THE HELP AMERICA VOTE ACT IN THE 2006 GENERAL ELECTION: PROVISIONAL BALLOTS AND OVERVOTING

A Thesis
submitted to the Faculty of the
Graduate School of Arts and Sciences
of Georgetown University
in partial fulfillment of the requirements for the
degree of
Master of Public Policy
in Public Policy

By

Amy L. Cohen, A.B.

Washington, D.C.
April 8, 2011
The Help America Vote Act of 2002 (HAVA) seeks to improve election administration by providing the states with funds to upgrade their voting machines, creates the Election Assistance Commission (EAC), mandates requirements in election administration practices such as the creation and maintenance of computerized voter registration rolls and the development of clear provisional ballot standards, and amends and improves military and overseas voting laws. It is not immediately clear, though, how effective HAVA has been at achieving these goals. Using data from the EAC’s 2006 Election Administration and Voting Survey as well as data from the Democratic National Committee’s voter file, this study looks specifically at the provisional ballot rate and the number of overvotes per county in the 2006 general election. Tobit regression methods are used to examine how effective HAVA has been at improving election outcomes on these metrics for minority voters. In spite of significant data quality issues, this analysis finds that the strongest predictors of the provisional ballot rate are the change in proportion of registered voters age 65 or older, the number of polling places, the number of new registrants, and whether or not the state previously had a form of provisional balloting. The strongest predictors of the number of overvotes are the number of polling places and precincts, and use of DRE ballot machines; the African American turnout rate is also significant, though less strongly.
To DNC Targeting and, especially, to Andrew Claster:
Thank you for your wisdom and your data.

To my friends:
You are my family. Please accept my eternal gratitude.
# Table of Contents

Introduction...................................................................................................................... 1

Definition of Key Terms................................................................................................. 7

Literature Review............................................................................................................ 14

Description of Data....................................................................................................... 27

Variables......................................................................................................................... 30

Model............................................................................................................................... 42

Results............................................................................................................................. 43

Policy Recommendations............................................................................................... 49

References....................................................................................................................... 53
LIST OF TABLES

Table 1: Descriptive Statistics on County Characteristics .......................................................... 37

Table 2: Results of a Logistic Regression Analysis of Counties That Did Not Report Data on Provisional Votes or Overvotes .......................................................... 39

Table 3: Results of a Logistic Regression Analysis of Counties That Reported Zero Provisional Ballots or Overvotes ................................................................. 41

Table 4: Results of a Tobit Regression Analysis of Provisional Ballot Rate on Demographic Variables ................................................................................................. 45

Table 5: Results of a Tobit Regression Analysis of Overvotes on Demographic Variables ...................................................................................................................... 47
Introduction

Without a doubt, the 2000 presidential election is infamous because of what happened on Election Day and not because of the campaign itself. By midnight on November 7, a few states remained close and the outcome of the election depended on Florida and its 25 electoral votes. The news media initially called Florida for Vice President Al Gore, but when the popular vote tally showed Texas Governor George W. Bush with a commanding lead, they retracted their call, and in the early morning of November 8, some media outlets called Florida for Bush. The votes for the entire state, however, had not been counted, and the remaining counties – Broward, Miami-Dade, and Palm Beach – were Democratic strongholds. As the numbers from these counties rolled in, Bush’s strong margin dwindled to a few hundred votes and the media outlets who called the state in his favor retracted their statements. Florida state law mandates a recount when the margin of victory or defeat is within 0.5 percent,\(^1\) and another statute provides for hand recounts if requested by a campaign.\(^2\) The Gore campaign requested hand recounts in Broward, Miami-Dade, Palm Beach counties, and Volusia county – another Democratic leaning county, and “hanging chads” and “butterfly ballots” entered the national lexicon. Legal wrangling kept the recount going for more than a month, and on December 12, 36 days after voters went to the polls, the Supreme Court of the United States called an end to the recounts, and called the election for Bush.

Although the Constitution worked as it was supposed to in 2000, between “hanging chads,” “butterfly ballots,” and the chaos of the Florida recount, the election system itself did not. National calls for election reform came from people on both sides of the aisle, and the result

---


was the Help America Vote Act of 2002 (HAVA), bipartisan legislation purporting to change
election administration practices across the country. Specifically, HAVA: provides all 50 states
with funds to upgrade and update their voting machines to eliminate the use of punch card and
lever voting machines; creates the Election Assistance Commission (EAC), a resource for state
and local election officials that provides technical advice, serves as a “national
clearinghouse…for the compilation of information and review of procedures with respect to the
administration of Federal elections,” certifies the accuracy of voting machines, and provides
innovation in the field of election administration; mandates requirements in election
administration practices, such as the creation and maintenance of computerized voter registration
rolls and the development of clear provisional ballot standards; and, amends and improves
military and overseas voting laws, requiring the Secretary of Defense to make greater efforts to
get overseas active duty members of the military their ballots on time, and centralizing the
administration of overseas absentee balloting for civilians (Library of Congress, 2002).

HAVA fills something of an institutional void in American government, as prior to this
legislation, there was “no adequate federal/state partnership in elections” (Saltman, 2006, p. 1).In late 1971, in response to Election Day problems in Ohio, Detroit, San Francisco, Atlanta, and
Los Angeles, Representative William Keating (R-OH) introduced an amendment to the bill that
would eventually become the Federal Election Campaign Act of 1972 calling for the
development of a Clearinghouse on Election Administration that he said would provide
“information on good and bad ideas in voting systems…” (as cited in Saltman, 2006, p. 2-3).
The Clearinghouse, originally a division of the Office of Federal Elections of the General
Accounting Office (GAO), and later a part of the Federal Election Commission (FEC), was a
five person organization with a $225,000 budget (Saltman, 2006, p. 2). They decided to issue
voluntary equipment standards in 1977 but the standards – for punch card, optical scan, and DRE machines and finally released in 1990 – were, according to Saltman, “voluntary and there was no provision of federal funding who acceptance would guarantee conformance or, at least, would guarantee the right of the federal government to demand conformance.” (Saltman, 2006, p. 3). The Clearinghouse later was renamed the Office of Election Administration (OEA), and in 2001, they had not yet reissued ballot technology standards. Thus, HAVA’s passage in late 2002 was a huge step forward for the advancement of election administration.

In Ohio in 2004, though, election administration issues were front and center again, despite HAVA’s efforts at reform. Ohio’s Secretary of State, J. Kenneth Blackwell, also served as the co-chair of President Bush’s campaign in Ohio, and partisan election administration problems plagued the process from beginning to end, including instructions from Blackwell to invalidate voter registrations not “printed on white, uncoated paper of not less than 80 lb. text weight,” although this order was not followed consistently across the state and ultimately was rescinded (House Judiciary Committee Democratic Staff, 2005, p. 36). Furthermore, the Secretary of State’s office and the local boards of election did not provide enough voting machines or adequately train poll workers to accommodate the increased turnout that accompanies presidential elections and particularly one as hotly contested as 2004. The lack of preparation especially was evident in Democratic precincts in Republican-controlled jurisdictions such as Columbus, where many voters waited in lines of up to ten hours just to cast their ballot; the wait likely caused thousands of people to leave without casting a ballot across the state (Powell and Slevin, 2004). Finally, there were multiple issues with provisional ballots due to Secretary of State Blackwell’s narrow interpretation of the HAVA provisional ballot statute; Blackwell declared that a voter had to be in the correct precinct in order to be eligible to
cast a provisional ballot – HAVA requires that voters swear that they are in the correct jurisdiction to be eligible to cast a provisional ballot but does not specify a definition for jurisdiction, and the EAC recommends that voters be eligible to cast a provisional ballot if they are in the correct county. Blackwell’s narrow interpretation of the word jurisdiction led to an increase in the number of provisional ballots issued, as well as an increase in the number of provisional ballots invalidated for being in the wrong precinct. According to a report issued by the Democratic staff of the House Judiciary Committee, the increase in discarded provisional ballots occurred at a rate not proportional to the increase in provisional ballots issued (House Judiciary Committee Democratic Staff, 2005, p.79).³

Election administration issues such as those in Florida in 2000 and those that occurred in Ohio in 2004 typically have the greatest effects in low income and minority areas, populations whose interests are often overlooked to begin with. According to the 2000 Census, the counties that were central to 2000’s election debacle in Florida have large concentrations of vulnerable populations, as does the city of Columbus, Ohio, site of many issues in 2004’s presidential election. Broward County, Florida was 20.5 percent African-American and 16.7 percent Hispanic or Latino,⁴ and Miami-Dade County was 20.3 percent African-American and 57.3 percent Hispanic or Latino.⁵ Palm Beach and Volusia counties were less ethnically diverse, with 13.8 percent and 9.3 percent African American, respectively, but both had older populations,

---

³ The increase occurred both because of Secretary of State Blackwell’s narrow legal interpretation of the provisional ballot statute in HAVA and his failure to issue provisional ballot guidelines per a United States District Court ruling, although the ruling was later overturned by the Sixth Circuit Court of Appeals. According to the EAC, Blackwell was the only Secretary of State in 2004 to interpret the HAVA provisional ballot statute as he did (House Judiciary Committee Democratic Staff, 2005, p. 33).
⁴ United States Census Bureau “Broward County Census 2000 Factsheet”
⁵ United States Census Bureau “Miami-Dade County Census 2000 Factsheet”
with more than 20 percent of the population over the age of 65 in both places.\textsuperscript{6} Columbus, Ohio was 24.5 percent African American in the 2000 Census.\textsuperscript{7}

Reform of the American election system is overdue, but for the most part, the media only focuses on the story when the outcome is very close or when things go very wrong, such as the presidential elections in 2000 in Florida or 2004 in Ohio, and most Americans think of these elections as outliers. But the reality is that there were also problems in primaries and general elections in several states in 2006 and in Ohio in 2007, and lest one think that this is purely a 21\textsuperscript{st} century problem, reports from the 1971 general election show voting problems in Detroit, San Francisco, Atlanta, and Los Angeles, as well as in Ohio in 1978 (Gerken, 2009; Saltman, 2006; Asher, Schussler, and Rosenfield, 1982); problems with their election machines in the 1968 election led IBM to leave the voting machine industry altogether (Caltech/MIT, 2001a, p. 7).

Given the long history of election administration problems, it is not difficult to conclude that almost every election administered in the United States is likely flawed from the start, although not every election is close enough to warrant a second look at election administration processes.

