Voting Technology and Residual Votes in the United States

Stephen Ansolabehere
Professor of Political Science
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139
sda@mit.edu
617.253.5236

Charles Stewart III
Professor of Political Science
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139
cstewart@mit.edu
617.253.3127
Facsimile: 617.258.6164

Please direct all correspondence to Stewart

June 1, 2004

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 valuable research assistance. We appreciate the helpfulness of election administrators and Election Data Services in our data collection efforts.

Abstract

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 U.S. 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.

Voting Technology and Residual Votes

xxxxx

The election of the president 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. And, 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. Legal and political problems of determining voter intent permeated the recount process throughout Florida.

Two questions: (1) performance. How well does election administration capture voters' intentions and wishes? Barriers to voting, such as registration, have the effect of lowering participation in the neighborhood of 2 to 3 percent of eligible voters. How does this compare to other sorts of performance factors? (2) fairness. Are some people's votes not counted because of election administration? What degree of unfairness are we willing to tolerate? How does this vary across jurisdictions? Specifically, are voters in jurisdictions that use punch cards less likely to have their votes counted?

The method used to cast and count ballots is surely the most mundane aspect of elections, but the possibility that equipment differs systematically immediately raises questions about the integrity of the electoral process in the United States. How bad are the methods for casting and counting votes in the United States? Equally troubling are questions of political equality and fairness. Are some technologies better at producing a more complete count of the vote? Does the lack of uniform voting equipment in the country mean that some voters are more likely to have their ballots counted than others?

xxxxx

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 of the performance of voting system 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, covering all counties in the United States, the years 1988 to 2000, and elections for president, US Senate, and governor. We examine two different dependent variables: the difference between total ballots cast and ballots cast for a specific office (called the residual vote) and the difference between ballots cast for president and ballots cast down the ballot (called rolloff or voter fatigue in past studies). We exploit panel structure of the data to hold constant a wide variety of town- and county-level factors that affect voting patterns, such as

[1] Several federal commissions issued substantial reports on the election process, most notably National Commission on 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).

demographics and administrative practices. We measure the effect of changes in technology on changes in residual vote and rolloff 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 and rolloff. The difference between the best performing and worst performing technologies is as much as 2 percent 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) also show promise, though they have a statistically higher residual vote rate than hand-counted paper and optically scanned ballots.

A somewhat different question is what explains the lion's share of the variation in residual votes and rolloff. Most of the variation – nearly 60 percent – is accounted for by the county, rather than by electoral competitiveness, demographics, or technology. Technology, competition, and demographics combined explain only about 15 percent of the variation in residual vote rates. Including indicators of county increases the percent explained to 70 percent. 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 became pervasive, not just an urban phenomenon (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 exploits cross-sectional variation only.

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 (1988) found a consistent 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 ballot type (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 over-votes (i.e., an excess of legal votes) and more under-votes (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 two cross-sectional studies of a national scope—Knack and Kropf (2001) study the 1996 election and Brady, et al, (2001) study the 2000 election.

This paper advances the methodology of past research in three ways. First, our study spans a long time frame, from 1988 to 2000, and we examine the entire nation. The results do not reflect the circumstances of one place or time. All of the previous research has been devoted to studying cross-sections of elections, and typically for a small range of political jurisdictions.

---

[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.

Second, we exploit the panel structure of electoral data. Use of voting technologies varies considerably across counties, but also within counties over time. This presents an opportunity to explore the relative performance of voting technologies across space and 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 Asher (1982) examines the effects of switching technology within counties, and he studied a handful of counties in Ohio.[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 and the rolloff 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.

Third, we examine two different indicators of "errors" in vote recording. Most past research focuses only on the difference between votes cast for president and votes cast for other offices down the ballot, i.e., rolloff. We examine the difference between total votes and votes for specific offices, i.e., residual votes, as well as rolloff. To the degree that problems with voting technologies may affect the votes cast on all races on a machine—most obviously when a machine malfunctions or when the ballot is too confusing—rolloff will mismeasure the magnitude of unrecorded votes for a particular office and the effects of technology.

The remainder of this paper is organized as follows. Section 1 describes different voting technologies used in the United States. Second 2 discusses our measure of uncounted votes and

---

[3] Mathew (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.

the factors that might explain this variable, including voting technologies. Section 4 reports the results of a series of panel regressions that assess the relative performance of voting technologies. Section 5 concludes with a discussion of the policy implications of our findings and directions for further research.

1. Variability in Voting Technologies

Five types of technologies are used to cast and count votes in the United States today. Three technologies, hand-counted paper, punch cards, and optically scanned paper, are based on paper ballots; two technologies, lever machines and electronics machines, involve machines that directly record, store, and tally the voters' preferences.

The oldest technology is the paper ballot. To cast a vote, a person makes a mark next to the name of the preferred candidates or referendum options. Paper ballots are counted manually.

Mechanical lever machines, introduced in the 1890s, are steel booths that the voter steps into. A card in the booth lists the names of the candidates, parties, or referenda options, and accompanying each option is a switch. The voter flicks the switch of their preferred options for each office or referendum. The voter, then, pulls a large lever, which registers their votes on series of counters inside the machine. At the end of the voting day, the election precinct workers record the tallies from the internal counters in each of the machines.

