United States General Accounting Office

GAO

Report to the Ranking Minority Member, Committee on Government Reform, House of Representatives

October 2001

ELECTIONS

Statistical Analysis of Factors That Affected Uncounted Votes in the 2000 Presidential Election

G A O

Accountability * Integrity * Reliability

GAO-02-122

Contents

Letter 1

Scope and Methodology 2
Results in Brief 3
Background 4
Ballots Cast and Uncounted Presidential Votes Varied by Type of Equipment 7
Voting Equipment Differed By Demographic Characteristics 7
Voting Equipment, Selected Demographic Characteristics, and State Differences Affected Counties' Percentages of Uncounted Votes 8
Our Results Are Generally Consistent With Those of Other Researchers 12

Appendix I **Technical Approach and Additional Results** 14

Our Database and Its Limitations 15
Regression Analysis: Approach and Results 17

Appendix II **GAO Contacts and Staff Acknowledgments** 21

GAO Contacts 21
Staff Acknowledgments 21

Tables

Table 1: Ballots Cast and Uncounted Presidential Votes by Type of Voting Equipment 7
Table 2: Characteristics of Counties Used in GAO's Analyses 8
Table 3. Coefficients of Various Regression Models Used to Estimate the Percent of Uncounted Votes across Counties 17

Figure

Figure 1: Factors Accounting for Variation in Uncounted Presidential Votes Across Counties 11

G A O
Accountability * Integrity * Reliability

United States General Accounting Office
Washington, DC 20548

October 15, 2001

The Honorable Henry Waxman
Ranking Minority Member
Committee on Government Reform
House of Representatives

Dear Mr. Waxman:

Following the 2000 presidential election, a number of issues have been raised concerning the election process, including the limitations of certain types of voting equipment[1] in rendering a complete and accurate vote count. Further, concerns have been raised about the possibility that minorities and disadvantaged voters were more likely to have their votes not counted because they may have cast their ballots using less reliable voting equipment than affluent white voters. A limited body of prior research exists that has studied these specific issues or a subset of these issues in a comprehensive, systematic, and empirical manner.

You asked us to provide information on uncounted presidential votes in the November 2000 general election and the extent to which these uncounted votes could be attributed to counties' voting equipment and demographic characteristics. We further examined how much of the difference in uncounted presidential votes across counties was related to the state in which the counties are located, as well as the potential role of error correction[2] in reducing uncounted votes due to voting errors. Information on the reason votes for President were not counted was not available, but may include voter error, equipment failure, election officials' errors, or intentional nonvoting for the office of President.

[1] For ease of presentation, we use the term "equipment" to refer to the five methods by which votes were cast and counted in the 2000 presidential election. The five methods were paper ballot, lever machine, punch card, optical scan, and electronic.

[2] Error correction refers to the ability of certain types of equipment to identify when a voting error has occurred (e.g., if the voter cast a ballot that registered more than one vote for the office of President). When error correction is used at the precinct level, voters are notified that they have made an error on their ballot that would prevent their vote from being counted and are given the opportunity to correct the error.

Scope and Methodology

To address your request, we matched selected demographic data from the U.S. Census Bureau with voting equipment, voter turnout, and presidential vote data obtained from Election Data Services (EDS) and the Internet web sites of state election officials. We statistically analyzed county-level data to investigate relationships among counties' demographic characteristics, their voting equipment, and their percentages of uncounted presidential votes. We also statistically controlled for the state in which counties are located. We included data from 43 states[3] and the District of Columbia, representing 78 percent of the counties in the United States. Our results should not be generalized beyond this set of locations. The county demographic characteristics included from the 2000 Census were population size, racial composition (percent of African American and Hispanic residents in the county), and age (percent of 18-24 year olds and residents over 65). We included estimates of median income and percent of residents living below the poverty level from a 1997 Census model, and education data (percent of high school graduates in a county) from the 1990 Census.

We measured uncounted presidential votes by subtracting the number of votes for President from the number of total ballots cast. Both numbers were included in EDS' data along with voting equipment information for each county. We supplemented the analysis using GAO survey data from a representative sample[4] of county election officials to obtain further information on the use of error correction in conjunction with the various types of voting equipment.

Because of the unavailability of comprehensive data, we could not determine why votes for President were not counted; could not distinguish between ballots cast at the polling place on election day and those cast by absentee ballot or through early voting; and could not assess the reliability of different models of the same type of voting equipment. Additional information on our methodology and its limitations is provided in appendix I. We conducted our work from March through October 2001 in accordance with generally accepted government auditing standards.

[3] We excluded all voting jurisdictions in Alaska because they did not correspond directly to election districts. Additionally, Arkansas, Maine, Mississippi, Missouri, Pennsylvania, and Wisconsin were excluded because they did not report the necessary data to calculate uncounted votes.

