereas they find mechanical lever machines to perform the best; in our analysis, mechanical lever machines are in the middle of the pack.

In terms of the policy debate now raging about how best to upgrade from antiquated voting equipment, our panel study provides different answers to policy makers about how to proceed than those provided in cross-sectional analyses. Our analysis favors abandoning mechanical lever machines in favor of optical scanners; the cross-sectional analysis would suggest holding on to these 900-pound beasts.

We estimate that the difference between the best technologies and the worst is about 2% of ballots cast. A margin of error that large must surely raise doubts about the outcome of many elections, past and future.

As a result of the election recount in Florida and studies done subsequently, including this one, it is now clear that close elections are ambiguous elections—even after the counting is done. This raises several troubling questions for democratic legitimacy. Do ambiguities in the counting of ballots themselves make people feel that their votes do not count? Will future legal battles lead to more public cynicism?

These problems extend further to the international efforts to propagate democracy. The international community widely criticized the conduct and legitimacy of 2002 Zimbabwe election. In defense of his election, President Mugabe cited the contentious 2000 U.S. presidential election. Lowering the rate of error attributable to voting technologies will improve the legitimacy of American elections, at home and abroad.

A more subtle implication of our analysis is that federalism and the decentralization of electoral administration in the United States produces political inequality. Local election officials retain most of the authority for the administration of elections in the United States. Until recently, they have been subject to little federal regulation. As a result, equipment usage and many other aspects of administration vary greatly in the United States. The consequence is that Americans' votes are not all counted the same.

Our results show this in two ways. First, voting equipment produces inequities. Not all voting equipment is equally reliable. Local election administrators choose technologies: they are the consumers (or demanders) of voting equipment. Over the century of its existence, the highly decentralized market for voting equipment in the United States has not driven error rates down. That is because new voting machines have been procured in recent years to speed up the count, not to improve accuracy. The oldest technology, hand-counted paper, performs the best. Punch cards are a relatively recent innovation (1960s), and they are the worst. The newest technology (DREs) does not show clear improvements over paper or optical scanning or, at the top of the ticket, lever machines.

Second, the incidence of uncounted and spoiled ballots depends strongly and systematically on "county," in addition to equipment. Our panel analysis revealed that almost all of the variation explained in the residual vote is not due to demographics, political factors, or technology, but to "county." We conjecture that this county effect is substantially the result of local *institutions* of electoral administration, such as the administration of polling places or advance instruction to voters. The data point to an administrative story, because demographic factors, like race and income, and political factors, like electoral competition and state, explain only a very small percent of the variability in residual votes. There are likely other factors that affect residual votes which would be difficult to measure at the county level in a study like this, such as political culture, the activity of political parties, etc. Why county matters for the rate of uncounted and spoiled ballots is, as yet, unexplained, and an important subject for future research. In addition, the most important demographic influence affecting residual vote is not a characteristic of voters, but of place—population. Rural counties have significantly higher residual vote rates than urban and suburban counties. This is an undeniable difference, quite apart from race, income, age, electoral competition, and equipment.

Since the 1960s, the doctrine of political equality has become the law of the land. The courts and Congress have asserted this principle repeatedly in the areas of districting and voter registration. While the degree of intentional discrimination is less clear with voting equipment, there is clear evidence that votes are not counted the same by different technologies. In *Bush v. Gore*, the U.S. Supreme Court skirted this issue. But the issue will surely resurface, as it goes to one of the core conflicts in the American polity—the conflict between the broad principles of political equality as it has been asserted by the national government and the practice of federalism and decentralized administration of government.