Election reform, though, is difficult, because of the partisan nature of the election process from the bottom (elected local election officials) to the top (elected state and federal legislators), and because admitting that the system is broken undermines Americans’ confidence in election results and casts doubt on the validity of our government institutions. The 2005 report issued by the Carter-Baker Commission on Federal Election Reform stated that, “[P]ublic confidence in the electoral system is critical for our nation’s democracy. Little can undermine democracy more than a widespread belief among the people that elections are neither fair nor legitimate”

\textsuperscript{6} United States Census Bureau “Palm Beach County Census 2000 Factsheet;” United States Census Bureau “Volusia County Census 2000 Factsheet”
\textsuperscript{7} United States Census Bureau “Columbus City, Ohio Census 2000 Factsheet”
(“Building Confidence,” 2005, p. 1). Thus, reforming the election administration process while retaining trust in the system is a difficult, yet critical undertaking, and HAVA has admirable goals.

In practice, however, HAVA’s impact has been blunted by its overly broad mandate, lack of funding, and the difficulty of making meaningful changes in an entrenched system without undermining confidence in the outcomes. HAVA tries to improve election outcomes for all voting citizens by improving voter registration methods, developing clearer provisional ballot standards, and replacing old lever and punch card ballot machines, but measuring the effectiveness of these improvements is difficult. Looking specifically at the provisional ballot rate and the number of overvotes cast, and particularly for African American and Hispanic voters, voters under 25 years old and over 65 years old, as well as the low income, this study examines whether the passage of HAVA improved election performance.
Definition of Key Terms

Because election administration uses terminology that is often unfamiliar to people outside of the field, it is useful to have a clear understanding at the outset of what terms mean to avoid confusion.

A provisional ballot is “a ballot issued to a voter at the polling place when their eligibility to vote has not been determined” (United States Election Assistance Commission, 2007, p. 83). Voters who are not on the registration rolls at a given precinct but believe that they should be, or whose information on the printed registration list is incorrect, are legally entitled by HAVA to cast provisional ballots. Provisional ballots typically are counted several days after the election, when election administrators have time to determine the individual voter’s eligibility. Because the text of the HAVA legislation requires that voters swear they are in the correct jurisdiction, but does not define jurisdiction, different states have different provisional ballot requirements. In 2006, 28 states and the District of Columbia only counted provisional ballots if they were cast in the correct precinct, and 15 states only counted provisional ballots if they were cast in the correct state or county (Electiononline.org, 2006a, p. 32).

Before HAVA, most states had some kind of recourse for voters who arrived at their polling place and found that they were not on the registration rolls or were unable to provide the required identification: provisional balloting, same-day registration, affidavit balloting, and fail-safe voting.

Same-day registration renders provisional balloting unnecessary. A voter who arrives at his polling place and finds he is missing from the list may simply re-register on the spot. Idaho, Maine, Minnesota, New Hampshire, Wisconsin, and Wyoming used same-day registration before HAVA, although Wisconsin and Wyoming issued provisional ballots to those voters trying to
register at their polling place but who were unable to provide the identification required,\textsuperscript{8} and Maine issued challenge ballots to voters whose eligibility to vote in that precinct is in question ("Maine Revised Statutes;" Electiononline.org and The Constitution Project, 2001, p. 3).\textsuperscript{9}

Affidavit ballots were given to voters not appearing on the registration rolls, and required the voter to swear a legal affidavit, under penalty of perjury, that he was eligible to vote. Generally, election administrators were not required to go back and check the voter’s eligibility before counting the ballot, though in some states, affidavit ballots were checked regularly for validity. Other states specially marked these ballots so that in the event of a recount, the eligibility of affidavit ballot voters could be confirmed. States that used affidavit balloting are: Alabama, Kentucky, Michigan, Mississippi, and Texas (Electiononline.org and The Constitution Project, 2001, p. 3, 5). Several states currently use the terms “affidavit ballot” and “provisional ballot” interchangeably.

Fail-safe voting was mandated by the National Voter Registration Act (NVRA), often referred to as “Motor Voter” because it allows voters to register at their state’s Department of Motor Vehicles. Voters who moved within the same jurisdiction but did not re-register were permitted to vote after an oral or written affirmation of their new address. Nebraska, New Jersey, and Ohio used fail-safe voting (Electiononline.org and The Constitution Project, 2001, p. 3, 5).

\textsuperscript{8} These provisional ballots were only counted if cast in the correct precinct.
\textsuperscript{9} Per the Bureau of Corporation, Elections, and Commissions in the Maine Secretary of State’s office, challenge ballots are issued when a voter’s eligibility to vote in that precinct is in doubt, such as if someone questions whether the voter lives in that town. The challenged voter is given a special ballot with an identification number known only to that voter and the warden of that town. The ballot is cast and counted just like a regular ballot, and is only reexamined in the event of a recount. Maine has issued challenge ballots since the mid-1980s.
Once a voter is given a ballot, there are several things that can occur: a valid vote, an overvote, or an undervote. A valid vote is one that is cast by a voter and counted by election administrators.

Overvoting is defined as a ballot where “…a voter makes more than the permitted number of selections in a single race/contest or when a voter makes a selection in a race/contest on which he/she was not eligible to vote” (United States Election Assistance Commission, 2007, p. 89). Thus, a voter overvotes when he votes for two candidates for President, and while the voter cast a ballot, this ballot will not be counted towards either candidate’s vote total. If a voter marks the box or bubble for a candidate or slate of candidates and writes-in the same name or names in the area reserved for write-in candidates, this often is considered an overvote because the voter has marked more than the allowed number of selections (Caltech/MIT, 2001a, p. 20). Overvoting may be a result of poor ballot design or low education levels.

Undervoting is defined as a ballot where “a voter makes less than the allowed number of selections in a single race/contest or when a voter votes on less than all of the races/contests for which he/she is eligible to vote” (United States Election Assistance Commission, 2007, p. 89). A voter undervotes, for example, when she goes to her polling place and does not cast a ballot for any candidate for President. This ballot is not counted in any presidential candidate’s vote total. If that voter, however, did mark a candidate in the Senate race, her vote counts towards her Senate candidate of choice. It is difficult to determine the cause behind undervoting because many people deliberately undervote to show their displeasure with the choice of candidates, such as the voter who does not mark a candidate for President but does mark a candidate for Senate, while others may undervote by accident because of poor ballot design or low levels of education.
Residual votes are the combined total of overvotes, undervotes, and any ballot that is cast but not counted for any reason (Caltech/MIT, 2001a, p. 20). Often, the total number of residual votes is divided by the total number of ballots cast for a residual vote rate.

There are five types of voting technology: paper balloting, punch card balloting, lever balloting, optical scan balloting, and DRE balloting.

Paper balloting is the most basic of all methods of voting. The voter marks the box of his candidate of choice on a piece of paper and puts the ballot in a box, where it is counted manually after voting has closed. Paper balloting was ubiquitous in the United States in the 19th century and remains in use today (Caltech/MIT, 2001b, p. 2). Overvoting and undervoting are possible with this technology.

Punch card machines automate the process of counting traditional paper ballots, and became popular in the 1960s (Caltech/MIT, 2001a, p. 18). There are two primary types of punch card balloting machines – DataVote and Votomatic. In both methods, voters are given a piece of heavy cardstock with perforated rectangles, called chads. In DataVote systems, the names of the candidates and ballot initiatives are listed on the cardstock next to a chad. Voters place the card into a hollow ballot holder and use a lever to punch the chads next to the candidates or initiatives that they wish to select so that the chads fall into the ballot holder (Smithsonian, 2004). After voting, the voter places the card in a box, and at the end of the day, the cards are mechanically sorted by a machine that counts the number of perforations next to each candidate’s name and records the information on internal software. In some cases, the voter may place her ballot into the sorter herself. In Votomatic systems, the cardstock does not have any candidate names written on it. Voters go into a booth and place the card into a ballot holder that is attached to a booklet that lists the candidates for each race. Each listed candidate in the booklet lines up with
a chad, and the voter uses a lever to punch the chad next to her chosen candidate, so that the chad falls into the ballot holder. To vote for the next office, the voter turns the page in the booklet and repeats the process. After voting, the procedure is the same as the DataVote system: the voter puts the card in a box, and the ballots are counted using a card sorter at the end of the day, or the voter puts the card in the sorter herself (Caltech/MIT, 2001b, p. 3). Overvoting and undervoting are possible with both DataVote and Votomatic punch card balloting. HAVA provides funding for states using punch card machines to switch to other technologies, but in 2006 two states – Idaho and New York – still used punch card technology in some of their counties.

Lever ballot machines require voters to step into a booth and flip the switches of the preferred candidates or ballot options; the ballot is not officially cast until the voter pulls a large lever that registers the votes on a counter on the back of the machine. While undervoting is possible, it is not possible to overvote on a lever machine because the machines are designed not to accept more than the appropriate number of candidates for a given office. The final vote tallies are taken from the back of each individual machine and added together to get a precinct total, which is then added to other precinct totals to get the total number of ballots cast for each candidate. There is no record of each individual vote. Originally introduced in New York in the late 19th century, lever ballot machines were used in all major metropolitan areas by 1930, although by the beginning of the 21st century, these machines were primarily used in New York and Connecticut (Caltech/MIT, 2001b, p. 2; Electiononline.org, 2003, p. 5). HAVA gives states using lever machines the money to upgrade to other technologies because the machines do not keep a record of each individual vote, but in 2006, five states – Connecticut, Delaware, Iowa, New York, and Missouri – were still using lever balloting, though primarily Connecticut, Delaware, and New York.
Optical scan balloting, which became popular in the 1990s, uses methodology familiar to any standardized test taker (Caltech/MIT, 2001a, p. 18). The voter is given a sheet with the names of candidates and brief descriptions of ballot initiatives, and next to each candidate or initiative is a bubble or an incomplete arrow. To mark her selection, the voter fills in the bubble or completes the arrow next to her preferred candidate. Depending on the technology in use at that particular precinct, the voter may be able scan her ballot on the spot to make sure that the vote is registering correctly before placing her ballot in a secure box. Often, though, the voter places her ballot in the box and precinct administrators scan all of the ballots after voting is closed, either at the polling place or at one central location, to count the totals for each candidate or initiative (Caltech/MIT, 2001b, p. 3). Overvoting and undervoting are possible with this technology.