Punch card machines, introduced in the 1960s, use a form of paper ballot, and this technology automates the counting process. Upon entering the polling place the voter is given a paper ballot in the form of a long piece of heavy stock paper. There are two variants of the punch card – one, the DataVote, lists the names of the candidates on the card; the other (Votomatic) does not. For Votomatic machines, the voter inserts the card into device that shows the voter a list of candidates for each office and alternatives for each ballot questions. The voter's card is aligned with the appropriate candidates and ballot questions. The voter uses a metal stylus to punch out the perforated rectangle beside the candidate of choice. With DataVote machines the voter punches a hole in an unperforated card to indicate a choice. When finished, the voter removes the card and puts it in the ballot box. At the end of the day, the election workers put the cards into a sorter that counts the number of holes next to each candidate.

Optically scanned ballots, also known as "marksense" or "bubble" ballots, offer another method for automating the counting of paper ballots. The voter is given a paper ballot that lists the names of the candidates and the options for referenda; next to each choice is small circle or an arrow with a gap between the fletching and the point. The voter darkens in the bubble next to the preferred option for each office or referendum, or draws a straight line connecting the two parts of the arrow. The ballot is placed in a box, and, at the end of the day, counted using an optical scanner. Some versions of this technology allow the voter to scan the ballot at the polling place to make sure that he or she voted as intended, or at least did not produce an over-vote.

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 presents voters with a panel of push buttons. The voter selects the button next to each candidate, and when finishes pushes the "VOTE" button. This is analogous to voting on a lever machine. A second variant presents voters with a touchscreen 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. Some electronic machines allow voters to check their ballots at the end of the session; others do not.

Each type of technology involves many variations based on specifications of manufacturers, ballot formats, and implementation. Our focus is on the five main types of machines. 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 occur most commonly Massachusetts, Michigan, Maine, New Hampshire, and Vermont, where town governments usually administer elections.

We examine the variation in usage across counties and over time. The bulk of our data come from Election Data Services (EDS), the leading vendor of data on elections and voting equipment. This data includes information about election returns, turnout, and voting equipment. We augment the EDS data with data from state, county, and municipal election officials, particularly for the 2000 election.

The voting technology data do not allow us to distinguish between the precise makes and models of voting technologies that are used by local jurisdictions, usually because the states themselves reported highly aggregated categories (e.g., "optical scan" instead of "Optech Eagle IIIP.") Therefore, we are unable to address the relative performance of precise implementations of these broad technology categories.[4]

Xxx From Here xxxx

Even without this additional level of detail, the pattern of equipment usage across the United States looks like a crazy quilt. Americans vote with a tremendous array of types of equipment. In the 2000 presidential election, one in five voters used the "old" technologies of paper and levers—1.3 percent paper and 17.8 percent levers. Punch cards were used by just over

[4] 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. 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 did test for correlation between the percent of ballots cast absentee and the county residual vote rate. It is statistically insignificant.

one-third of voters (34.4%). Over one-in-four used optically scanned ballots. One in ten used electronic devices. The remaining 8.1 percent were in counties that used a mix of systems.

Xxx to Here xxx

Within states there is typically little uniformity. This is illustrated in Table xx1, which reports the percent of the population in each state that use the various types of voting technologies. Some states use 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 equipment (Hawaii, Oklahoma, and Rhode Island). At the other extreme, states such as Arkansas, Georgia, Indiana, Kansas, Maryland, Mississippi, North Carolina, South Carolina, and Tennessee do not have one dominant voting technology. In some states, such as Arkansas, North Carolina, Pennsylvania, Texas, and Virginia, at least one county uses each type of technology available.

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, optically scanned ballots and DREs have grown from a combined 3.2 percent of the population covered to 38.2 percent of the population covered. There has been little change in the use of punch cards. Paper ballots have fallen from 9.7 percent of all people in 1980 to just 1.3 percent in 2000. Lever machines, by far the dominant mode of voting in 1980, covered 43.9 percent of the electorate. Today, only 17.8 percent of people reside in counties using lever machines.[5]

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 faster tabulation of ballots. Punch cards, optical scanners, and DREs use computer technology to produce a speedy and,

---

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

hopefully, more reliable count. Some locales have, however, gone back to the older technologies. For example, several towns in Massachusetts went back to lever machines after a difficult recount in 1996 involving punch card ballots.

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 adopted new technologies (1476 out of 3155 counties). And, today each of four technologies (lever machines, punch cards, optical scanning, and electronic machines) are widely used across counties.

2. Uncounted Ballots: Measures and Causes

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."[6]

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

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

---

[6] 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; see National Commission on Election Reform (2001), H.R. 3295 (Ney-Hoyer Bill), and S.565 (Dodd bill) of the 107th Congress. 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, as occurs if a lever machine is broken or punch card machine is jammed with chad.

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.

Residual Votes as a Measure of Uncounted Votes

To calculate residual votes, we assembled data on the total number of votes cast in each county or municipality and the total number of ballots counted with a valid vote for president, for U.S. Senate, and for governor.

The residual presidential vote in the average county equaled 2.3% from 1988 to 2000.[7] Because county populations vary dramatically, this does not equal the fraction of people who cast an under- or over-vote 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.[8]

The residual gubernatorial and senatorial vote rates are somewhat higher. The county average residual vote rates in gubernatorial and senatorial elections is 4.2 percentage points, and the percent of all ballots cast (population weighted county average) is 4.1 percent. The standard

[7] We exclude from this calculation counties in which the county reports shows more presidential votes cast than total ballots cast, that is, cases with negative residual vote rates. This affects about 2 percent of the counties in our analysis.