[4] The sample is a stratified random sample of election jurisdictions nationwide. See Appendix I for more details.

Results in Brief

While the state in which counties are located had more of an effect on the number of uncounted presidential votes than counties' demographic characteristics or voting equipment, all three factors had statistically significant effects on uncounted presidential votes.

The type of voting equipment that counties used in the 2000 general election, for example, had an effect on uncounted presidential votes. The largest percentages of uncounted presidential votes tended to occur in counties that used punch card equipment. Counties that used optical scan equipment with error correction had percentages of uncounted presidential votes that were about 1.1 percentage points lower than counties with punch card equipment. Potentially, an estimated 300,000 additional presidential votes may have been counted if counties that used punch card equipment had, instead, used precinct-based optical scan equipment with error correction. We did not have data available to assess the extent to which other equipment changes, such as error correction with punch card equipment, could have reduced the total number of uncounted presidential votes.

Counties' demographic characteristics also affected their percentages of uncounted presidential votes. Specifically, counties with higher percentages of minority residents tended to have higher percentages of uncounted presidential votes, while counties with higher percentages of younger and more educated residents tended to have lower percentages of uncounted presidential votes. Counties that used punch card equipment did not generally have higher percentages of minority, less educated, or lower-income residents.

We found that the state in which counties are located had a greater effect on counties' percentage of uncounted presidential votes than did counties' voting equipment and demographic characteristics combined. State differences accounted for 26 percent of the total variation in uncounted presidential votes across counties. State differences may have included such factors as statewide voter education efforts, state standards for determining what is a valid vote, the use of straight party ballots, the number of candidates on the ballot, the use of provisional ballots, and the extent to which absentee or early voting occurred. County demographic characteristics accounted for 16 percent of the variation, and voting equipment accounted for 2 percent of the variation. A supplemental analysis of a subset of 404 counties showed that using optical scan equipment with error correction accounted for an additional 4 percent of the variation in counties' uncounted presidential votes. The remaining 52 percent of the variation was due to unknown factors for which we had no

data, such as whether a county switched to a new type of voting equipment or the number of inexperienced voters in a county.

Background

Each state and the District of Columbia play a role in elections by establishing election laws, policies, and procedures. In most states, counties are responsible for conducting elections, including selecting countywide voting equipment, counting ballots, and reporting elections results. In separate reports, we provide more in-depth information on election issues relating to people, processes, and technology at the county and state levels.[5]

Types of Equipment Used in the November 2000 Election

The equipment on which votes were cast and counted in the November 2000 election can be placed into five categories: paper ballots, lever machines, punch cards, optical scan, and electronic. Three of these five types of equipment—lever, optical scan, and electronic—have some capability or can be used to prevent or allow for the correction of voting errors.

Paper ballots. Paper ballots list the names of the candidates and the issues to be voted on. Voters generally complete their ballots in the privacy of a voting booth, recording their choices by placing marks in boxes corresponding to the candidates' names and the issues. After making their choices, voters drop the ballots into sealed ballot boxes. Election officials gather the sealed boxes and transfer them to a central location, where the ballots are manually counted and tabulated.

Lever machines. Lever machine "ballots" consist of a rectangular array of levers. Printed strips listing the candidates and issues are placed next to each lever. Voters cast their vote by pulling down the levers next to the candidates or issues of their choice. After voting, the voter moves a handle, which automatically records the vote and resets the levers. Votes are tallied by mechanical counters, which are attached to each lever. At the close of the election, election officials tally the votes by reading the counting mechanism totals on each lever voting machine. A feature inherent to lever machines is that they prevent voters from overvoting (i.e., voting more than once for the same office, unless the ballot explicitly

[5] *Elections: Perspectives on Activities and Challenges Across the Nation* (GAO-02-03, Oct. 15, 2001).

allows for more than one choice to be made). Overvoting is prevented by the interlocking of the appropriate mechanical levers in the machine.

Punch cards. Punch card voting equipment generally consists of a ballot, a vote recording device, a privacy booth, and a computerized tabulation device. Votes are cast by inserting the ballot into the vote recording device and punching a hole through the ballot such that the hole corresponds to the voter's ballot choice. Votes cast on punch card equipment are machine readable. Votes are tabulated using vote tabulation machines, and software is used to program each vote tabulation machine to correctly assign each vote read into the computer to the proper race and candidate or issue. The two basic types of punch card devices are Votomatic and Datavote.