Direct record electronic devices, or DREs, essentially are electronic lever machines and use technology similar to ATMs, making them familiar technology for most voters. Although the process differs slightly by manufacturer, generally, on touch screen DRE machines, the voter goes into the booth and touches the name on the screen of the candidate she wishes to vote for. The display can either be paginated, where each office on the ballot gets its own page, or unpaginated, whereby all the offices are listed on one page. When the voter is finished, she may review all of her selections before pressing a button to officially cast her ballot. Each machine tallies the number of votes for each candidate and at the end of the day, the totals for each machine are sent, usually on a disk, to a central location where they are combined with other machine totals for a total vote tally. Push button DREs operate similarly, although instead of touching the screen, the voter pushes a button next to the name of the candidate she wishes to select. Like touch screen machines, the list may be paginated or unpaginated, and the voter is
given the opportunity to check over her ballot before officially casting her vote. DRE machines are programmed not to accept overvotes, and because the voter is asked to review her ballot before it is cast, accidental undervoting is less frequent on DRE machines (Caltech/MIT, 2001b, p. 3). HAVA mandates a paper record of each vote cast, known as a voter verifiable paper audit trail, or VVPAT, be created for use in the event of a recount, although some state laws require that the paper record be produced for review by each individual voter (Tokaji, 2006, p. 6).
Literature Review

Although HAVA was passed in 2002 and initially implemented in 2003, the current literature on the subject primarily focuses on the merits of different kinds of voting machines. A 2006 article, however, examines the impact of provisional balloting by examining the results of the EAC’s 2004 Election Administration and Voting Survey. According to the EAC’s report, in 2004, 1.9 million voters cast provisional ballots, and 64.5 percent of those ballots were counted, a huge step towards providing the voters with the safety net intended by the authors of HAVA. However, as Wendy Weiser, author of “Are HAVA’s Provisional Ballots Working?” points out, that means that 35.5 percent, or more than half a million votes, were rejected, and this number does not take into account the large numbers of people who were denied access to provisional ballots in the first place – provisional ballot access problems were among the top five complaints to the Election Protection Coalition’s voter hotline (Weiser, 2006, p. 2-3). Because “states did not plan for provisional ballots until shortly before the election,” provisional ballots were not available, misinformed pollworkers did not offer provisional ballots when appropriate, and, in several instances, pollworkers denied provisional ballots to those voters who requested them (Weiser, 2006, p. 3). It is easy to blame these missteps on individual pollworkers, but in most cases, the fault ultimately lies with the Secretaries of State, who did not issue standards for casting and counting provisional ballots until weeks – or in some cases, days – before the election; election administrators are also at fault for being slow to respond to voter questions about their polling location. As an experiment, a writer from The New York Post emailed the New York City Board of Elections on November 1, 2004 to confirm polling place information and heard back five months later, on March 23, 2005 (Gaskell, 2005).
Weiser’s article clearly elucidates the pros and cons of provisional balloting. On the one hand, it seems unfair to disenfranchise a registered voter for an honest mistake, especially because in many cases, voter confusion is a result of sloppy election administration. At the same time, however, provisional balloting raises valid concerns. Many election administrators assume that voters will want to cast their ballot at the most convenient precinct, even if it is not the precinct they are assigned to; if provisional ballots are counted when they are cast in the correct county, regardless of precinct, there would be no way for election administrators to allocate resources appropriately, since there would be no way to predict where voters would appear. Further, provisional ballots cast in the wrong precinct only count for top of the ballot and jurisdiction-wide races, and, as a result, lower-tier candidates who are not running in jurisdiction-wide races would lose voters (Weiser, 2006, p. 7). Weiser points out, though, that there are legal means of ensuring that voters do not deliberately go to the wrong precinct, including making it a punishable offense for voters to purposely vote in the wrong precinct, or asking voters to swear an affidavit that they believe they are in the correct precinct.

Most of the literature, however, focuses on voter registration, and it is not unrelated to the matter of provisional balloting because of the relationship between improper registrations and provisional ballots.

The ultimate goal of voter registration, according to Eric Fisher and Kevin Coleman, is “…to ensure that only those people entitled to vote in a given jurisdiction can do so, and that they each vote only once” (Fisher and Coleman, 2006, p. 1). The fundamental question in voter registration, though, is whether it is more important for all eligible voters to be able to vote, or whether it is more important to prohibit ineligible voters from voting; unfortunately, HAVA does not provide a clear answer that question. Registration has a history of being used by both the
parties and the states to limit participation according to the whims of the majority. Historically, the Democratic Party in many Southern states implemented a “white primary” that excluded African Americans from being members of the Democratic party, and the states also made liberal use of poll taxes, literacy tests, and the grandfather clause, which restricted voting with exceptions for residents whose ancestors were eligible to vote prior to the Civil War. The state-to-state differences in registration laws reflect the federalist system of government in the United States, and election administration is an area typically reserved for the states. Most have a patchwork system of voter registration rules and regulations, and HAVA gave all of the states an opportunity to develop “new generations of registration systems” that function efficiently (Fisher and Coleman, 2006, p. 2). This is a much larger and more onerous task than many people understand.

Much of the focus from HAVA is on developing a system of voting that is easy and secure, but it is important to remember that people vote, on average, one time per year while registering to vote is a task that happens even less frequently: “Registration happens much less often than voting on average, forms and rules vary substantially among jurisdictions, there is often very little feedback, and a mistaken rejection may mean disenfranchisement until the next election” (Fisher and Coleman, 2006, p. 4). Between 18 year olds aging into our electoral system, deaths, and the highly mobile society that we live in today, there are approximately 30 to 35 million potential registration record changes each year, leaving a lot of room for error (Fisher and Coleman, 2006, p. 3).

Database errors are not uncommon in any system, let alone one as large as a statewide voter registration database. To minimize errors, states run their registration lists against other state databases, such as a list from the Department of Motor Vehicles, to ensure that information
matches up; when mistakes are found, the record is examined, and often, the registration is
discarded, resulting in thousands of so-called administrative disenfranchisements each year
(Fisher and Coleman, 2006, p. 5). The EAC recommends that states not require a perfect match
with other state databases, but, like all EAC recommendations, this is just a guideline, and many
states require very close, although not quite exact, matches. Registrants are not always informed
that their application has been rejected, particularly if the information that does not match
between lists is their address, and upon going to their polling places and finding themselves not
registered, are required to cast provisional ballots, or, in states with same day registration, re-
register in order to cast their vote. Although both of these methods are somewhat effective in
blunting the impact of administrative disenfranchisement, Fisher and Coleman rightly conclude
that national consensus standards for state voter registration systems would improve election
administration overall without the federal government imposing on the rights of states to run
elections as they see fit.

Although the difficulties with voter registration databases do raise some legitimate
concerns with HAVA, the legislation has not been completely ineffective in changing the way
elections are administered. January 1, 2006 was the deadline for many of the key requirements
of the HAVA legislation, and Doug Chapin, Director of the Electiononline.org initiative for the
Pew Center on the States, notes that “…between 2000 and 2006, jurisdictions containing over 80
million registered voters saw … new voting technology...” (Chapin, 2006, p. 2). This number
represents almost two-thirds of all of the people who voted in the 2004 presidential election and
approximately 75 percent of all of the people who voted in the 2000 presidential election.
Similarly, in 2000, only 11 states had what we now know as provisional balloting, and about
two-thirds of the states had something resembling provisional balloting; now, HAVA mandates
provisional balloting in every state (Chapin, 2006, p. 2). Thus, election administration has made
some important steps forwards, but the very partisan nature of the current electoral process often
results in extremely close elections, and the large number of changes that have occurred in a
short period of time “…breeds uncertainty and uncertainty almost always is fertile ground for
error” (Chapin, 2006, p. 6).

Much of the attention after the 2000 election focused on different types of voting
machines in use across the United States, and Title I of HAVA allocates $325 million to the
states for the replacement of punch-card and lever voting machines (Tokaji, 2006, p. 3).
According to the text of the legislation, the original deadline for states to replace these types of
machines was November 2004, but many states were given extensions; the extended deadline
was 2006, in time for each state’s primary election, and any state that failed to meet either the
first deadline where applicable, or the second deadline, was required to return the federal funds
they took “in proportion to the number of precincts that…failed to get rid of the old equipment”
(Tokaji, 2006, p. 3). In spite of the financial incentive, many states were still not compliant in
the 2006 general election because the increased usage of DRE machines led to growing concerns
about fraud and the security of electronic voting machines.

The focus on voting machines in the wake of HAVA, though, exposed publicly the fact
that certain machines have higher rates of overvoting and undervoting than others, a fact known
to researchers since the 1970s. In their 2004 article “Punch Card Technology and the Racial Gap
in Residual Votes,” Justin Buchler, Matthew Jarvis, and John E. McNulty find that minorities,
and especially African Americans, have higher rates of residual voting, and thus higher rates of
invalidated ballots, particularly when using Votomatic punch card machines. Higher residual
vote rates, though, are a result of more than just race. Lower education levels are associated with
higher residual vote rates, and compounding this effect, the average education level of minority voters tends to be less than the average education of white voters, increasing the likelihood that minority voters will cast ballots that are ultimately declared invalid. Even when the authors control for education, though, they find that minorities still cast more residual ballots than non-minorities. Furthermore, in 2000, all punch card machine counties in their study had higher residual vote rates than non-punch card counties, and Buchler, Jarvis, and McNulty note that the rate at which minority ballots are invalidated by punch card machines is higher than it is for other voting methodologies (Buchler, Jarvis, and McNulty, 2004, p. 518).

Buchler, Jarvis, and McNulty’s model looks at the relationship between county and precinct-level residual vote rates and county-level demographics, and compares residual vote rates in counties that switched all of their precincts from punch card ballot systems to optical scan systems (Buchler, Jarvis, and McNulty, 2004, p. 517-518). Because good election data generally, and precinct-level data especially, are difficult to come by, Buchler, Jarvis, and McNulty use data from California and Illinois for their analysis. In Fresno County, California, which switched from Votomatic punch card ballots to optical scan ballots in time for the 2000 presidential election, “…every single tract had smaller residual vote rates, regardless of racial composition” after the switch, and any disparity in the residual vote rates of whites and African Americans all but disappeared. The results are the same for Marin County, California as well (Buchler, Jarvis, and McNulty, 2004, p. 518-519).

Another striking example of the benefits of switching from a Votomatic system to an optical scan system comes from Illinois. In 1996, Madison County, Illinois used Votomatic machines and the undervote rate was higher in precincts with higher minority populations; however, in 2000, when the county used optical scan ballots, “…there was no consistent
relationship between the percent of a precinct that was minority and the undervote rate,” the undervote rate for all voters decreased, and minority voters were as likely as white voters to undervote (Buchler, Jarvis, and McNulty, 2004, p. 520). The authors observed the same result in other counties in their study that changed from the Votomatic system to optical scan balloting, as well, and in both years, no clear pattern emerged with respect to overvoting.