[8] Logarithmic transformation of the data eliminates the skew, and makes the distribution nearly normal.

deviations are 3.5 percentage points for the county average and 2.9 percentage points for the population weighted data. The skew statistic is 2.8.

The residual vote is not a pure measure of uncounted votes. The residual vote rate is too generous a measure of uncounted ballots because it includes abstentions.

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. Precise figures on intentional abstention do not exist, because ballots are secret. However, exit polls and post election surveys indicate that from 1988 to 2000 approximately one-half of one-percent of voters intentionally abstain from voting for president in the voting booth.[9] The residual vote rate is 2.2 percent of total ballots cast. That leaves approximately 1.7 million votes (1.7 percent of total ballots cast) "lost" because of technological malfunctions and voter confusion.

Second, the residual vote is the dependent variable and noise in that measure due to variation in abstention rates will not produce bias. Noise in the dependent variable lowers efficiency, and makes it less likely to find differences across technologies.

Third, we ultimately care about whether technology leads fewer votes to be recorded. We care less about intentionality, than about the extent to which technology interferes with voters' attempts to vote. Some of what we care about is actual machine breakdown. But, some of what we care about is poor overall design that intimidates, confuses, and, ultimately, discourages voters. In fact, psychological factors might mean that some technologies produce higher rates of intentional abstention. Human factors research in the area of technology is relatively new, and in the area of voting equipment the research is nascent. Some technologies

[9] Responses to the American National Election Study help to provide an order-of-magnitude estimate of the frequency of conscious abstention. Among respondents who reported having voted, 0.3% reported not voting for president in 1988, 0.7% in 1992, 1.0% in 1996, and 0.3% in 2000. Therefore, the rate of actual abstention in presidential elections is roughly ½%.

might encourage people to engage in the relevant activities more than other technologies. A voting machine, for example, may be sufficiently confusing or intimidating that the voter elects not to vote, as some laboratory analyses have documented (Roth, 1988; Shocket, Heighberger, and Brown 1992).

To put this matter differently, the proof of the usefulness of residual votes 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, once we have done our best to hold other factors constant. In fact, there are substantial differences, as we show below.

Two other measures were possible, but we judged them inferior. Most prior research on this subject has examined rolloff. Also, overrvotes were widely cited in Florida.

Rolloff is the difference between the vote for president and the vote for another office, down the ballot, such as governor or Senator. The objections above apply equally to rolloff – that is, to all past research. More importantly, though, there is less information in rolloff than in the residual vote. Rolloff only captures failure to vote for one office down the ballot. It does not capture technology failures that affect voting for president. And, rolloff misses any technology failures or confusion that lead to a voters' entire ballot not being counted, such as occur with a general machine failure. The analysis of residual votes for different offices encompasses rolloff, because rolloff is the difference between the residual governor vote or the residual senator vote and the residual presidential vote.

Overvotes are also too restrictive a measure. Overvotes occur when someone votes twice for the same office. Such double votes are only part of the problem. Technology can enable or interfere with voting in many ways, especially general voter confusion. Indeed, voter confusion may account for most of what occurs. The residual vote will capture some of this effect; the overvote will not. Very few jurisdictions report enough information to construct the overvote and other measures. Researchers wishing to use measures other than residual vote and roll-off will be forced to study a very small subset of cases.

Explanations for Residual Votes

Having received a ballot and proceeded to a voting booth, a voter may not have a vote recorded for a particular office for three general types of reasons—reasons relating to machines, to individuals, and to local administrative practices.

Machine effects. 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 an "odometer" that records the votes cast on a mechanical lever machine stops working 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—although vendors are increasingly adding "paper backup" features to their DREs to address this problem.

One of the failures of Votomatic punch cards in Palm Beach County was a mechanical failure. In that case, controversies over "dimpled, pregnant, and hanging chads" were really about the failure of the punch cards to perform as designed.[10] When a voting machine fails mechanically, election officials will record a voter <u>who intends to vote</u> as having received a ballot, but when the ballot is counted (or the machine memory retrieved), the vote will not register.

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 the case of Palm Beach County, the flaw was the infamous "butterfly ballot" that apparently confused voters in the presidential election. (See Darcy 1986; Darcy and Schneider 1986; Bullock and Dunn 1996; Wattenberg, McAllister, and Salvanto 2000; Wand, et al 2001; Herron and Sekhon 2001.) 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.

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

<u>Individual reasons</u>. Some voters may have greater difficulty voting than other voters. Literacy and language are common explanations for such problems (Posner 2002). Quite apart from what machine is being used, a county may have higher residual votes because it has more voters with low literacy. Direct measures of literacy are not available; however, education level, income level, and several other demographic characteristics are correlated with literacy. Also,

[10] 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.

percent Hispanic and percent foreign-born may indicate populations that are likely to encounter language problems. Factors relating to aging, such as poor eyesight, might also affect residual vote rates across all technologies.

One reason why a voter might not have a vote register for an office is that the voter intends not to vote in that race, still choosing to vote in other races. As we have discussed, intentional abstention accounts for about one-half of one-percent of total ballots cast. Intentional abstention also varies across people, and demographic and geographic indicators help to account for some of the variation in residual votes that occurs through abstention.