Appendix. States included in residual vote analysis

State	Counties	1988	1992	1996	2000	Total
Alabama	67	0	0	0	0	0
Alaska	40	27	40	40	29	136
Arizona	15	15	15	15	15	60
Arkansas	75	0	0	0	27	27
California	58	57	58	58	58	231
Colorado	63	62	63	63	0	188
Connecticut	8	8	8	8	8	32
D.C.	1	1	1	1	1	4
Delaware	3	3	3	0	0	6
Florida	67	0	66	66	67	199
Georgia	159	0	0	154	159	313
Hawaii	5	4	4	4	4	16
Idaho	44	44	44	43	44	175
Illinois	102	102	101	102	102	407
Indiana	92	90	86	89	83	348
Iowa	99	0	82	98	99	279

Appendix. *continued*

State	Counties	1988	1992	1996	2000	Total
Kansas	105	0	82	79	94	255
Kentucky	120	116	115	112	107	450
Louisiana	64	0	55	64	62	181
Maine	16	0	0	0	0	0
Maryland	24	23	23	24	24	94
Massachusetts	351*	351	351	351	351	1,404
Michigan	83	19	20	20	29	88
Minnesota	87	56	76	78	79	289
Mississippi	82	0	60	2	3	65
Missouri	115	0	0	0	114	114
Montana	57	54	55	56	51	216
Nebraska	93	93	93	91	91	368
Nevada	17	17	17	16	17	67
New Hampshire	10/234*	7	7	6	225	245
New Jersey	21	15	17	19	21	72
New Mexico	33	27	28	31	33	119
New York	62	61	61	61	62	245
North Carolina	100	0	25	32	29	86
North Dakota	53	53	53	53	53	212
Ohio	88	88	88	88	88	352
Oklahoma	77	76	77	0	7	160
Oregon	36	29	36	36	36	137
Pennsylvania	69	0	0	0	1	1
Rhode Island	5	0	0	0	0	0
South Carolina	46	45	39	43	45	172
South Dakota	66	0	0	62	65	127
Tennessee	95	0	11	11	11	33
Texas	254	0	0	0	153	153
Utah	29	29	29	29	29	116
Vermont	14/246*	8	0	8	246	262
Virginia	135	0	0	0	134	134
Washington	39	39	38	39	37	153
West Virginia	55	55	0	55	0	110
Wisconsin	72	0	0	0	0	0
Wyoming	23	19	20	21	22	82
Total		1,693	2,047	2,228	3,015	8,983

*Massachusetts has 351 municipalities, which is the universe for analysis for all years. New Hampshire has 10 counties and 234 municipalities; counties are the universe in 1988, 1992, and 1996; municipalities are the universe in 2000. Vermont has 14 counties and 246 municipalities; counties are the universe in 1988, 1992, and 1996; municipalities are the universe in 2000.

Acknowledgment

We gratefully acknowledge the financial support of the Carnegie Corporation of New York, which underwrote the first phase of the Caltech/MIT Voting Technology Project. Lee Carpenter, Michele Nyein, Meena Untawale, James Wagner

Catherine Wilson, and Melanie Wong provided invaluable research assistance. We appreciate the helpfulness of election administrators and Election Data Services in our data collection efforts.

Manuscript submitted July 8, 2003
Final manuscript received June 4, 2004

References

Asher, Herb. 1982. "The Effect of Voting Systems on Voter Participation." Presented at the annual meeting of the Midwest Political Science Association, Milwaukee.

Brace, Kimball, and Doug Chapin. 1993. "Lies, Damned Lies, and Statistics: Observations on Using Election Data." In *The Election Data Book*, ed. Kimball Brace. Lanham, MD: Bernan Press, pp. xiii–xviii.

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. Survey Research Center and Institute of Governmental Studies, University of California at Berkeley.

Bullock, Charles S., and Richard E. Dunn. 1996. "Election Roll-Off: A Test of Three Explanations." *Urban Affairs Review* 32(1): 71–86.

Caltech/MIT Voting Technology Project. 2001. *Voting: What Is, What Could Be.* Pasadena and Cambridge: Caltech and MIT.

Center for American Politics and Citizenship. 2003. "Characteristics of Contemporary Voting Machines." Accessed 14 November 2003. <http://www.capc.umd.edu/rpts/md_evote_ContempVotingMach.pdf>.

Darcy, R. 1986. "Position Effects with Party Column Ballots." *Western Political Quarterly* 39(4): 648–62.

Darcy, R., and Anne Schneider. 1989. "Confusing Ballots, Roll-Off, and the Black Vote." *Western Political Quarterly* 42(3): 347–64.

Election Data Services. 2002. "Election Data Services Inc. Unveils 2002 Voting Equipment Study." Press release, 9 October 2002.