Using the Fresno data, Buchler, Jarvis, and McNulty calculate the probability of casting a residual ballot based on the percentage minority in a census tract. In a Votomatic punch card tract with 0 percent minority voters, the residual vote rate will be approximately 2.5 percent; in a 100 percent minority punch card tract, the residual vote rate will be about 5 percent. In contrast, in an optical scan census tract with 0 percent minority voters will have a residual vote rate of about 0.5 percent; in a 100 percent minority optical scan tract, the residual vote rate will be about 1 percent (Buchler, Jarvis, and McNulty, 2004, p. 521). Assuming that the residual vote percentages are estimates of the percent of votes, dividing the percentage of valid white votes by the percentage of valid minority votes provides ratios equal to the weight of the white vote. For Votomatic machines, this works out to needing 1,026 minority votes to produce the same number of valid votes as 1,000 white voters: 97.5/95 = 1.026; in other words, the Votomatic system gives white votes 1.026 times the weight of a minority vote. For optical scan machines, this works out to needing 1,005 minority votes to produce the same number of valid votes as 1,000 white voters: 99.5/99 = 1.005. (Buchler, Jarvis, and McNulty, 2004, p. 521). Putting this in perspective, the Supreme Court rejected a state’s redistricting plan that gave one district 0.6984 percent more voters than another in the same state on the grounds that it violated the “one person, one vote” decision in *Gray vs. Sanders* (1963) (Buchler, Jarvis, and McNulty 522).
Furthermore, the authors find a strong correlation between educational attainment and the percent of ballots not counted on Votomatic machines. Because there is also a substantial racial gap in educational attainment, it is more likely that minority votes will be invalidated on Votomatic machines, which the authors suggest is akin to a modern day literacy test. In 1970, literacy tests were declared unconstitutional by the Supreme Court, which argued that due to a history of unequal educational opportunities, literacy tests were racially unfair. Ten years later, the Court held that a policy must demonstrate the intent to discriminate to be considered discriminatory, but in the 1982 reauthorization of the Voting Rights Act, Congress declared that a policy could be considered discriminatory absent intent. Votomatic voting machines do not intend to discriminate against minority voters, but because their design does discriminate against these voters, the effect is similar to that of a literacy test (Buchler, Jarvis, and McNulty, 2004, p. 522-523).

The authors, however, make no mention of the fact that one analysis of polls and exit polling that predates their study finds that intentional undervoting is more common among African Americans than among whites (Tomz and Van Houweling, 2003, p. 57). This analysis, by Michael Tomz and Robert P. Van Houweling, of American National Election Survey (ANES) and Voter News Service (VNS) polls, found a statistically significant difference in the rates of reported undervoting between African Americans and whites, suggesting that perhaps the results found by Buchler, Jarvis, and McNulty are a result of deliberate action, or rather, inaction, on the part of the African American community (Tomz and Van Houweling, 2003, p. 57-58).\textsuperscript{10}

While Buchler, Jarvis, and McNulty look exclusively at the racial gap between Votomatic punch card ballots and optical scan ballots, Tomz and Van Houweling’s 2003 article, “How Does

\textsuperscript{10}In spite of the racial gap the authors find among those who intentionally undervote, it is important to keep in mind that intentional undervoting is still a rare occurrence.
Voting Equipment Affect the Racial Gap in Voided Ballots?” looks at the racial gap on the four primary types of machines in use across the United States – punch cards, optical scan, lever, and DREs. They hypothesize that the racial gap among residual votes should be smaller with lever and DRE machines than with punch cards and optical scan ballots because these systems do not allow overvoting, and DRE machines make it harder to accidentally undervote because the software reminds the voter if he has not cast a ballot for a particular office (Tomz and Van Houweling, 2003, p. 48-49). Punch cards and optical scan ballots are more prone to human error – on punch cards, if the hole is not punched completely and the chad is hanging, the ballot may not be counted, and optical scan ballots can be invalidated if the bubble is not completely filled in, the arrow is not completed all the way, or the marks are made using the wrong kind of pen or pencil. In the 2000 election, only 1.7 percent of presidential ballots cast on lever or DRE machines were not counted, an average better than punch cards and slightly worse than optical scan ballots (Tomz and Van Houweling, 2003, p. 48).

Tomz and Van Houweling use individual-level data from South Carolina and Louisiana, states which also record the race of voters, aggregated to the precinct-level to determine the exact numbers of voters who used each type of voting machine. Having the race of those who voted allows the authors to use a race variable that reflects the actual turnout percentage by race at the precinct-level, eliminating some of the measurement error that plagues similar studies that use proxy variables to get a measure of the African American turnout, such as percentage African American for the whole county or the percentage of the registered voter population that is African American. Using these proxies introduces “nonrandom measurement error that depresses estimates of African American and white invalidation” (Tomz and Van Houweling, 2003, p. 49-50).
The authors were also able to work with the state and local Boards of Election to get accurate measures of the number of invalid absentee ballots at the precinct level. Typically, absentee ballots are aggregated and considered their own precinct, unrelated to the actual precinct in which the voter is registered, so there is no real means of determining whether absentee ballots from one precinct had a higher invalidation rate than those from another. This is particularly significant because absentee ballots have high rates of invalidation; in Louisiana in 2000, for example, absentee ballots represented 4 percent of voter turnout, but 31 percent of voided presidential ballots. Further, if whites are more likely to vote absentee than African Americans, the official statistics will lead researchers to overestimate the ballot invalidation rate for white voters because of the high rates of invalidation among absentee voters (Tomz and Van Houweling, 2003, p. 50). Like using the actual African American turnout percentages, the additional step of assigning absentee ballots to their appropriate precinct gives their results a higher degree of statistical accuracy than many other studies. Additional socio-economic variables in their model, such as measures of the poverty rate, education, and income, are measured at the county-level.

The authors run five separate models – one for each machine type used in the 2000 presidential election in each state\(^\text{11}\) – and the dependent variable in each case is the proportion of invalid ballots out of the total number of ballots, leaving out extreme cases, which they classify as the top and bottom 1 percent (Tomz and Van Houweling, 2003, p. 50). They find “that DRE and lever machines nearly eliminate the difference between black and white invalidation rates,” (Tomz and Van Houweling, 2003, p. 59) and that the \(R^2\) values are much larger for the punch card and optical scan ballot models than for the DRE and lever models. The authors suggest that difference in \(R^2\) values means that race explains a larger portion of the variation in counties that

\(^{11}\) In South Carolina: punch card, optical scan, and DRE; in Louisiana: lever and DRE.
use punch card and optical scan balloting than counties that use DRE and lever balloting (Tomz and Van Houweling, 2003, p. 53).

Tomz and Van Houweling run an additional model to see if it matters where the votes are counted, that is, whether counting the ballots at a central location or counting the ballots at the polling place has an impact on the invalidation rate. This is a particularly significant question because polling places that have the ability to count individual ballots often allow voters to check their ballots before they leave to make sure that they marked their ballot appropriately, perhaps decreasing the invalidated ballot rate. Unfortunately, the authors are unable to draw any clear conclusions about whether it matters where the ballots are counted (Tomz and Van Houweling, 2003, p. 56).

The Tomz and Van Houweling study paints a positive picture of DRE ballot machines, but other studies are somewhat less optimistic about DRE machines. The Caltech/MIT Voting Technology Project released a paper in March 2001 entitled “Residual Votes Attributable to Technology: An Assessment of the Reliability of Existing Voting Technologies.” They look at incidences of overvotes and undervotes by looking at election returns and voting methods in four presidential elections – 1988, 1992, 1996, 2000 – and find that “manually counted paper ballots have the lowest average incidence of spoiled, uncounted, and unmarked ballots, followed closely by lever machines and optically scanned ballots” (Caltech/MIT, 2001b, p. 2). The authors also compare the average residual vote rates of all five types of voting systems; paper ballots, lever machines, and optical scan ballots have the lowest average residual vote rates, and there is a statistically significant difference in the average residual vote rates of these methods and the average residual vote rates of punch card machines and DREs (Caltech/MIT, 2001b, p. 9).
Specifically, this study looks at the change in the residual ballot rate in counties that changed voting method between any of the four presidential races of interest, which works out to approximately 50 percent of all counties (Caltech/MIT, 2001b, p. 12). Additionally, the authors remove any county using more than one system from their dataset because the study uses a natural experiment to examine what happens when a county changes technology, and in mixed counties it is unclear which system is being used. The models are run on the full dataset, as well as on a subset that excludes extreme values in the dependent variable, residual vote rate, which is defined as counties in the top and bottom 5 percent. The authors do not provide their exact methodology, but, using lever machines as a baseline, find that changing to a paper ballot system decreases the residual vote rate, although this relationship disappears when extreme values are removed from the dataset. Changing from a lever machine to a punch card system – either the Votomatic system or the DataVote system – results in a statistically significant increase in residual vote rate for both punch card systems; this relationship remains even when the model is run on the subset that excludes extreme values. Interestingly, the authors find that switching from a lever machine to a DRE results in a statistically significant increase in residual vote rate, even when extreme values are excluded. Changing from a lever system to an optical scan system has no statistically significant effect on the residual vote rate (Caltech/MIT, 2001b, p. 14).

Because the design of DREs makes it impossible to overvote, the increases in residual vote rate the authors find are driven by undervotes (Caltech/MIT, 2001b, p. 14). They suggest that their findings could be a result of any number of problems, such as maintenance issues or a voter learning curve. DREs can be confusing to the voters, difficult to set up, or poorly

---

12 Note that they include Massachusetts and Vermont, where voting technology is uniform within each town; for these states, the authors use town level voting methodology information. For more information, see Caltech/MIT, 2001b, p. 12.
designed, and there are security concerns with using a computerized system for voting. The machines need to be tested regularly to ensure that the software has not been hacked or corrupted, and most states and localities do not have the resources to run the machines “…through all the possible situations that [it] might be faced with” (Selker, 2004, p. 94).

Clark Miller, however, notes that there is more to the administration of elections than just machines: “…the objectivity of electoral outcomes derives not only from the mechanics of counting but also from the social and political processes that help to assure that machines work as they should, both technically and politically,” and that a valid vote is more than just marking a ballot (Miller, 2001, p. 455). We have a legal system that allows for challenges and recounts, and while our elections are frequently messy, they usually sort themselves out. In recounts, state election administration laws typically direct officials to try to divine the intention of the voter before discarding the vote, because disenfranchisement is a severe punishment for not following the exact letter of the law. Generally, the courts have supported this tactic, and try only to get involved when voting discrepancies would affect the outcome of the election; as Miller notes, “[T]he American electoral system tallies votes fairly well, when it is allowed to function” (Miller, 2001, p. 457).
Description of Data

The data in this analysis come from a combination of two primary sources: the EAC and the Democratic National Committee (DNC).