<u>Jurisdiction-specific factors.</u> Local jurisdictions vary significantly in how they administer elections, and some of this variation likely affects the degree to which votes are actually counted. Analysis of the public finances of county election offices suggest 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.[11] Residual votes, then, will likely vary systematically from county to county. Some of this is predictable on the basis of county population.

---

[11] 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 on, producing the appearance in some cases that more ballots had been counted than actually cast. In 1996, for instance, 26 of Kansas's 105 counties report more total ballots cast for president than total ballots cast overall. In explaining to us why this discrepancy occurred, an election official from Kansas wrote us the following in a personal communication (dated 28 June 2001):

> Figures on the number of persons who went to the polls are taken from unofficial reports we collect just because it's useful information for our office and political researchers to have. Because it is unofficial, counties aren't required to report it, and some of it is less reliable than official reports. For instance, turnout numbers are sometimes election-night figures that don't take into account provisional/challenged ballots that haven't been resolved as of election night. Some counties don't go back and add them back into their reports to us.

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. Wealthier counties, in particular, are more likely to have more resources to staff elections properly.

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 there is high turnout, not only will more voters need instruction, but poll workers will likely have less time to instruct voters on the way to use the voting equipment.

3. 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 U.S., 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—

ballots for which no vote is cast or counted—across counties using different sorts of technologies. Second, we examine how changes in technologies within localities over time explain changes in the incidence of ballots that are spoiled, uncounted, or unmarked. If existing technology does not affect the ability or willingness of voters to register preferences, then the incidence of over- and under-votes 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 collected election returns from governor 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 states (Massachusetts and Vermont) we were able to collect the requisite data for this analysis, and we have included their town-level data.[12] In total, there are over 20,000 county-year observations in the data set. In the Appendix we report which states fall within our sample during the elections for this time period, in addition to average state residual vote rates for 2000.

All told, there are nearly 9000 cases for which we have been able to identify the machines used and to collect data on total ballots and presidential ballots cast. There are approximately 11,000 cases for which we can identify the equipment used and calculate the

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

residual vote for senators and governors. The large number of observations produces high levels of precision in estimating average residual vote rates associated with each machine type. The data come from state election offices, typically within the office of the Secretary of State, and from Election Data Services.

Beyond equipment, many other factors may explain rates of uncounted votes and abstentions. As discussed earlier, turnout, county wealthy, 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 take two approaches. First we include indicators for demographics as well as for state and year. State and year effects capture the competitiveness of elections within the state. They also capture the effects of state laws defining what counts as a vote. Our data for county level demographics come from the U.S. Census of 1990. These include population, median income, percent over 65, percent 18-25 years old, percent white, and percent Hispanic.

Second, and more importantly, 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 voting, 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. 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.

4. Results

Basic descriptive statistics about residual votes for various technologies capture many of the principle results of this investigation. Table 4 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.

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 percent. In gubernatorial and senatorial elections, those voters average a residual vote rate of 2.1 percent and 3.0 percent, 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 studied. Third, 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%)

The statistics in Table 4 are based on cross-sectional analysis of the data, whereas our real interest is in how residual vote rates change when voting technologies change in counties. Before turning to a series of regressions, it is instructive to look directly at the experience of counties that have moved away from the most common obsolete technologies (lever machines and punch cards) since 1988.

What happened in counties that used levers or punch cards in 1988 in the subsequent three presidential elections? About half the lever machine counties and a quarter of punch card counties adopted other technologies after 1988. How did the residual vote rate change in counties that changed, compared to counties that stood pat?

Table 5 provides an answer to this question. The top part of the table presents three types of counties. The first row shows counties that used lever machines in 1988 and stayed with lever machines in 1992, 1996, and 2000. The second row represents counties that had lever machines in 1988 but switched to optical scanning in one of the subsequent elections. The third row represents counties that had lever machines in 1988 but switched to DREs in one of the succeeding elections.

The first column presents the average change in the residual vote rate from 1988 to the current year. (The standard deviation is in parentheses.) We then average over all years. Consider, for example, a county that had levers in 1988 and 1992, but scanners in 1996 and 2000. The first row includes the observed change in the residual vote rate from 1988 to 1992 for such a county. The second row contains the average change in the residual vote rate from 1988 to 1996 and from 1988 to 2000—the two elections in which the county used scanners.

On average, counties that kept their lever machines saw a slight improvement in their residual vote rates from 1988 to 1992, 1996, and 2000, by -0.3% on average. Counties that switched to scanners had their residual vote rates fall by even more than the counties that stuck with levers, by -0.6% on average. Counties that switched to DREs saw their residual vote rates increase above the residual vote rate that they had in 1988, by +0.2% on average.

A similar story emerges among the counties that started out with punch cards. Counties that stuck with punch cards enjoyed a 0.4% decrease in residual vote rate. Those that switched to optical scanning saw a four-fold improvement over this baseline. The few counties that switched to DREs also saw a reduction in residual vote, but not nearly as much as counties that stood pat.