Optical scan. An optical scan voting system is comprised of computer-readable ballots, appropriate marking devices, privacy booths, and a computerized tabulation machine. The ballot lists the names of the candidates and the issues. Voters record their choices using an appropriate writing instrument to fill in boxes or ovals, or to complete an arrow next to the candidate's name or the issue. Like punch card software, the software for optical scan equipment is used to program the tabulation equipment to correctly assign each vote read into the computer to the proper race and candidate or issue. Optical scan equipment based in precincts can be programmed to detect and reject both overvoting and undervoting (i.e., not registering a vote for every race and/or issue on the ballot). Using such error correction technology could allow voters to fix their mistakes before leaving the polling place. If ballots are tabulated centrally, voters do not have the opportunity to correct mistakes that may have been made.

Electronic. Electronic equipment (also called Direct Recording Electronic or DRE) comes in two basic types, pushbutton or touchscreen, with the pushbutton being the older and more widely used of the two. For pushbuttons, voters press a button next to the name of the candidate or the issue, which then lights up to indicate the selection. Similarly, voters using touchscreens make their selections by touching the screen next to the candidate or issue, which is then highlighted. When voters are finished making their selections, they cast their votes by pressing a final "vote" button or screen. Because all electronic equipment is programmable, it does not allow overvotes. In addition, voters can change their selections before hitting the final button to cast their votes.

Other Studies of Uncounted Votes

There have been several broad-based studies that have examined relationships among voter demographics, voting equipment, and/or uncounted votes. These studies, whose methods and findings we did not independently verify, included the following.

- A recent research study estimated that about 1.5 million voters thought they had voted for President but did not have their votes for President counted in the 2000 election. Faulty voting equipment and confusing ballots were among the stated reasons for the ballots being unmarked, spoiled, or too ambiguous to count. The study reported that punch card and electronic voting equipment were associated with uncounted votes for President exceeding 2 percent of all ballots cast. (CalTech/MIT, July 2001.)

- Another recent research study reported that, despite the perception that minorities and poor people were disproportionately more likely to vote on antiquated voting machinery and therefore have their ballots invalidated, the data did not support this contention. The study found that in the majority of states, whites and non-poor voters were more likely than African Americans and poor voters to reside in counties that used punch card equipment, based on 1998 voter equipment data. (Knack & Kropf, Jan. 2001.)

- A study of invalidated ballots in the 1996 presidential election found that counties with more African Americans and Hispanics were more likely to have higher rates of invalidated ballots, particularly in counties using punch card machines, optical scanners with centralized (as opposed to precinct-based) counting, and hand-counted paper ballots. When counties used equipment that can be programmed to prevent overvoting (i.e., lever technology, electronic voting technology, and precinct-count optical scan systems), racial differences in the rate of invalidated votes disappeared. (Knack & Kropf, May 2001.)

- A study of the 2000 presidential election found that the percentage of uncounted votes in 20 congressional districts with low-income/high-minority populations were higher, regardless of the type of voting equipment used, than in 20 congressional districts with high-income/low-minority populations. In both types of districts, the percent of uncounted votes was highest when punch card equipment was used. (House Committee on Government Reform, Minority Staff, Special Investigations Division, July 2001)

Ballots Cast and Uncounted Presidential Votes Varied by Type of Equipment

In the November 2000 presidential election, there were over 85 million votes cast in the 2,455 counties in our analysis and, of those, 1.6 million votes for President were not counted. The percentage of uncounted votes ranged from 0 percent to 23 percent, with an average of 2.3 percent. Only 12 counties had percentages of uncounted votes that exceeded 10 percent. Of the 2,455 counties, 284 (or 12 percent) used electronic voting equipment, 381 (16 percent) used lever equipment, 1,095 (45 percent) used optical scan equipment, 213 (9 percent) used paper ballots, and 482 (20 percent) used punch card equipment.[6] Furthermore, Table 1 shows that while 35 percent of the ballots cast came from counties using punch card equipment, 49 percent of the uncounted presidential votes were cast on punch card equipment.

Table 1: Ballots Cast and Uncounted Presidential Votes by Type of Voting Equipment

Voting equipment	Ballots cast		Uncounted votes	
	Number	Percent	Number	Percent
Electronic	11,604,770	14	184,132	11
Lever machines	13,557,499	16	255,196	16
Optical scan	29,338,967	34	386,011	23
Paper ballots	634,407	1	12,010	1
Punch card	30,195,730	35	805,635	49
Total	**85,331,373**	**100**	**1,642,984**	**100**

Source: GAO analysis of EDS data.

Voting Equipment Differed By Demographic Characteristics

Counties with different voting equipment differed demographically. (See table 2.) Counties that used punch cards, for example, had larger populations; higher median incomes; and smaller percentages of residents over 65 years of age and persons living below the poverty level than counties using other types of voting equipment. Our analysis did not show that minorities, or persons with less education or income, were more likely than others to be found in counties that used punch card voting equipment, the equipment associated with higher percentages of uncounted presidential votes. As the final row of table 2 shows, before controlling for demographic characteristics or state differences, the average percent of uncounted presidential votes was higher in counties

[6] The sum of the percentages of types of voting equipment does not equal 100 percent due to rounding.

that used punch cards (2.9 percent) than in other counties (2.1 percent to 2.3 percent).