The EAC data are from the 2006 Election Administration and Voting Survey, a combination of the 2006 Election Day Survey, the 2005-2006 National Voter Registration Act Survey, and the Uniformed and Overseas Citizens Absentee Voting Act (UOCAVA) Survey which, until 2006, were administered as three separate surveys. The survey was sent to election administrators in all 50 states, as well as the District of Columbia and the territories, via an internet based survey application. Each state was given a secure username and password in early December 2006 to use with the survey, and the deadline for submissions was March 7, 2007, although states could submit amended information until November 19, 2007. Twenty-eight of the survey’s questions pertain to statewide data only, with the remaining 30 questions pertaining to state and jurisdictional information, namely counties, although there are independent cities included in the survey as well. Some state administrators chose to share their login information with county or local officials to aid in the timely collection of data, while other states collected the county-level data themselves and then submitted the data to the survey. Almost half of the states submitted their data to the EAC in spreadsheet form, instead of using the online survey instrument (United States Election Assistance Commission, 2007, p. 5-6).

There are significant holes in the data because some states do not keep records of the information requested in the survey or required by HAVA; in total, however, the survey has at least some information on 3,123 jurisdictions of the 3,140 counties and independent cities in the United States, including the territories. No question in the survey has more than 3,004 responses (United States Election Assistance Commission, 2007, p. 6). Alaska did not report information
on any of its 18 boroughs, and instead only reported statewide numbers, and thus is not included in this analysis; New York did not include individual information on each of the five counties within New York City, but did include New York City; Illinois provided information on eight cities that are not recognized by the federal government as independent cities; and, Missouri provided information on one city that is not a recognized independent city. For purposes of this analysis, American Samoa, the District of Columbia, Guam, Puerto Rico, and the Virgin Islands have been removed from the dataset.

In addition to providing data by county, the EAC data include a row for each state. The state rows do not equal the sum of the individual counties in that state, and these additional state rows have been removed from the data for this analysis. Because some states did not answer every question at the county level, deleting these state rows means that some states are missing a lot of information; New York, for example, reported so few questions at the county level that this analysis only contains data on the total number of voter registration applications that were duplicates, changes of name, address, or party, and invalid or otherwise rejected and the type of voting technology in use.

Data from the DNC come from the DNC voter file, a compilation of publicly available voting records and demographic information for every registered voter in the United States. Vote history and individual voter information such as address and age comes from state Secretaries of State and, in some cases, local Boards of Election. Ethnicity is provided in some states by Secretaries of State and in others, likely ethnicity is acquired from commercial vendors. For the purposes of this analysis, all data are aggregated at the county level to match the data from the EAC. Additional information at the county level comes from the United States Census Bureau’s 2000 Census. Due to varying file acquisition dates, imperfect compliance with
HAVA reporting requirements, and other data quality issues at the state and local level, 2002 vote history is imperfect – in most cases, data are available only for those 2002 voters who were still on the voter file in 2006. Further, there are some people in both 2002 and 2006 for whom there is no county information available and these records are not included in the data used in this analysis.

Although the EAC has Election Administration and Voting Survey results dating back to 2004 and the DNC has many years of back data, this study looks specifically at the 2006 general election for several reasons. 2006 was four years after HAVA became law, allowing the states to have a little bit more time to work out some HAVA-related kinks in election administration that might have shown up in the 2004 data. The next presidential election after 2004 was 2008, a year in which voter turnout and enthusiasm were exceptional, even as compared to other presidential years. Ultimately, the goal of an analysis such as this one is to draw conclusions about election administration generally, and because 2008 was such an incredible year politically, it is unlikely that the results of an analysis using those data would be generalizable.

Ideally, a study such as this one would compare the 2006 data to data collected before HAVA’s passage and implementation so that the results would speak to HAVA’s effectiveness. Before HAVA, however, there was no uniform election administration data collection, making a pre/post comparison impossible. Some individual states and private-sector firms collected detailed election administration data, but before 2002 there was no nation-wide policy or standard for collecting the information necessary for this analysis, thus this study only examines 2006 data and tries to extrapolate HAVA’s impact from the results.
Variables

The primary dependent variable is the likelihood of casting a provisional ballot, or the number of provisional ballots cast divided by the total number of ballots cast. The average provisional ballot rate for counties in this dataset is 0.62 percent;\(^\text{13}\) Los Angeles County, California reported the most number of provisional ballots cast, with 110,915. All of Iowa’s 99 counties and 19 other counties throughout the country reported data incorrectly; these 118 jurisdictions reported provisional ballot totals equal to the total number of ballots cast in that county. Because this information is so obviously incorrect, these counties are recoded as missing. Taking this into account, this analysis contains provisional ballot rate for 2,312 counties, or 70.1 percent of jurisdictions in the dataset. States that did not provide information are: Connecticut, Idaho, Kentucky, Maryland, all but three counties in Massachusetts, Minnesota, New Hampshire, New York, North Dakota, Pennsylvania, Rhode Island, Tennessee, and Wisconsin. Several of these missing states, however, have valid reasons – Idaho, Minnesota, and New Hampshire had same day voter registration in 2006, eliminating the need for provisional balloting, and North Dakota does not have voter registration – so these states have no provisional ballots (Electiononline.org, 2006a, p. 32).\(^\text{14}\) Wisconsin and Wyoming both have same day registration, but provide provisional ballots to those who do not have the identification required to register; Wisconsin did not provide the number of provisional ballots they issued in 2006, while Wyoming did (Electiononline.org, 2006a, p. 32). Maine also has same day registration, but provided the number of challenge ballots cast. Thus, the primary offenders in this dataset are Connecticut, Kentucky, Maryland, Massachusetts, New York, Pennsylvania,

\(^{13}\) Mean = 0.006, Standard Deviation = 0.032

\(^{14}\) North Dakota is exempt from 1993’s National Voter Registration Act and HAVA because its statewide system of small precincts where election administrators know most of the voters is a state tradition rooted in its rural values. North Dakota was one of the first states to enact voter registration, and the first to eliminate it in 1951. For further information, see “What Does HAVA Mean for North Dakota?”
Rhode Island, and Tennessee. Many other states provide data that are, at best, incomplete, either because they did not provide any information on provisional balloting, because they did not keep a separate record of the number of provisional ballots cast, or because they reported zero provisional ballots when it is unlikely that there were no provisional ballots cast in that county.

There are several factors that likely contribute to the provisional ballot rate in a given county: the race or ethnicity of voters, the age of voters, the change in the proportions of registered African Americans and Hispanics, the change in proportions of registered youth and seniors, the number of precincts and the number of polling places, the number of people in that county who are new registrants, and whether the state had any kind of provisional balloting before HAVA. Economic indicators – median income and unemployment rate – are also included in the model as proxies for voter education levels.

The data from the DNC voter file include county demographic percentages for age and race from the 2000 Census as well as the demographic characteristics for voters who actually went to the polls. Michael Tomz and Robert P. Van Houweling argue that using the actual voter proportions eliminates a source of nonrandom measurement error from the analysis (2003, p. 49-50), and per their discussion, this analysis will use the proportion of African Americans, Hispanics, youth, and senior voters who turned out on Election Day.

The change in registration variables are calculated from the differences in actual registration proportions on the national voter file from 2002 to 2006. The median income and unemployment rate variables are from the 2000 Census, thus it is not possible to calculate change from 2002 to 2006. To make interpretation easier, median income is divided by 10,000.

The number of new registrants comes from the DNC data and includes every person who registered between January 1, 2003 and November 1, 2006. The EAC provides similar figures,
from the close of registration in 2002 to the close of registration in 2006, but like the other data from the EAC, several states chose not to provide this information at the county level. Thus, for completeness, the data used in this analysis come from the DNC voter file. The raw number of new registrants is divided by 10,000 to facilitate interpretation.

It is possible that counties in states that used some form of provisional balloting before HAVA might have higher provisional ballot rates because they were already familiar with the process and aware of what kinds of voters are eligible to cast provisional ballots. Before HAVA, states either used provisional balloting, affidavit balloting, or fail safe voting. Based on information from a late 2001 report issued by Electiononline.org and The Constitution Project, a dummy variable is created, where 1 indicates that the state did have some form of provisional balloting and 0 indicates that the state did not (Electiononline.org and The Constitution Project, 2001, p. 5). States with same day registration, including Wisconsin and Wyoming, are coded as not having a form of provisional balloting.\textsuperscript{15} 58.4 percent of states offered provisional ballots, affidavit ballots, or fail safe voting before HAVA required them to do so.

One expects that the proportion of African American voters and the proportion of Hispanic voters in a county will increase the likelihood of casting a provisional ballot because minority voters are more likely to be less informed about their polling place and election laws, and more likely to be taken advantage of by polling place workers. The proportion of young voters and senior voters in a county will increase the likelihood of casting a provisional ballot for reasons similar to those of minority voters. These four groups of voters are also more prone to registration errors. The number of precincts and polling places will both increase the likelihood of casting a provisional ballot because there will be more opportunities for voters to go to the

\textsuperscript{15} For more detailed information on state-by-state nuances, see Electiononline.org and The Constitution Project, 2001, p. 5.
wrong polling place and thus need to cast a provisional ballot. The number of new registrants in a county will increase the likelihood of casting a provisional ballot because new registrants may not be familiar with the voting process and could go to the wrong polling place; further, first time voters may not be familiar with the rules for voting, such as showing identification the first time you vote, or their registration may not appear on the list at the polling place, all reasons why they might have to cast a provisional ballot. It is not clear exactly what effect median income and unemployment will have on the provisional ballot rate. Median income could increase the provisional ballot rate, because voters with higher median incomes are more likely to demand a provisional ballot if they feel they are eligible to vote in a given precinct; however, it is also possible that lower income voters will have more registration errors resulting in an increase in the provisional ballot rate. Similarly, as the unemployment rate increases, the provisional ballot rate could decrease because voters will not be aware of their right to a provisional ballot, or it could increase because these voters may have more registration errors.

A second model will examine the likelihood of overvoting. Other analyses use a measure of residual votes to examine the performance of specific voting technologies, even though this measure “…cannot separate intentional nonvoting on a ballot choice from an error or voting machine malfunction” (Alvarez, 2009, p. 3-4). However, in their analysis of ANES and VNS polls, Tomz and Van Houweling (2003, p. 57) find that intentional undervoting is more common among African Americans than among whites, so including undervote rates in this model would be inappropriate. Thus, the number of undervotes is excluded, and this model uses the number of overvotes instead of residual ballot rate.