We see in this analysis some patterns that will bear further scrutiny. Switching from an old technology to scanning leads to a significant improvement in the residual vote rate. Switching to DREs appears to make the residual vote rate worse, but not by much.

xxxx We need to rewrite From Here xxxx

To perform this analysis more generally, we estimated the fixed effects regression we previously described. There is a fixed effect (i.e., a dummy variable) for each county, which captures unmeasured county factors – such as administration and demographics. Table 6 reports the results of these regressions.[13] The first two columns compare non-fixed effects and fixed effects estimation of the presidential ballots; the last two columns compare non-fixed effects and fixed effects estimation of a combination of the gubernatorial and senatorial ballots.[14] In all

[13] We used the STATA command areg to perform these regressions.

[14] We combined the gubernatorial and senatorial analysis for the sake of simplicity. Separate analysis of gubernatorial and senatorial ballots show substantially similar results.

regressions, counties are weighted by overall turnout, so the interpretation of the dependent variable is the percent of ballots cast. Also, because much of the literature looks at cross-sectional regressions, Table 6 presents the OLS regression without fixed effects, which mainly reflects across county correlation in residual votes and technology and other factors. Comparison of results with and without fixed effects is instructional for understanding the advantages of the panel design of our analysis.

Estimated effects of voting technology in the panel analysis reflect how the residual vote in each county changes when the county changes its technology. The indicators for each technology equal 1 in the years when the county uses that technology and 0 when the county does not. Because the 6 technology categories are linearly dependent with the constant in the regression, we exclude one of the categories. We chose to exclude the oldest "modern" technology -- lever machines. Therefore, the equipment coefficients measure how much higher or lower is the average residual vote of that equipment type, compared to lever machines.

Xxx I'm unclear on the numbers in this paragraph:

The OLS regressions show relatively higher residual voting rates of counties using punch cards and DREs. In the non-fixed effects estimates for president, punch cards produced an average 0.77 greater residual vote rate, once we've also controlled for shifts in technology, the presence of the governor or senator on the ballot, and turnout. DREs produced a rate that was 0.8% higher. The paper and optical scan coefficients show virtually no difference from lever machines. Overall, these results are virtually identical to the zero-order effects that can be discerned in Table 4.

Turning to the fixed effects estimates, the results are quite different. Beginning with the equipment effects, an important consequence of introducing the fixed effects is that the differences between lever machines (on the one hand) and paper ballots and optically scanned ballots (on the other hand) become more pronounced; the difference between lever machines and DREs becomes significantly less.

Paper ballots turn out to be the champion in presidential ballots in this fixed effects analysis, producing 1.3% fewer residual vote than mechanical lever machines. Next in efficacy is optical scan ballots, which produced 0.5% 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. Bringing up the rear, by a significant amount, are punch cards, which produced 0.8% 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 (0.0082+0.0045 =) 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 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 one percent of presidential turnout, which is a significant proportion of election-to-election variability in turnout.

Turning to gubernatorial and senatorial elections, the results are quite different from those associated with presidential voting. 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 bottom of the heap. 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 simply presents in graphical form the fixed effects voting equipment coefficients from Table 6. (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) does not promise such unambiguous gains.

Xxxxx rewrite to Here xxxx

Returning to Table 6, we also see some interesting results that also pertain voting technologies. First, it seems 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. Before dismissing the importance of a shift in technology, we should note two problems with this variable that may attenuate the estimated effect. First, because the data only contain information about gross categories of voting technologies, the technology shift variable can only measure changes in technology categories. Consequently, there is measurement error in this variable that undoubtedly leads to a downward bias in the coefficient. Second, and more substantively, 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 correctly. 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 new technologies minus the effects of greater awareness due to voter education.

Finally, the behavior of the turnout variable reveals an important subtle effect that turnout and size of the electorate have on voter errors. In the simple non-fixed effects analysis, turnout is primarily measuring cross-sectional differences in the population of the county. The negative coefficient in the presidential analysis quantifies the moderate negative relationship between jurisdiction size and residual vote in the cross-section. The addition of the county-specific dummy variables takes care of the cross-sectional relationship between residual vote and population. The turnout variable, therefore, measures the effects on residual vote that occur because of fluctuating turnout within jurisdictions. For all kinds of elections in this analysis, a surge in turnout within a county is associated with more ballots not being counted.

Xxxx From Here xxxx

Alternative specifications are possible. One reviewer suggests using lagged dependent variables to address possible dynamics and other omitted variables. First, residual votes might involve a dynamic component. For example, a county that had an unusually high residual vote rate might redouble its election administration efforts in the succeeding election. Second, there might be important demographic characteristics that change dramatically within a county over a decade. In the panel model the fixed effects capture nearly all of the cross-sectional effect of the demographic characteristics measured by the U.S. census – including income, education, and race. The reason these factors likely did not matter much is that they change relatively little and very slowly within counties over a decade, so almost all of the only variability to exploit to estimate the effects demographics is captured in the county fixed effects. Nevertheless, there might be some unmeasured factors that are important but not captured in the decennial census. Lagged values of the residual vote might capture these factors.

Including the lagged residual vote alters the panel estimates somewhat, but the same substantive conclusions still emerge. The effect of lagged residual vote is statistically very strong, with a coefficient of .16 and a t-statistic of 11.xxx. The coefficient on turnout remains negative, as expected. The Finally, county fixed effects remain very important. Most of the variation in the residual vote occurs across counties. Why some counties regularly have substantially higher residual vote rates, then, is a puzzle.