Table 2: Characteristics of Counties Used in GAO's Analyses

	Voting equipment used				
Averages for County Characteristics	**Punch card (482)**	**Electronic (284)**	**Lever machine (381)**	**Optical scan (1,095)**	**Paper ballots (213)**
Population	172,612	108,913	96,389	70,464	6,382
Percent African American	6.3	10.8	17.8	7.6	1.3
Percent Hispanic	6.5	6.1	3.3	7.6	3.0
Percent high school graduates	46.3	41.7	40.4	45.7	48.6
Percent 18 to 24 years old	9.3	9.2	9.5	8.9	6.6
Percent over 65 years of age	13.9	13.4	13.8	15.0	18.5
Percent below poverty level	13.1	16.1	16.9	14.8	14.5
Median income	35,513	32,692	31,587	33,066	28,963
Percent uncounted votes	**2.9**	**2.3**	**2.2**	**2.1**	**2.3**

Note: Numbers of counties for which we had complete data on voting equipment, uncounted presidential votes, and demographic characteristics are given in parentheses. Counties with mixed equipment were not used in our analysis.

Source: GAO's analysis of EDS' and the Census Bureau's data.

Voting Equipment, Selected Demographic Characteristics, and State Differences Affected Counties' Percentages of Uncounted Votes

Overall, while we found that counties' percentages of uncounted presidential votes were related to their voting equipment and demographic characteristics, these factors accounted for less of the variation in uncounted votes across counties than did the state in which the county is located.

To determine how the percentages of uncounted votes across the counties for which we had data were affected by voting equipment, demographic characteristics, and the state in which counties are located, we used robust regression models that adjusted for the clustering (i.e., the lack of independence) of observations within states. Our statistical model included type of voting equipment, county demographic variables, and a set of variables to control for differences across states in which counties are located. (See app. I, table 3 for a more detailed discussion of all models and effects.)

Counties That Used Punch Card Equipment Had Higher Percentages of Uncounted Presidential Votes

Our statistical model indicated that there were no significant differences in uncounted presidential votes among counties that use electronic, paper, and optical scan voting equipment. Counties with punch cards had percentages of uncounted presidential votes that were roughly 0.6 percentage points higher than those counties, and counties with lever machines had percentages of uncounted presidential votes that were 0.7 percentage points lower than those counties. Given that the average of the uncounted presidential votes across all counties was roughly 2 percent, these represent sizable, as well as statistically significant, differences.

When the same statistical model was run for the subset of 404 counties that we surveyed, we found an additional equipment effect. The survey asked county election officials if they used equipment that either prevented errors or identified errors for voters so the ballot might be corrected. Since both electronic and lever equipment prevent overvotes, almost all of the counties using those types of equipment reported using error correction. In addition, almost all of the counties using punch card equipment and paper ballots reported not having or using error correction capabilities. Therefore, responses to the survey allowed us to distinguish between counties with optical scan equipment that used error correction and those that did not use it. Doing so resulted in significant differences between types of equipment. Counties using punch cards had uncounted presidential votes that were 1.1 percentage points higher than counties using error-corrected optical scan equipment. If we apply these results to the larger set of 2,455 counties, an estimated 300,000 additional votes may have been counted if counties that used punch card equipment had, instead, used precinct-based optical scan equipment with error correction.

Race, Education, and Age Affected Uncounted Presidential Votes

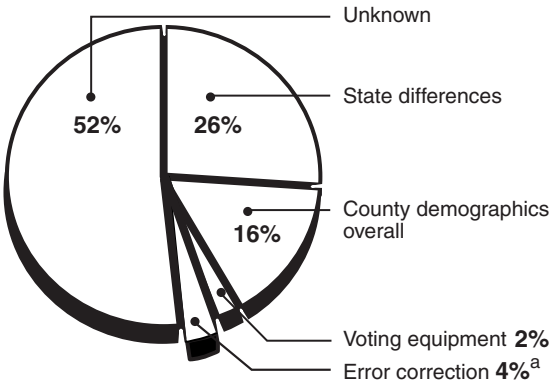
After we statistically controlled for the effects of state differences and voting equipment, uncounted presidential votes in our dataset of 2,455 counties were significantly higher in counties with higher percentages of African Americans and Hispanics. Each percentage point increase in a county's population of African Americans was associated with a 0.02 percentage point increase in the county's uncounted presidential votes. Each percentage point increase in a county's population of Hispanics was associated with a 0.01 increase in the county's uncounted presidential votes. This means, for example, that we would expect that a county where African Americans made up 35 percent of the population would have had uncounted presidential votes that were 0.6 percentage points higher than a county where African Americans made up 5 percent of the population.