Many states reported that they did not collect information on overvoting or simply reported that there were no overvotes recorded; even within states that did report data for most
counts, some individual jurisdictions did not collect the necessary data and reported 0 overvotes or left the response blank. The EAC survey requested that states report overvote data by federal contest so that there would be a record of the number of overvotes per Senate race and per Congressional race. Some jurisdictions, however, miscoded the type of race and/or the Congressional district of the race; before publishing its final report on the 2006 election the EAC made efforts to correct the coding, but as a result, there are even more missing data than might otherwise be present. 68.4 percent of counties did not report any data on overvoting; among the 985 counties that reported data, 50.7 percent reported 0 overvotes. Of counties that reported overvotes, there was an average of 15.2 overvotes per county.\footnote{Mean = 15.2, Standard Deviation = 179.6} Madison County, Illinois recorded the most, with 4,529.

The data in this analysis match the overvote totals reported by the EAC in their final report for all states except Iowa, Illinois, and Maryland, although for three different reasons. The EAC data used in this analysis report 199 overvotes in Iowa that cannot be assigned to a specific county, so they are not included in this analysis; the EAC data in this analysis report 7,964 overvotes in Illinois, while the EAC final report shows only 7,917, and there is no obvious explanation for the discrepancy; finally, Maryland only reported overvotes at the congressional district level, so while it is possible to get a state total, it is not possible to get county level totals because the congressional districts do not follow county lines.

The variables expected to impact overvoting are very similar to those that impact the likelihood of casting a provisional ballot: the race or ethnicity of voters, the age of voters, the change in the proportion of registered African Americans, Hispanics, youth, and seniors in a county, and the number of people in that county who registered between January 1, 2003 and November 1, 2006. Median income and unemployment rate are included again as proxies for
education. In addition, the model includes variables describing the types of voting machines used in each county.

To measure the impact of the different voting technologies on the number of overvotes cast, the model uses dummy variables for each voting method: paper, lever, punch card, optical scan, and DRE. For each method, the counties were manually coded based off of their responses to the EAC survey, and 1 indicates that they use that type of technology. Each county may use multiple technologies, so the categories are not mutually exclusive, and, in fact, 821 of all counties in this study use multiple voting methodologies. Because these data are based on self-reported information from local election officials, in some cases it was impossible to determine the method of voting from the information provided, and in such cases, the county is left uncoded. Generally speaking, coding erred on the side of caution and was more, rather than less, generous. Counties that use vote by phone were coded as paper. In spite of these leniencies, however, missing data are pervasive in this area: Colorado, Florida, Indiana, Maine, Maryland, all but two counties in Massachusetts, Montana, Ohio, all but one county in Rhode Island, South Dakota, Texas, Washington, and all but two counties in Wyoming provide no information on the voting technologies used in the 2006 general election. Five of these states – Colorado, Florida, Indiana, Maryland, and Ohio – had substantial voting technology issues during their primaries or, in the case of Montana and Washington, made major changes in voting methodology, making their lack of data somewhat troubling (Electiononline.org, 2006a, p. 7-10). When possible, the EAC data in this analysis are supplemented with data from www.electiononline.org’s 2006 Election Preview, released in October of that year, and their post-election briefing, released in

---

17 For example, some administrators reported things like "paper ballots and m100." In this situation, it is unclear if they mean that they are using paper optical scan ballots or paper ballots and optical scan ballots. This was coded as paper and optical scan.

18 Montana switched to same day registration for the 2006 General Election; Washington’s King County used touch screen DRE machines for the first time in November 2006.
November 2006; doing so adds voting methodology for Maryland, Massachusetts, Rhode Island, South Dakota, as well as individual counties in Indiana, Florida, and Washington. Data are still missing or very incomplete in Colorado, Florida, Indiana, Maine, Montana, Ohio, Texas, Washington, and Wyoming.

Overall, though, the numbers reported in this dataset are mostly consistent with the aforementioned post-election report issued by www.electiononline.org that noted 62 counties used lever machines, 13 counties used punch card machines, and 1,142 counties used DRE machines (Electiononline.org, 2006b, p. 7); the data in this analysis show 71 counties used lever machines,19 15 used punch card machines, and 1,393 used DRE machines. Additionally, the data show that 281 counties used paper ballots and 1,475 used optical scan machines. However, the data in this analysis are not quite consistent with the numbers reported in the EAC’s 2006 report, likely due to differences in coding and the EAC’s ability to follow up with the states directly for clarification when necessary. The EAC report shows: 67 counties used lever machines, 13 used punch card machines, 1,671 used DREs, 99 used paper ballots, and 1,332 used optical scan; 927 jurisdictions reported using multiple voting methods (United States Election Assistance Commission, 2007, p. 68). Because the number of counties that used punch card ballots in 2006 is so small, the indicator variable for punch card balloting is omitted from this analysis.

Table 1 details the descriptive statistics for the demographic variables used in this analysis.

---

19 Note that this number counts New York City as one county, and every county in New York used lever machines in this election (United States Election Assistance Commission, 2007, p. 69).
Table 1: Descriptive Statistics on County Characteristics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Observations</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion African American Voted in County, 2006</td>
<td>3,107</td>
<td>0.085</td>
<td>0.150</td>
</tr>
<tr>
<td>Proportion of Hispanic Voted in County, 2006</td>
<td>3,107</td>
<td>0.031</td>
<td>0.082</td>
</tr>
<tr>
<td>Proportion of 18-25 Voted in County, 2006</td>
<td>3,107</td>
<td>0.039</td>
<td>0.019</td>
</tr>
<tr>
<td>Proportion of 65+ Voted in County, 2006</td>
<td>3,107</td>
<td>0.240</td>
<td>0.060</td>
</tr>
<tr>
<td>Number of New Registrants/10,000, 2002 to 2006</td>
<td>3,107</td>
<td>1.395</td>
<td>4.433</td>
</tr>
<tr>
<td>Change in Proportion Registered African American, 2002 to 2006</td>
<td>3,107</td>
<td>0.001</td>
<td>0.006</td>
</tr>
<tr>
<td>Change in Proportion Registered Hispanic, 2002 to 2006</td>
<td>3,107</td>
<td>-0.014</td>
<td>0.032</td>
</tr>
<tr>
<td>Change in Proportion Registered 18-25, 2002 to 2006</td>
<td>3,107</td>
<td>0.010</td>
<td>0.016</td>
</tr>
<tr>
<td>Change in Proportion Registered 65+, 2002 to 2006</td>
<td>3,107</td>
<td>0.020</td>
<td>0.011</td>
</tr>
<tr>
<td>Number of Precincts</td>
<td>2,837</td>
<td>55.4</td>
<td>186.2</td>
</tr>
<tr>
<td>Number of Polling Places</td>
<td>2,510</td>
<td>40.2</td>
<td>144.0</td>
</tr>
<tr>
<td>Median Income/10,000 ($), 2000</td>
<td>3,104</td>
<td>3.527</td>
<td>0.884</td>
</tr>
<tr>
<td>Unemployment Rate, 2000</td>
<td>3,104</td>
<td>0.057</td>
<td>0.027</td>
</tr>
</tbody>
</table>

Source: EAC 2006 Election Administration and Voting Survey and the DNC Voter File
Notes: Total Unweighted N = 3,118

Because missing data are so prevalent in both dependent variables, logistic regression models are used to determine what kinds of counties did not report data for these metrics at all. The dependent variables are coded as 1 if the county did not report any data on the number of provisional ballots or overvotes, and are 0 if the county reported any data on these metrics, even if the reported number is zero. The significant differences between those counties that reported data and those that did not, as shown in Table 2, suggest that while the findings in this analysis are instructive and a good starting point for further research, they are limited by poor data quality and should not be considered the final word on the subject.

Counties that did not report the number of provisional ballots have a significantly lower African American turnout in 2006, a higher rate of youth voting, a lower rate of senior voting, a
negative change in the proportion of registered youth from 2002 to 2006, a positive change in the proportion of registered seniors from 2002 to 2006, a lower median income, a lower unemployment rate, more polling places, fewer precincts, fewer new registrants, and are more likely to have previously had provisional balloting. Counties that did not report the number of overvotes have higher African American, youth, and senior turnout rates in 2006, a negative change in the proportion of registered African Americans, a higher median income, a higher unemployment rate, fewer new registrants, and are less likely to use paper, optical scan, or DRE balloting. Put succinctly, there is no consistent pattern that emerges to describe counties that did not report data on either dependent variable.
## Table 2: Results of a Logistic Regression Analysis of Counties That Did Not Report Data on Provisional Votes or Overvotes

<table>
<thead>
<tr>
<th>Provisional Vote Rate Model (Missing = 1)</th>
<th>Overvote Model (Missing = 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion of African American Voted in County, 2006</td>
<td>-1.475*** (0.392)</td>
</tr>
<tr>
<td>Proportion of Hispanic Voted in County, 2006</td>
<td>-6.007 (3.995)</td>
</tr>
<tr>
<td>Proportion of 18-25 Voted in County, 2006</td>
<td>6.385** (3.007)</td>
</tr>
<tr>
<td>Proportion of 65+ Voted in County, 2006</td>
<td>-7.695*** (1.118)</td>
</tr>
<tr>
<td>Change in Proportion Registered African American, 2002 to 2006</td>
<td>-4.652 (9.779)</td>
</tr>
<tr>
<td>Change in Proportion Registered Hispanic, 2002 to 2006</td>
<td>-5.053 (8.172)</td>
</tr>
<tr>
<td>Change in Proportion Registered 18-25, 2002 to 2006</td>
<td>-10.8*** (3.475)</td>
</tr>
<tr>
<td>Change in Proportion Registered 65+, 2002 to 2006</td>
<td>17.3*** (5.202)</td>
</tr>
<tr>
<td>Median Income/10,000 ($), 2000</td>
<td>-0.435*** (0.083)</td>
</tr>
<tr>
<td>Unemployment Rate, 2000</td>
<td>-7.038*** (2.327)</td>
</tr>
<tr>
<td>Number of Polling Places</td>
<td>0.009*** (0.002)</td>
</tr>
<tr>
<td>Number of Precincts</td>
<td>-0.004** (0.002)</td>
</tr>
<tr>
<td>Number of New Registrants/10,000, 2002 to 2006</td>
<td>-0.088* (0.052)</td>
</tr>
<tr>
<td>Previous Provisional</td>
<td>0.710*** (0.106)</td>
</tr>
<tr>
<td>Paper Balloting</td>
<td>--</td>
</tr>
<tr>
<td>Optical Scan Balloting</td>
<td>--</td>
</tr>
<tr>
<td>DRE Balloting</td>
<td>--</td>
</tr>
<tr>
<td>Lever Balloting</td>
<td>--</td>
</tr>
<tr>
<td>Constant</td>
<td>1.799*** (0.538)</td>
</tr>
<tr>
<td>N</td>
<td>2,494</td>
</tr>
<tr>
<td>Pseudo R²</td>
<td>0.069</td>
</tr>
</tbody>
</table>

Standard errors in parentheses. *** indicates significance at 0.01, ** indicates significance at 0.05, * indicates significance at 0.1
Many counties that did report data on the number of provisional ballots and the number of overvotes cast in the 2006 election reported 0 on these metrics. In most cases, there is no way to determine if these counties actually had no provisional ballots or overvotes, or if they reported 0 because they did not know the answer or did not understand what information was required of them. To determine what kinds of counties reported zero, logistic regressions are used again, this time using a binary variable where 1 indicates that the county reported zero provisional ballots or overvotes, and 0 indicates that the county reported more than zero provisional ballots or overvotes.