Xxx To here xxxx

The importance of holding the county constant, using fixed effects regression, can also be seen by examining the county-level effects that are estimated by the regression routine. These dummy variables, one for each county, can be thought of as the "baseline" residual vote rate for each county, taking into account a host of unmeasured variables, such as county elections budget, voter political knowledge, machine maintenance routines, etc. 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 factors are can be seen by simply comparing the r^2 statistics of the OLS estimates with the fixed effects estimates in Table 6. From a variance-explained perspective, it appears that most of what influences whether votes get counted as cast 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 (2001), Brady, Buchler, Jarvis, and McNulty (2001), and others have recent done cross-sectional analysis similar to what is presented in this paper. The analysis presented in this paper suggests that unless researchers are lucky enough to control for the relevant non-technological, jurisdiction-specific factors affecting residual vote rates, the risk of encountering omitted variables bias is high. Not only does such bias affect the estimated size of the technology effect, but can affect other variables, too, as the sign change on the population variable in Table 6 attests to.[15]

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). The pattern of results is always the same.

---

[15] Knack and Kropf (2001) are aware of the problem with the cross-sectional analysis, and report that the incorrect sign on turnout in their statistical analysis puzzles them.

5. Discussion

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, and 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. Lever machines are among the best technologies at the top of the ticket, but perform significantly less well further down. Conversely, DREs fall behind optical scanning in voting for the presidency, but make up for that deficiency further down the ballot.

The difference between the best technologies and the worst is about 2 percent of ballots cast. A margin of error that large must surely raise doubts about the outcome of many elections past, and to come.

Five of the last 20 presidential elections (the post war elections) have been determined by less than 2 percent of the vote. Roughly 1 in 10 statewide elections, such as for governor and attorney general, have been determined by less than 2 percent of the vote over the last twenty years. And, 1 in 20 U.S. Congressional elections have been determined by less than 2 percent of the vote since 1980.

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 of 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. They are 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 data show this two ways.

First, voting equipment produces inequities. Voting equipment clearly record votes with different degrees of reliability. 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. There is no relationship between the generations of technologies and their performance. 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) do 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 explained not by demographics or political factors or technology, but by "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. 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 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. Clearly, citizens in some counties, especially rural counties, regularly have higher residual vote rates than citizens in other counties.

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 principle 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
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	1404
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.

Figure 1. Comparison of voting technology performance coefficients for president and senator/governor.

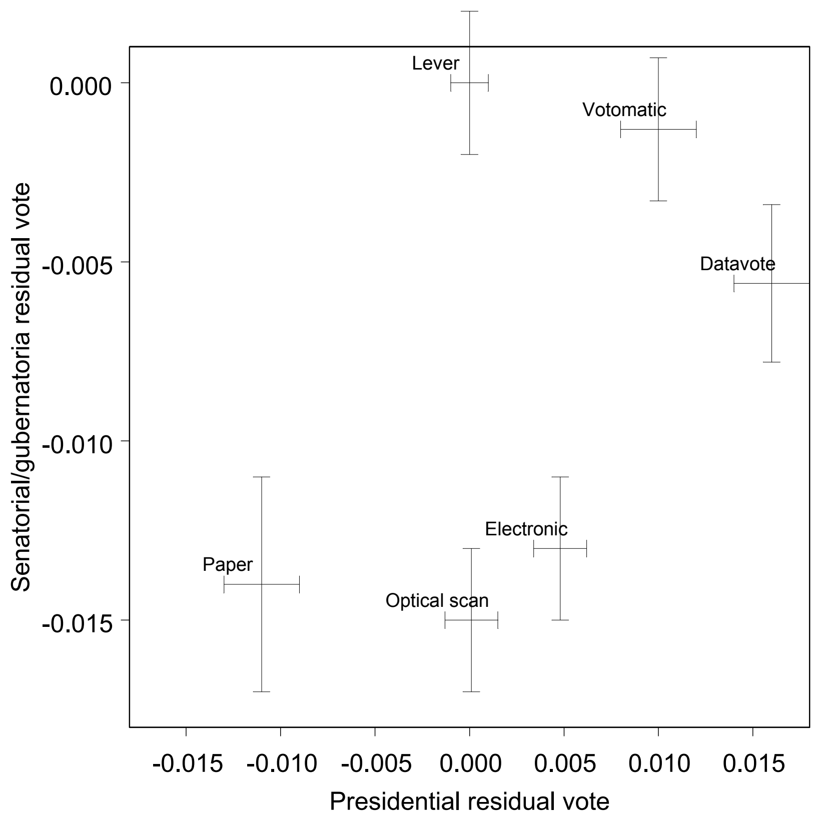

Figure 2. County-specific effects against turnout in county/town.