After we statistically controlled for the effects of state differences and voting equipment, uncounted presidential votes in our dataset of 2,455 counties were significantly lower in counties with higher percentages of high school graduates and 18- to 24 year-olds. Each percentage point increase in a county's population of high school graduates was associated with a 0.06 percentage point decrease in the county's uncounted presidential votes. Likewise, each percentage point increase in a county's population of 18- to 24-year-olds was associated with a 0.03 percentage point decrease in the county's uncounted presidential votes. This means, for example, that we would expect that a county where high school graduates made up 50 percent of the population would have had uncounted presidential votes that were 1.8 percentage points lower than a county where high school graduates made up 20 percent of the population.

State Differences Accounted for More Variation in Uncounted Presidential Votes Than Voting Equipment and Demographic Variables Combined

We next determined the incremental effects of voting equipment, county demographics, and state differences on counties' percentage of uncounted presidential votes. When we just included type of equipment in the statistical model, we found that equipment alone explained 2 percent of the variation in uncounted presidential votes across counties. When we added demographic variables to that model, the county demographics explained an additional 16 percent of the variation. Next, we included a set of variables to statistically control for differences across the states in which counties are located. This made it possible to account for an additional 26 percent of the variation in uncounted presidential votes. A supplemental analysis of a subset of 404 counties that we surveyed showed that including a county's use of error correction with optical scan equipment would explain an additional 4 percent of the variation in uncounted votes across counties.

Figure 1: Factors Accounting for Variation in Uncounted Presidential Votes Across Counties



Source: GAO's analysis of EDS' data, the Census Bureau's data, and GAO's survey data.

Differences across states were of considerable importance in determining the prevalence of uncounted presidential votes and accounted for more of the variability (26 percent) in uncounted presidential votes across counties than demographic characteristics and type of voting equipment used combined. The following factors, for which we had no data because they have not been measured in a comprehensive, systematic way, are among those that may have contributed to differences among states: (1) voter education efforts, such as making sample ballots available prior to election day; (2) the use of straight party ballots that enable voters to make one entry to cast votes for all offices on the ballot; (3) the number of candidates on the ballot (including presidential, gubernatorial, or congressional candidates); (4) the number of provisional ballots cast,[7] and percentage of provisional ballots that were not counted; and (5) the extent to which absentee and/or early voting occurred and if such ballots were counted using a different voting equipment than ballots cast on election day.

The remaining 52 percent of the variation was due to unknown factors for which we had no data, such as whether a county switched to a new type of voting equipment or the number of inexperienced voters in a county.

[7] Provisional ballots are ballots that are cast by voters who may not be properly registered when they arrive at the polling place. After election day, their situation is reviewed and election officials make a decision as to whether the vote should be counted or not.

Our Results Are Generally Consistent With Those of Other Researchers

Like all four of the studies cited earlier in this report, we found that punch card equipment was associated with higher percentages of uncounted votes in counties, although our findings did not indicate, as did those of CalTech/MIT, that electronic voting equipment was similarly problematic. We also found, like Knack and Kropf, that minorities and persons with lower income were not more likely than others to reside in counties that used punch cards, and that counties with higher percentages of African Americans had higher percentages of uncounted presidential votes. We did not find, however, that the racial difference "disappears" in counties with certain voting equipment. Also, while there were differences between our study and that of the Special Investigation (e.g., our analytic methods did not involve making the same specific comparisons, and we analyzed counties while they analyzed congressional districts, our results do indicate, like theirs, that regardless of voting equipment, percentages of uncounted presidential votes were higher in high minority areas than in other areas.

To the extent that our results are not consistent with the findings of others, factors that may account for these differences include the variables included in the analyses, the number of counties included in the dataset, and the age of the data used by the different studies.

This report is one of several GAO studies addressing election issues. Our other reports discuss in greater depth election issues such as the scope of congressional authority in election administration, voter registration, absentee and early voting, voting assistance for military and overseas voters, election day administration, voting accessibility for voters with disabilities, vote counts and certification, Internet voting, and voting equipment standards.

We are sending copies of this report to the Chairman of your Committee and to other congressional committees. Staff members who contributed to this review are acknowledged in appendix II. If you or your staff have any questions about this report, please contact me on (202) 512-8777.

Sincerely yours,

Laurie E. Ekstrand
Director, Justice Issues

Appendix I: Technical Approach and Additional Results

This appendix provides information on our analyses of uncounted presidential votes in the November 2000 general election and the extent to which these uncounted votes were affected by counties' voting equipment, demographic characteristics, and state differences. It also discusses a separate analysis of a subset of counties in which we explored the potential of using optical scan equipment with error correction capability to reduce uncounted votes.