Counties that reported zero provisional ballots, as shown in Table 3, had a significantly higher rate of senior voting in 2006, a negative change in the proportion of registered youth from 2002 to 2006, a higher median income and unemployment rate, more polling places, fewer new registrants, and were less likely to have used a form of provisional balloting prior to the 2006 election. Counties that reported 0 overvotes had a significantly lower rate of seniors voting in 2006, a negative change in the proportion of registered youth voters from 2002 to 2006, fewer polling places, more precincts, and were less likely to use optical scan balloting and more likely to use DRE balloting than counties that reported a non-zero number of overvotes.

---

20 However, counties in Idaho, Minnesota, New Hampshire, and North Dakota – states with same day registration and without any form of provisional balloting – should, and do, report 0 provisional ballots.
Table 3: Results of a Logistic Regression Analysis of Counties That Reported Zero Provisional Ballots or Overvotes

<table>
<thead>
<tr>
<th></th>
<th>Provisional Vote Rate Model</th>
<th>Overvote Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion of African American Voted in County, 2006</td>
<td>-0.404 (0.432)</td>
<td>1.474 (1.045)</td>
</tr>
<tr>
<td>Proportion of Hispanic Voted in County, 2006</td>
<td>1.178 (3.767)</td>
<td>7.987 (7.175)</td>
</tr>
<tr>
<td>Proportion of 18-25 Voted in County, 2006</td>
<td>2.839 (3.458)</td>
<td>2.816 (8.194)</td>
</tr>
<tr>
<td>Proportion of 65+ Voted in County, 2006</td>
<td>4.769*** (1.495)</td>
<td>-6.682*** (2.115)</td>
</tr>
<tr>
<td>Change in Proportion Registered African American, 2002 to 2006</td>
<td>-15.5 (11.6)</td>
<td>26.9 (25.4)</td>
</tr>
<tr>
<td>Change in Proportion Registered Hispanic, 2002 to 2006</td>
<td>7.273 (11.9)</td>
<td>15.5 (15.0)</td>
</tr>
<tr>
<td>Change in Proportion Registered 18-25, 2002 to 2006</td>
<td>-26.3*** (4.324)</td>
<td>-22.4*** (8.436)</td>
</tr>
<tr>
<td>Change in Proportion Registered 65+, 2002 to 2006</td>
<td>-5.938 (6.201)</td>
<td>0.487 (8.225)</td>
</tr>
<tr>
<td>Median Income/10,000 ($), 2000</td>
<td>0.226* (0.118)</td>
<td>-0.208 (0.142)</td>
</tr>
<tr>
<td>Unemployment Rate, 2000</td>
<td>6.007*** (2.542)</td>
<td>0.990 (3.518)</td>
</tr>
<tr>
<td>Number of Polling Places</td>
<td>0.016** (0.007)</td>
<td>-0.015*** (0.005)</td>
</tr>
<tr>
<td>Number of Precincts</td>
<td>-0.007 (0.005)</td>
<td>0.002* (0.001)</td>
</tr>
<tr>
<td>Number of New Registrants/10,000, 2002 to 2006</td>
<td>-1.052*** (0.222)</td>
<td>-0.034 (0.085)</td>
</tr>
<tr>
<td>Previous Provisional</td>
<td>-2.122** (0.151)</td>
<td>--</td>
</tr>
<tr>
<td>Paper Balloting</td>
<td>--</td>
<td>-0.187 (0.384)</td>
</tr>
<tr>
<td>Optical Scan Balloting</td>
<td>--</td>
<td>-1.147*** (0.362)</td>
</tr>
<tr>
<td>DRE Balloting</td>
<td>--</td>
<td>1.746*** (0.242)</td>
</tr>
<tr>
<td>Lever Balloting</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Constant</td>
<td>-1.636** (0.670)</td>
<td>2.378*** (1.026)</td>
</tr>
<tr>
<td>N</td>
<td>1,899 (479)</td>
<td>679</td>
</tr>
<tr>
<td>Pseudo R²</td>
<td>0.247</td>
<td>0.235</td>
</tr>
</tbody>
</table>

Standard errors in parentheses. *** indicates significance at .01, ** indicates significance at .05, * indicates significance at .1

Note: Lever balloting was omitted from this analysis because, of the 70 counties that reported using this voting methodology in 2006, only 5 of those counties also reported a non-zero number of overvotes
Model

Although both dependent variables are continuous, each has a large number of cases at the lower limit of zero. The models used for analysis must account for clustering around 0. Thus, tobit models are used in both cases to control for the probability of a county being at 0. The resulting coefficients and standard errors are larger than with ordinary least squares, but the model overall is a better fit because of the large numbers of counties reporting 0 provisional ballots or overvotes.
Results

The strongest predictors of the provisional ballot rate are the change in proportion of registered voters age 65 or older, the number of polling places, the number of new registrants, and whether or not the state previously had some form of provisional balloting. The African American turnout rate also has a significant positive effect on the provisional ballot rate, though the relationship is not as strong as expected. On the one hand, this suggests that African American voters are taking advantage of provisional balloting options when they might otherwise have been denied their right to vote; however, this also suggests that African Americans have more registration problems that necessitate provisional ballots.

Seniors have the most experience voting of any other demographic group, so it is not surprising that increases in the proportion of registered seniors would have a negative impact on the provisional vote rate – these voters have voted in the past, know what they are doing, and know the location of their precinct and polling place, and are thus unlikely to require a provisional ballot. One could also surmise that because these voters traditionally have high levels of political awareness, they are knowledgeable about changes in voter identification laws and are conscientious about updating their registration information, situations that might otherwise necessitate provisional ballots. It is also unsurprising that the number of polling places has a positive impact on the provisional ballot rate – as the number of polling places increases, so does the likelihood that a voter will accidentally go to the wrong polling place to cast their ballot.

For each 10,000 new registrations, the number of provisional ballots decreases by 0.0011489. While it might seem counterintuitive that more new registrants would have a significant negative impact on the provisional ballot rate, new registrants might be more cautious
about going to the correct polling place and more inclined to double check the identification they need to bring with them because they are inexperienced and do not want to make a mistake.

States that had some form of provisional balloting prior to HAVA have a significantly higher provisional ballot rate than states did not have provisional balloting before 2003. County election officials and pollworkers in pre-HAVA states likely are more familiar with the provisional ballot process and eligibility than election administrators in states with less provisional ballot experience. The voters in the pre-HAVA states might also be more familiar with provisional balloting and the circumstances under which it is appropriate to ask for one.
Table 4: Results of a Tobit Regression Analysis of Provisional Ballot Rate on Demographic Variables

<table>
<thead>
<tr>
<th>Provisional Ballot Rate Model</th>
<th>Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion of African American Voted in County, 2006</td>
<td>0.011* (0.006)</td>
</tr>
<tr>
<td>Proportion of Hispanic Voted in County, 2006</td>
<td>-0.012 (0.037)</td>
</tr>
<tr>
<td>Proportion of 18-25 Voted in County, 2006</td>
<td>-0.009 (0.052)</td>
</tr>
<tr>
<td>Proportion of 65+ Voted in County, 2006</td>
<td>-0.015 (0.019)</td>
</tr>
<tr>
<td>Change in Proportion Registered African American, 2002 to 2006</td>
<td>0.157 (0.124)</td>
</tr>
<tr>
<td>Change in Proportion Registered Hispanic, 2002 to 2006</td>
<td>-0.078 (0.084)</td>
</tr>
<tr>
<td>Change in Proportion Registered 18-25, 2002 to 2006</td>
<td>0.062 (0.055)</td>
</tr>
<tr>
<td>Change in Proportion Registered 65+, 2002 to 2006</td>
<td>-0.208** (0.085)</td>
</tr>
<tr>
<td>Median Income/10,000 ($), 2000</td>
<td>0.000 (0.001)</td>
</tr>
<tr>
<td>Unemployment Rate, 2000</td>
<td>0.012 (0.034)</td>
</tr>
<tr>
<td>Number of Polling Places</td>
<td>0.000*** (0.000)</td>
</tr>
<tr>
<td>Number of Precincts</td>
<td>0.000 (8.70e-06)</td>
</tr>
<tr>
<td>Number of New Registrants/10,000, 2002 to 2006</td>
<td>-0.001** (0.000)</td>
</tr>
<tr>
<td>Previous Provisional</td>
<td>0.022*** (0.002)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.009 (0.009)</td>
</tr>
<tr>
<td>N</td>
<td>1,899</td>
</tr>
<tr>
<td>Pseudo R²</td>
<td>-0.050</td>
</tr>
</tbody>
</table>

Standard errors in parentheses. *** indicates significance at .01, ** indicates significance at .05, * indicates significance at .1
The strongest predictors of the number of overvotes are the number of polling places and precincts, and use of DRE ballot machines; the African American turnout rate is also significant, though less strongly. As the number of polling places increases, so does the number of overvotes, likely because more polling places provides more opportunities for error. The number of precincts, however, has a significant negative impact on the number of overvotes, a relationship for which there is no obvious explanation, suggesting that this could be related to poor data quality.

The large, significantly negative impact of DRE ballot machines suggest that these machines might provide the best opportunity for election administrators to ensure that every ballot cast is counted, as these machines do not permit overvoting.