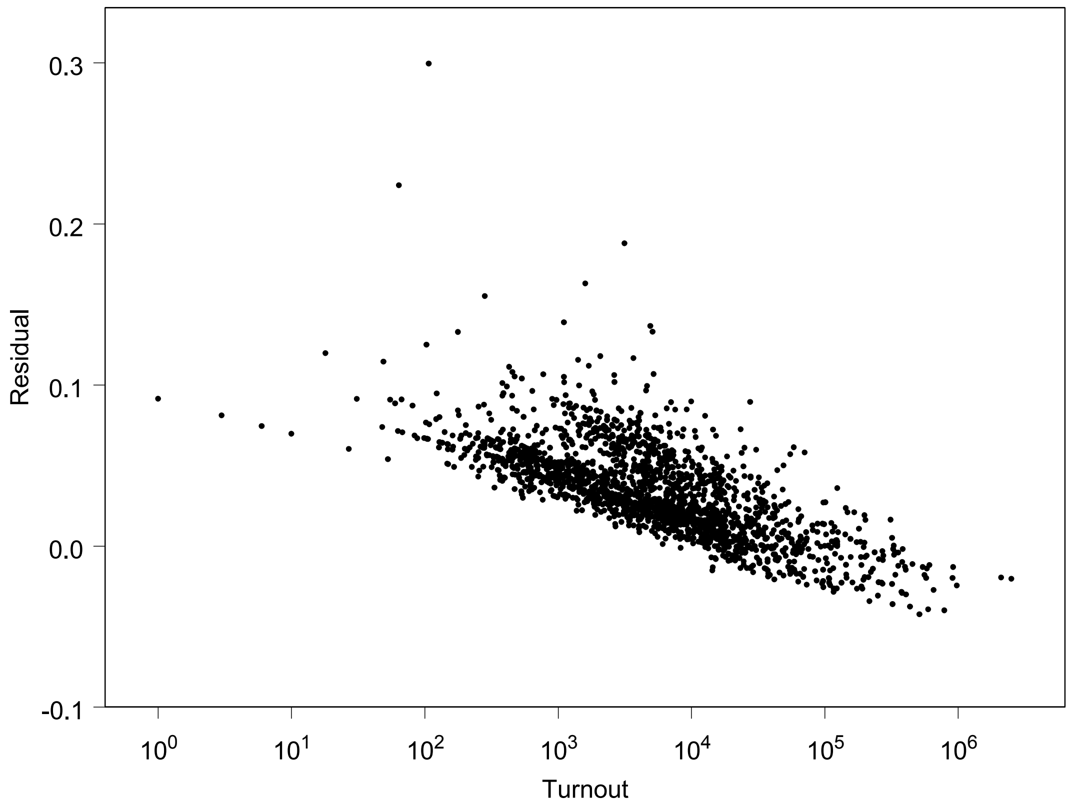

Table 1xx. State population using types of voting technologies, 2000.

	Voting Technology					
State	Punch	Lever	Paper	Scan	Electronic	Mixed
Alaska				76.6%		23.4%
Alabama		3.1%	0.2%	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%		0.5%	29.5%	24.3%	
Connecticut		100.0%				
D.C.	100.0%					
Delaware					100.0%	
Florida	64.5%	0.1%	0.1%	35.4%		
Georgia	43.8%	18.1%	0.1%	38.0%		
Hawaii				100.0%		
Iowa		10.5%	0.9%	82.1%	6.5%	
Idaho	58.7%		8.0%	33.3%		
Illinois	99.9%			0.1%		
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%	0.4%	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%	0.2%	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%	
Nevada	82.0%			18.0%		
New York		100.0%				
Ohio	74.4%	2.3%		12.6%	10.8%	
Oklahoma				100.0%		
Oregon	47.1%		0.4%	52.5%		
Pennsylvania	12.7%	62.6%	0.3%	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%	0.5%		
Virginia	20.4%	43.0%	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.

Table 2xx. Average population of jurisdictions using different voting machine types, 2000.

Equipment	Avg. pop.	N
Punch card	157,370	607
Lever machine	97,470	465
Paper	8,980	395
Optical can	60,992	1,267
Electronic	104,486	281
Mixed	154,124	138
Total	83,350	3,153

Table 3xx. 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%)

Table 4xx. Change in residual vote as counties change voting technologies after 1988. (Standard deviations in parentheses.)

Counties with lever machines in 1988….	Change in residual vote	N
Retained lever machines	-0.32% (1.16%)	510
Changed to optical scan	-0.61% (1.40%)	137
Changed to DREs	+0.22% (1.76%)	243
Counties with punch cards in 1988		
Retained punch cards	-0.42% (1.28%)	1,165
Changed to optical scan	-1.73% (1.85%)	322
Changed to DREs	-0.14% (2.82%)	12

Table 5xx. Residual vote multivariate analysis, presidential, gubernatorial, and senatorial elections, 1988-2000