Data Sources

To obtain this information, we purchased data on equipment used in and results of the November 2000 election from Election Data Services (EDS), a company that compiles data on election administration and election results from the election jurisdictions of each state. Using EDS' election results data, we could calculate the number of uncounted presidential votes by subtracting the number of votes for President from the number of total ballots cast.

For the most part, EDS' data files for the 2000 presidential election are county level tabulations of election returns, voter participation, election official contact information, and voting equipment information. For Alaska, data are provided for election districts and regions, rather than counties; for the New England states, additional data are included for cities and townships within counties.

We matched data from the U.S. Census Bureau on selected demographic characteristics of each county with data on voting equipment and election results from EDS. From the 2000 Census, we included the following demographic variables in our analyses: population size, racial composition (percent African American and percent Hispanic), and age (percent 18 to 24 and over 65). From a 1997 Census model, we used estimates of median income and percent of residents living below the poverty level. Because more current data were not available, we used education (percent high school graduates) from the 1990 Census. We selected these demographic variables to include in our analysis because they have been included in prior studies of uncounted votes.

We also analyzed data for a subset of 404 counties whose election officials were surveyed by GAO in May 2001. The sample frame consisted of (1) all county election jurisdictions, or their equivalents, in 39 states that delegate election responsibilities primarily to counties; (2) the largest minor civil division in each county in the nine states that delegate election responsibilities to minor civil divisions; (3) the District of Columbia; and (4) Alaska. The sample was a stratified random sample of 607 election

jurisdictions nationwide selected from three strata—jurisdictions that used electronic voting equipment; those that used optical scan; and those that used any other method, including punch cards, lever machines, and hand-counted paper ballots. Of the 607 questionnaires sent, 513 usable questionnaires were returned. In our analyses of the questionnaire data, we included responses from the 404 counties and excluded responses from minor civil divisions to remain consistent with the unit of analysis in our larger county level analysis. One question in the survey asked: "Did the voting equipment used for votes cast at precincts on Election Day for the November 2000 general election either prevent errors or identify errors for voters so they could correct their ballots at the polling place?" From the responses to this question, we were able to distinguish, for these 404 counties, those using optical scan equipment with error correction and without.

Data Quality

We verified EDS' voting equipment data using several sources. Specifically, we (1) checked the Internet sites of 10 secretaries of state, (2) reviewed 2 state reports that provided information on the voting equipment used by counties and/or minor civil divisions, and (3) reviewed responses to a nationwide mail survey of election jurisdictions for other elections work GAO undertook. We made corrections where necessary.

To verify and augment election results data EDS provided us, we checked the Internet sites of secretaries of state and spoke with several state and county election officials. As result of these efforts, we verified a substantial portion of EDS' election results data and added 230 counties in 4 states (Delaware, Oklahoma, Tennessee, and West Virginia) and 147 counties in Texas to our database.

Based on the extent and nature of our data verification, we are confident that the data used in our analyses are of sufficient quality to support our conclusions.

Our Database and Its Limitations

Our database consisted of demographic, voting equipment and election results data for each of 2,455 counties in 43 states and the District of Columbia. The database included 78 percent of the nation's 3,141 counties at the time of the 2000 presidential election. To our knowledge, these data were the most recent, comprehensive, and valid data available to address the research questions specified for our study. Notwithstanding the strengths of our database, the precision of our analytic results and our

ability to explain why they occurred are limited by a number of factors, including missing data, omitted variables, and measurement error.

For several reasons, we did not include a number of states and counties in our database. Specifically, we excluded (1) all counties in 6 states (Arkansas, Maine, Mississippi, Missouri, Pennsylvania, and Wisconsin), 107 counties in Texas, 1 county in Alabama and 1 county in Oklahoma because these counties did not report the necessary data to calculate uncounted votes; (2) all voting jurisdictions in Alaska because they did not correspond directly to election districts; (3) counties that used a mix of voting equipment; (4) counties in which the reported numbers of votes cast for President exceeded the number of persons who turned out to vote; and (5) 1 county in which it appeared that only half the persons who turned out to vote cast a vote for President.

Our results should be interpreted with caution for the following reasons: (1) The available data did not distinguish between votes cast at the polling place on election day and those cast by absentee ballot or through early voting. Because some locations used different equipment for absentee and/or early voting, we could not assess the impact of such differences on our results. (2) We did not have information on the particular model of voting equipment used, so uncounted presidential votes, even within a single county, may have been affected by differences in the reliability of different models of the same equipment. (3) We used aggregate county-level demographic data as a proxy for the characteristics of voters because we did not have data on individual voters. (4) We could not determine why votes for President were not counted. For example, we could not discern if uncounted presidential votes were due to voter error, equipment failure, errors on the part of election officials, or intentional nonvoting for the office of President.[1] (5) In the absence of more current data, we analyzed 1990 Census data on education, which may have had different relationships with other variables in 2000 than it did in 1990. The extent to which such relationships may have changed is unknown. (6) Because our data on income and poverty were estimates derived from statistical models, they contained an unknown amount of measurement error that could not be accounted for in our statistical models.