The African American turnout rate has a significant negative effect on the number of overvote ballots cast, yet the published research suggests the exact opposite – that controlling for education or socioeconomic factors, African Americans have higher invalidation rates than whites. Tomz and Van Houweling’s research shows that on certain types of voting methodologies, namely punch cards and optical scan ballots, African Americans are more likely than whites to cast residual ballots, even when controlling for education or other socioeconomic factors (2003). Buchler, Jarvis, and McNulty also find that minority voters have higher residual ballot rates, although their research shows that punch cards are far worse for minority voters than optical scan ballots (2004). Nevertheless, using these articles and others as a guide, the results of this analysis are surprising, though it is possible that this result is due to the poor quality of the overvote data from the EAC.
### Table 5: Results of a Tobit Regression Analysis of Overvotes on Demographic Variables

<table>
<thead>
<tr>
<th></th>
<th>Overvote Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion of African American Voted in County, 2006</td>
<td>-343.7**</td>
</tr>
<tr>
<td></td>
<td>(146.2)</td>
</tr>
<tr>
<td>Proportion of Hispanic Voted in County, 2006</td>
<td>-748.8</td>
</tr>
<tr>
<td></td>
<td>(910.5)</td>
</tr>
<tr>
<td>Proportion of 18-25 Voted in County, 2006</td>
<td>-822.8</td>
</tr>
<tr>
<td></td>
<td>(1041.6)</td>
</tr>
<tr>
<td>Proportion of 65+ Voted in County, 2006</td>
<td>91.7</td>
</tr>
<tr>
<td></td>
<td>(272.2)</td>
</tr>
<tr>
<td>Change in Proportion Registered African American, 2002 to 2006</td>
<td>-3412.5</td>
</tr>
<tr>
<td></td>
<td>(3154.5)</td>
</tr>
<tr>
<td>Change in Proportion Registered Hispanic, 2002 to 2006</td>
<td>-513.3</td>
</tr>
<tr>
<td></td>
<td>(1719.4)</td>
</tr>
<tr>
<td>Change in Proportion Registered 18-25, 2002 to 2006</td>
<td>1429.4</td>
</tr>
<tr>
<td></td>
<td>(1040.3)</td>
</tr>
<tr>
<td>Change in Proportion Registered 65+, 2002 to 2006</td>
<td>539.3</td>
</tr>
<tr>
<td></td>
<td>(1103.4)</td>
</tr>
<tr>
<td>Median Income/10,000 ($), 2000</td>
<td>10.2</td>
</tr>
<tr>
<td></td>
<td>(17.6)</td>
</tr>
<tr>
<td>Unemployment Rate, 2000</td>
<td>121.9</td>
</tr>
<tr>
<td></td>
<td>(490.3)</td>
</tr>
<tr>
<td>Number of Polling Places</td>
<td>2.191***</td>
</tr>
<tr>
<td></td>
<td>(0.387)</td>
</tr>
<tr>
<td>Number of Precincts</td>
<td>-0.397***</td>
</tr>
<tr>
<td></td>
<td>(0.126)</td>
</tr>
<tr>
<td>Number of New Registrants/10,000, 2002 to 2006</td>
<td>9.444</td>
</tr>
<tr>
<td></td>
<td>(7.083)</td>
</tr>
<tr>
<td>Paper Balloting</td>
<td>34.1</td>
</tr>
<tr>
<td></td>
<td>(40.0)</td>
</tr>
<tr>
<td>Optical Scan Balloting</td>
<td>43.4</td>
</tr>
<tr>
<td></td>
<td>(44.7)</td>
</tr>
<tr>
<td>DRE Balloting</td>
<td>-101.2***</td>
</tr>
<tr>
<td></td>
<td>(29.2)</td>
</tr>
<tr>
<td>Lever Balloting</td>
<td>33.7</td>
</tr>
<tr>
<td></td>
<td>(133.0)</td>
</tr>
<tr>
<td>Constant</td>
<td>-176.4</td>
</tr>
<tr>
<td></td>
<td>(133.0)</td>
</tr>
<tr>
<td>N</td>
<td>683</td>
</tr>
<tr>
<td>Pseudo R²</td>
<td>0.030</td>
</tr>
</tbody>
</table>

Standard errors in parentheses. *** indicates significance at .01, ** indicates significance at .05, * indicates significance at .1
Although data quality issues make it difficult to draw robust conclusions, the results of the provisional ballot rate analysis demonstrate that the number of African American voters does have a significant impact on the provisional ballot rate, possibly suggesting that HAVA’s efforts to improve registration list accuracy are not working. At the same time, though, the mere fact that these voters now have a legal right to a provisional ballot marks an important step forward. Given that counties in states that had some form of provisional ballot access prior to the passage of HAVA have significantly higher provisional ballot rates than states that did not have provisional balloting before 2002, this suggests that with the passage of time, counties and individual voters will become more comfortable with provisional balloting procedures and the number of provisional ballots issued will increase, hopefully leading to decreased rates of administrative disenfranchisement among all voters, but particularly among those this analysis and others show to be most at risk – minority voters.

While HAVA encourages the replacement of lever and punch card ballot machines, it does not dictate what should replace these balloting methods. Looking at the results of the overvote analysis, and with the same data quality issues in mind, casting ballots on DRE machines has a very strongly significant impact on the number of overvotes. These machines appear to provide voters of all races, ethnicities, and ages with the best possible chance to have their votes counted, which at the end of the day, is the ultimate goal of the HAVA legislation.
Policy Recommendations

Based on this analysis, the most obvious, and perhaps the most impactful, policy recommendation is to collect better data. Although HAVA mandates that the EAC periodically conduct surveys on election administration issues, such as methods of voting and registration, there is no provision in HAVA mandating that the states respond to the survey with complete and accurate data, although the NVRA and UOCAVA, both of which fall under the EAC’s purview, do statutorily require the states to provide a base level of information for collection. As this analysis clearly demonstrates though, the data collected by the EAC, through no fault of their own, are riddled with holes and errors that make any kind of meaningful analysis difficult.

At first blush, it is easy to feel something akin to sympathy for the states when thinking about the large quantities of information they must collect from the counties and precincts to provide complete and meaningful data to the EAC and to the public; this feeling quickly disappears, though, once you realize that private sector companies collect significantly more information, and more detailed information, on their millions of customers every day. Walmart has, as of March 2010, 4,300 stores in the United States and serves millions of customers in every state in America, as well as in 14 other countries across the globe (Walmartstores.com 2010; Walmartstores.com 2011). In spite of their size, the company’s customer information is so extensive that they can tell you that bananas are the most frequently purchased items in their stores, or that they sell more than seven times their usual rate of Strawberry Pop-Tarts right before a hurricane (Gerken, 2009, p. 49). Although this is an extreme example, it certainly proves the point – accurate, detailed data collection is not just possible, it is already being done very effectively in other sectors.
It would be difficult for Congress to amend HAVA to mandate sanctions for states that provide incomplete or incorrect data; the legislation already represents an unprecedented federal incursion into election administration, an area historically left to the states per the 10th Amendment. There are, however, non-legal means of encouraging better data collection. In her 2009 book, *The Democracy Index: Why Our Election System is Failing and How to Fix It*, Heather K. Gerken, a law professor at the Yale University School of Law, proposes that a non-profit organization develops an index, similar to other high profile indices like the Environmental Performance Index, to rank the states on their election administration practices. Although missing data would be problematic, it would actually work in the index’s favor: states receiving low rankings because they did not collect the applicable data would have an incentive to start collecting that data to improve their rankings. Creating an index would certainly not be a magic bullet for election reform, and would raise other issues such as falsifying data or states competing on the wrong dimensions of the index, but it would encourage better data collection, which is an important place to start.

Along the same lines, it is clear that there is still much work to be done in improving the registration process. African Americans are more likely to cast provisional ballots than other groups in this analysis, suggesting that it is possible that these voters have more problems with their voter registration to begin with, and many problems with the registration process are directly related to the registration lists.

HAVA mandates the creation of statewide voter registration databases, which sounds straightforward enough, but in practice, there is a high level of variation among the states in terms of the quality and capability of their lists. As of 2009, 41 states have top-down registration lists, or lists that are organized and maintained at the state level, and 8 states – including
California, New York, and Illinois, home to approximately one-fifth of the American population – have bottom-up registration lists, in which the registration lists are organized and maintained at the county level; Texas has a hybrid system (American University’s Center for Democracy & Election Management, 2009, p. 9). Even states with top-down systems typically allow local jurisdictions to take control of the final list. The Carter-Baker Commission on Federal Election Reform’s 2005 report strongly recommended that all states create top-down systems because “…bottom-up systems are not capable of providing a complete, accurate, current and valid voter registration list. They are ineffective in removing duplicate registrations of individuals who move from one county to another and in coordinating with databases from other state agencies” (“Building Confidence,” 2005, p. 11). The counties should be assisting the states, not the other way around. Additionally, very few states engage in persistent list maintenance, which is critical to ensuring the accuracy of the registration list, particularly in our highly mobile society (American University’s Center for Democracy & Election Management, 2009, p. 9).

Furthermore, there is little or no interoperability among the state lists; that is, each state uses its own database design, and there is little or no exchange of information among the states. Lack of interoperability results in voters being registered in more than one state and a list riddled with inaccuracies. Fewer than half of the states conducted data exchanges to identify duplicate registrations before the 2008 election, and, as of 2009, only four states explicitly authorize data exchange with other states; Kansas performs the best in this area, with exchange agreements with 11 states (American University’s Center for Democracy & Election Management, 2009, p. 10). List interoperability would provide the states with a means to check and update the information on their lists, making Election Day backups and mishaps less frequent, and, if executed perfectly, would only require a voter to register one time in her life.
Although problems with data quality do mean that the results of this analysis must be
taken with a grain of salt, the results do demonstrate the benefits of DRE balloting systems in
decreasing the number of overvotes. No balloting method is perfect, but because DRE machines
do not permit overvoting and warn the voter about undervoting, these machines provide the voter
with the best opportunity to have her ballot counted. Much of the hesitation about DRE systems
seems to stem from the lack of a paper record, but manufacturers make DRE machines with a
VVPAT, either for use in a recount or for the voter to confirm that her vote was cast as she
intended, and such a record should assuage the concerns of many people. States that used
HAVA funds to purchase DREs that do not provide a VVPAT should replace these machines
with DREs that do provide auditable records or retrofit these machines to provide a record of
each vote, both because the law demands it and because doing so will encourage public
confidence in these machines.

That being said, there are still significant technological and administrative hurdles to
overcome with this type of balloting. Numerous accounts of elections over the past decade tell
of polling places that opened late because pollworkers forgot a disk or the DRE machines were
on the fritz due to loose cables. This technology has also been plagued with security fears
because the states do not have the resources to properly check and maintain the software. If the
states can commit to regularly auditing and inspecting their machines, and can commit to
keeping the machines in genuinely secure locations, then DRE machines could be the best
chance for every vote to be counted.
References