	President		Governor & Senator	
	Without fixed effects	With fixed effects	Without fixed effects	With fixed effects
Equipment effects:				
Punch card	0.0077	0.0082	-0.021	-0.0030
	(0.0005)	(0.0015)	(0.001)	(0.0018)
Lever machine	Excluded	Excluded	Excluded	Excluded
Paper	-0.0012	-0.014	-0.022	-0.014
	(0.0014)	(0.002)	(0.002)	(0.003)
Optical scan	0.00071	-0.0045	-0.032	-0.014
	(0.00070)	(0.0014)	(0.001)	(0.002)
Electronic (DRE)	0.0080	0.0022	-0.0097	-0.012
	(0.0010)	(0.0015)	(0.0013)	(0.002)
Shift in tech.	0.00005	0.0010	-0.0021	-0.0004
	(0.00067)	(0.0007)	(0.0013)	(0.001)
Log(turnout)	-0.0004	0.0095	0.0005	0.031
	(0.0001)	(0.0026)	(0.0002)	(0.003)
Gov. or Sen. on ballot	-0.003	-0.0011	0.005	-0.004
	(0.001)	(0.0007)	(0.001)	(0.001)
Senator	—	—	0.009	0.008
			(0.001)	(0.001)
Percent Over 65	0.047	--	0.104	--
	(0.008)		(0.009)	
Percent 18 – 24	-0.012	--	0.027	--
	(0.009)		(0.010)	
Percent White	-0.030	--	-0.045	--
	(0.002)		(0.003)	
Percent Hispanic	0.011	--	0.005	--
	(0.004)		(0.005)	
Median Income (10,000s)	-0.002	--	-0.001	--
	(0.001)		(0.001)	
Constant	0.025	-0.11	0.027	-0.29
	(0.002)	(0.03)	(0.016)	(0.03)
N	8,982	8,982	11,625	11,625
R^2	.14	.79	.43	.74
Fixed effect: (not shown)	Year x State	Year x State County	Year x State	Year x State County
Number of categories	—	3,346	—	2,245
F test	—	F(3345,5572) = 2.971 (p < .0001)	—	F(2244,9318) = 3.705 (p < .0001)

Citations

Asher, Herb. 1982. "The Effect of Voting Systems on Voter Participation." Paper presented at the annual meeting of the Midwest Political Science Association, April 28–May 1, 1982.

Asher, Herb and Karen Crauder Snyder. 1990. "Effects of Voting Systems on Political Participation: Three Technologies Compared." Paper presented at the annual meeting of the Midwest Political Science Association, Chicago, Illinois, April 5–7, 1990.

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. 1996. "Election Roll-Off: A Test of Three Explanations." Urban Affairs Review 32:71–86.

Burnham, Walter Dean. 1965. "The Changing Shape of the American Political Universe." American Political Science Review 59:7–28.

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

Darcy, R. 1986. "Position Effects with Party Column Ballots." Western Political Quarterly 39:648–662.

Darcy, R. and Anne Schneider. 1989. "Confusing Ballots, Roll-Off, and the Black Vote." Western Political Quarterly 43:347–64.

Florida Governor's Select Task Force on Election Procedures, Standards, and Technology. 2001. "Revitalizing Democracy in Florida," March 21, 2001. url: www.collinscenter.org/info-url2660/info-url_list.htm?cat=FINAL%20REPORT accessed December 4, 2001.

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

Georgia Secretary of State. 2001. “The 2000 Election: A Wake-Up Call For Reform and Change.” Report to the Governor and Members of the General Assembly. url: www.sos.state.ga.us/elections/2000_election_report.pdf accessed December 4, 2001.

Herron, Michael C. and Jasjeet S. Sekhon. Forthcoming. “Overvoting and Representation: An Examination of Overvoted Presidential Ballots in Broward and Miami-Dade Counties.” Electoral Studies.

Iowa Secretary of State. 2001. “Iowa’s Election 2000: Facts, Findings, and Our Future.” A Report to the People of Iowa, Governor Vilsack, Lt. Governor Pederson, Members of the General Assembly, and Iowa’s Congressional Delegation. March 12, 2001. url: http://www.sos.state.ia.us/pdfs/ElectionReport.pdf accessed December 4, 2001

Knack, Stephen and Martha Kropf. 2001. “Invalidated Ballots in the 1996 Presidential Election: A County-Level Analysis.” Mimeo. University of Maryland.

Maryland Special Committee on Voting Systems and Election Procedures in Maryland. 2001. “Report and Recommendations.” url: http://www.sos.state.md.us/sos/admin/pdf/reportall1.pdf accessed December 4, 2001.

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.” May 2001. url: http://www.sos.state.mi.us/election/elecadmin/2001web/uniformvotinginmichigan.pdf accessed December 4, 2001.

Missouri Secretary of State. 2001. “Making Every Vote Count.” January 29, 2001. url: http://mosl.sos.state.mo.us/sos-elec/bluntcommission/mevc-report.pdf accessed December 4, 2001.

Nichols, Stephen M. and Gregory A. Strizek. 1995. “Electronic Voting Machines and Ballot Roll-Off.” American Politics Quarterly 23:300–318.

Roth, Susan King. 1994. “The Unconsidered Ballot: How Design Affects Voting Behavior.” Visible Language vol. 28, no. 1, pp. 48–67.

Roth, Susan King. 1998. “Disenfranchised by Design: Voting Systems and the Election Process.” Information Design Journal vol. 9, no. 1 (reprint; no pagination).

Rusk, Jerrold G. 1970. “The Effect of the Australian Ballot Reform on Split Ticket Voting: 1876–1908.” American Political Science Review 64:1220–38.

Saltman, Roy G. 1988. “Accuracy, Integrity, and Security in Computerized Vote-Tallying.” National Bureau of Standards Special Publications 500-158, August 1988.

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 45:521–37.

Thomas, Norman. 1968. “Voting Machines and Voter Participation in Four Michigan Constitutional Revision Referenda.” Western Political Quarterly 21:409–19.

Wand, Jonathan N., et al. Forthcoming. “The Butterfly Did It: The Aberrant Vote for Buchanan in Palm Beach County, Florida,” American Political Science Review.

Wattenberg, McAllister, and Salvanto. 2000. “How Voting is Like Taking an SAT Test: An Analysis of American Voter Rolloff.” American Politics Quarterly 28:234–250.

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