Fischer, Eric A. 2001. "Voting Technologies in the United States." *CRS Report for Congress* RL30773.

Florida Governor's Select Task Force on Election Procedures, Standards, and Technology. 2001. "Revitalizing Democracy in Florida," Accessed 4 December 2001. <http://www.collinscenter.org/info-url2660/info-url_list.htm?cat=FINAL%20REPORT>.

Garner, Phillip, and Enrico Spolaore. 2001. "Why Chads? Determinants of Voting Equipment Use in the United States." Mimeo. Department of Economics, Brown University.

Georgia Secretary of State. 2001. "The 2000 Election: A Wake-Up Call For Reform and Change." Accessed 4 December 2001. <http://www.sos.state.ga.us/elections/2000_election_report.pdf>.

Greene, William H. 2000. *Econometric Analysis*, 4th ed. Upper Saddle River, NJ: Prentice Hall.

Herron, Michael C., and Jasjeet S. Sekhon. 2001. "Overvoting and Representation: An Examination of Overvoted Presidential Ballots in Broward and Miami-Dade Counties." *Electoral Studies* 22(1): 21–47.

Iowa Secretary of State. 2001. "Iowa's Election 2000: Facts, Findings, and Our Future." Accessed 4 December 2001. <http://www.sos.state.ia.us/pdfs/ElectionReport.pdf>.

Kimball, David C., Chris T. Owens, and Katherine M. Keeney. Forthcoming. "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, pp. 135–50.

Knack, Stephen, and Martha Kropf. 2003. "Voided Ballots in the 1996 Presidential Election: A County-Level Analysis." *Journal of Politics* 65(3): 881–98.

Maryland Special Committee on Voting Systems and Election Procedures in Maryland. 2001. "Report and Recommendations." Accessed 4 December 2001. <http://www.sos.state.md.us/sos/admin/pdf/reportall1.pdf>.

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.

Michigan Secretary of State. 2001. "Uniform Voting in Michigan: A Report to the Legislature." Accessed 4 December 2001. <http://www.sos.state.mi.us/election/elecadmin/2001web/uniformvotinginmichigan.pdf>.

Missouri Secretary of State. 2001. "Making Every Vote Count." Accessed 4 December 2001. <http://mosl.sos.state.mo.us/sos-elec/bluntcommission/mevc-report.pdf>.

National Commission on Federal Election Reform. 2001. Accessed 4 December 2001. <http://www.tcf.org/Publications/ElectionReform/99_full_report.pdf>.

Nichols, Stephen M., and Gregory A. Strizek. 1995. "Electronic Voting Machines and Ballot Roll-Off." *American Politics Quarterly* 23(3): 300–18.

Posner, Richard A. 2001. *Breaking the Deadlock: The 2000 Election and the Courts.* Princeton: Princeton University Press.

Shocket, Peter A., Neil R. Heighberger, and Clyde Brown. 1992. "The Effect of Voting Technology on Voting Behavior in a Simulated Multi-Candidate City Council Election: A Political Experiment of Ballot Transparency." *Western Political Quarterly* 45(2): 521–37.

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(1): 46–60.

Traugott, Michael. 2003. "Why Electoral Reform Has Failed: If You Build It, Will They Come?" In *Rethinking the Vote: The Politics and Prospects of American Election Reform*, eds. Ann N. Crigler, Marion R. Just, and Edward J. McCaffery. New York: Oxford University Press, pp. 167–84.

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

Wattenberg, Martin P., Ian McAllister, and Anthony Salvanto. 2000. "How Voting is Like Taking an SAT Test: An Analysis of American Voter Rolloff." *American Politics Quarterly* 28(2): 234–50.

White, John P. 1960. *Voting Machines and the 1958 Defeat of Constitutional Revision in Michigan.* Ann Arbor: Institute of Public Administration, The University of Michigan.

Stephen Ansolabehere (sda@mit.edu) is Elting R. Morison Professor of Political Science, Massachusetts Institute of Technology, Cambridge, MA 02139-4307. Charles Stewart III (cstewart@mit.edu) is professor and head of the department of political science, Massachusetts Institute of Technology, Cambridge, MA 02139-4307.