[1] Research by others has indicated that the percentage of voters who reported deliberately not voting for President in the 2000 election was small (0.34 percent).

Regression Analysis: Approach and Results

Our analyses included, along with descriptive statistics, analysis of variance methods and robust regression models that account for the clustering.

To determine how the percentage of uncounted presidential votes was affected by the voting equipment employed in and the demographic characteristics of the counties for which we had data, we used a series of four robust regression models that adjusted for the clustering (i.e., the lack of independence) of observations within states. Model 1 in table 3 indicates that when demographic and other differences across counties are ignored, the average percentage of uncounted votes was significantly higher in counties that used punch card equipment than in counties that used optical scan equipment (which is the deleted referent category). Counties that used electronic, paper, or lever equipment, on the other hand, were not significantly different from those that used optical scan equipment. The R-squared value (i.e., the value representing the proportion of variation that the statistical model explained) for Model 1 indicates that differences in voting equipment accounted for only 2 percent of the variation in the percentage of uncounted votes across counties. This effect of voting equipment on uncounted votes may be due to various differences between types of equipment such as the design of the equipment by the manufacturer, the operation of the equipment by voters, or the processes that election officials used to prepare and operate the equipment.

Table 3. Coefficients of Various Regression Models Used to Estimate the Percent of Uncounted Votes across Counties

Characteristic	Model 1	Model 2	Model 3	Model 4
Intercept	2.07	5.59	6.03	3.39
Punch card	0.80**		1.08**	0.63**
Electronic	0.20		0.11	-0.32
Lever	0.10		-0.36	-0.72**
Paper	0.19		-0.10	-0.35
Population (logged)		-0.20**	-0.27**	-0.13
Percent African American		0.03*	0.03**	0.02**
Percent Hispanic		0.00	0.00	0.01*
Percent high school graduates		-0.03	-0.04	-0.06**
Percent 18 to 24 years old		-0.02	-0.02	-0.03**
Percent over 65 years old		0.01	0.02	0.04
Percent below poverty level		0.00	0.01	0.02
Median income (in 1000s)		0.00	0.00	0.00
R- squared	0.02	0.12	0.18	0.44

Notes: All models are robust regression models that account for clustering, and the lack of independence of observations, within states. Model 4 differs from Model 3 by including a set of 42 dummy variables to allow for effects of unmeasured state characteristics. We have omitted the coefficients associated with the dummy variables to simplify our presentation.

* Statistically significant at the 0.05 confidence level.

** Statistically significant at the 0.01 confidence level.

Source: GAO's analysis of EDS' and the Census Bureau's data.

When assessing the effects of demographic characteristics while ignoring differences in voting equipment across counties, as in Model 2, we found that the percentage of uncounted presidential votes was significantly higher in counties with smaller populations and in counties with higher percentages of African Americans. Other factors were not statistically significant. The demographic measures we considered, taken together, accounted for about 12 percent of the variability in the percentage of uncounted votes across counties. When we considered voting equipment and demographic factors jointly in Model 3, (1) we were able to account for 18 percent of the variation across counties in the percentage of uncounted presidential votes, and (2) punch card equipment, population size, and percent African American remained statistically significant. That is, regardless of county demographics, counties that used punch card equipment had higher percentages of uncounted presidential votes. Additionally, regardless of voting equipment, counties with higher percentages of African Americans had higher percentages of uncounted votes, and counties with larger populations had lower percentages of uncounted presidential votes.

In our final model, Model 4, we estimated these same effects after allowing not only for clustering but also for differences across counties that were due to the unmeasured effects of the states they are located in. Using dummy variables (the coefficients for which are deleted from table 3) to allow these effects made it possible to account for about 44 percent of the variation in uncounted presidential votes. Moreover, Model 4 indicates that once this full set of differences was accounted for, there were no differences in uncounted presidential votes among counties that use electronic, paper, or optical scan voting equipment. Counties with punch cards had roughly 0.6 percentage points higher percentages of uncounted presidential votes than those counties, and counties with lever equipment had 0.7 percentage points lower percentages of uncounted presidential votes than those counties. Given that the average uncounted votes across all counties was roughly 2 percent, these represent sizable, as well as statistically significant, differences.

The only demographic variables that were associated with significantly higher percentages of uncounted presidential votes when the state and voting equipment effects were controlled, were higher percentages of residents who were African American and Hispanic. The demographic variables that were associated with significantly lower percentages of uncounted presidential votes when the state and voting equipment effects were controlled included higher percentages of high school graduates and 18 to 24 year-olds in the county. Characteristics of voters did not appear to interact with voting equipment to affect the percentage of uncounted votes, although our aggregated data were not well suited to addressing this issue. Models that included interactions between voting equipment and demographic characteristics (not shown) accounted for only about 1 percent of the variation in uncounted votes across counties.

An additional key finding of our study was that differences across states were of considerable importance in determining the prevalence of uncounted presidential votes and accounted for more of the variability across counties in uncounted presidential votes (26 percent) than demographic characteristics (16 percent) and type of voting equipment (2 percent) combined. The following factors for which we had no data are among those that may have contributed to differences among states:

1. voter education efforts, such as making sample ballots available prior to election day;

2. the use of straight party ballots that enable voters to make one entry to cast votes for all offices on the ballot;

3. the number of candidates on the ballot (including presidential, gubernatorial, or congressional candidates);

4. the number of provisional ballots cast, and percentage of provisional ballots that were not counted; and

5. the extent to which absentee and/or early voting occurred and if such ballots were counted using a different voting equipment than ballots cast on election day.

When we ran Model 4 for a subset of 404 counties that GAO surveyed, we found an additional equipment effect. This survey asked county election officials if they used equipment that either prevents errors or identifies errors for voters so the ballot might be corrected. Since both electronic and lever equipment prevent “overvotes,” almost all of the counties using

those types of equipment reported using error correction. In addition, almost all of the counties using punchcard equipment and paper ballots reported not having or using error correction capabilities. Therefore, responses to the survey allowed us to distinguish between counties with optical scan equipment that used error correction and those that did not use it. Doing so resulted in significant differences between types of equipment. Counties using punch cards had significantly higher percentages of uncounted presidential votes than counties using error corrected optical scan equipment by 1.1 percentage points. If the relationship that we found in these 404 counties holds true for the larger set of 2,455 counties, an estimated 300,000 additional votes may have been counted if counties that used punch card equipment had, instead, used precinct-based optical scan equipment with error correction.

Appendix II: GAO Contacts and Staff Acknowledgments

GAO Contacts

Laurie E. Ekstrand (202) 512-8777
Evi L. Rezmovic (202) 512-8777

Staff Acknowledgments

In addition to the above, Wendy Ahmed, Douglas Sloane, David Alexander, Amy Lyon, and Tanya Cruz made key contributions to this report.

GAO's Mission

The General Accounting Office, the investigative arm of Congress, exists to support Congress in meeting its constitutional responsibilities and to help improve the performance and accountability of the federal government for the American people. GAO examines the use of public funds; evaluates federal programs and policies; and provides analyses, recommendations, and other assistance to help Congress make informed oversight, policy, and funding decisions. GAO's commitment to good government is reflected in its core values of accountability, integrity, and reliability.

Obtaining Copies of GAO Reports and Testimony

The fastest and easiest way to obtain copies of GAO documents is through the Internet. GAO's Web site (www.gao.gov) contains abstracts and full-text files of current reports and testimony and an expanding archive of older products. The Web site features a search engine to help you locate documents using key words and phrases. You can print these documents in their entirety, including charts and other graphics.

Each day, GAO issues a list of newly released reports, testimony, and correspondence. GAO posts this list, known as "Today's Reports," on its Web site daily. The list contains links to the full-text document files. To have GAO E-mail this list to you every afternoon, go to our home page and complete the easy-to-use electronic order form found under "To Order GAO Products."

Order by Mail or Phone

The first copy of each printed report is free. Additional copies are $2 each. A check or money order should be made out to the Superintendent of Documents. GAO also accepts VISA and Mastercard. Orders for 100 or more copies mailed to a single address are discounted 25 percent. Orders should be sent to:

U.S. General Accounting Office
P.O. Box 37050
Washington, D.C. 20013

To order by phone:	Voice:	(202) 512-6000
	TDD:	(301) 413-0006
	Fax:	(202) 258-4066

Visit GAO's Document Distribution Center

GAO Building
Room 1100, 700 4th Street, NW (corner of 4th and G Streets, NW)
Washington, D.C. 20013

To Report Fraud, Waste, and Abuse in Federal Programs

Contact:

Web site: www.gao.gov/fraudnet/fraudnet.htm,
E-mail: fraudnet@gao.gov, or
1-800-424-5454 (automated answering system).

Public Affairs

Jeff Nelligan, Managing Director, NelliganJ@gao.gov (202) 512-4800
U.S. General Accounting Office, 441 G. Street NW, Room 7149,
Washington, D.C. 20548

PRINTED ON RECYCLED PAPER