How Does Voting Equipment Affect the Racial Gap in Voided Ballots?

Michael Tomz   Stanford University
Robert P. Van Houweling   University of Michigan

An accumulating body of research suggests that African Americans cast invalid ballots at a higher rate than whites. Our analysis of a unique precinct-level dataset from South Carolina and Louisiana shows that the black-white gap in voided ballots depends crucially on the voting equipment people use. In areas with punch cards or optically scanned ballots, the black-white gap ranged from four to six percentage points. Lever and electronic machines, which prohibit overvoting and make undervoting more transparent and correctible, cut the discrepancy by a factor of ten. Judging from exit polls and opinion surveys, much of the remaining difference could be due to intentional undervoting, which African Americans profess to practice at a slightly higher rate than whites. In any case, the use of appropriate voting technologies can virtually eliminate the black-white disparity in invalid ballots.

An accumulating body of research points to a racial gap in voided ballots. African Americans, several studies suggest, cast invalid presidential votes at a higher rate than whites (e.g., Brady et al. 2001; Herron and Sekhon 2001; Posner 2001). This article investigates whether the black-white gap depends on the voting equipment people use. The bulk of existing evidence on this issue pertains to punch cards and optical scanners, the two technologies that predominated in Florida during the 2000 presidential election. Both technologies apparently admit a wide racial gap, although the difference narrows considerably when ballots are counted at the polling place (e.g., Fessenden 2001; Keating and Mintz 2001; USCCR 2001). The racial consequences of alternative equipment, such as lever and direct-recording electronic (DRE) machines, are not yet known. The few studies that include these technologies reach inconsistent conclusions (HCGR 2001; Knack and Kropf 2001; GAO 2001), and none isolates and compares invalidation rates for blacks and whites. In this article we introduce a unique dataset and employ ecological regression methods that allow, for the first time, a direct assessment of the black-white gap under all major types of voting machines. We find that, unlike centrally counted optical ballots and punch cards, DRE and lever machines nearly eliminate the racial gap in voided ballots.

Our analysis is based on an original dataset, constructed from millions of voter history records in South Carolina and Louisiana. Relative to other datasets, ours is especially appropriate for several reasons. First, it contains accurate measures of turnout by race, our key explanatory variable. Among U.S. states only South Carolina and Louisiana officially report this information, instead of imperfect proxies such as the racial composition of the registered electorate or the general population. Second, our data distinguish between in-person voters, who use the machines in question on election day, and absentee voters, who do not. Third, we aggregate the data by precinct (rather than county or congressional district) to maximize variation on our explanatory and dependent variables, thereby increasing the precision of our estimates. Finally, our dataset covers the main technologies in use nationwide: punch cards, optical scanners, DREs, and lever machines. For these four reasons our

data permit more accurate estimates of black-white differences for a wider range of voting equipment than previously possible. Of course our focus on South Carolina and Louisiana comes at a potential cost: racial patterns of invalidation in these states may differ from those in other parts of the country, an issue we discuss in the conclusion.

The article not only introduces an original dataset, but also departs from other research by employing methods designed to infer individual behavior from aggregate data.[1] Using the quadratic ecological regression model of Achen and Shively (1995), we find that the black-white discrepancy in uncounted ballots is widest (four to six percentage points) in areas with centrally counted optical ballots and punch cards. Lever and DRE machines, which prohibit overvoting and make undervoting more transparent and correctable, cut the discrepancy by a factor of ten, leaving a gap of only 0.3 to 0.7 percentage points. To deepen the analysis we then investigate whether a gap of such small magnitude could be due to racial patterns of *intentional* undervoting. Using a special collection of opinion surveys and exit polls, we find that African Americans profess to undervote at a slightly higher rate than whites, the estimated difference being a few tenths of a percentage point. Thus, DRE and lever machines nearly eliminate the portion of the black-white gap that arises from human error, rather than willful undervoting.

Our analysis proceeds in several steps. In the second section we review existing work on the racial gap and explain why, in theory, DRE and lever machines should narrow the difference in invalidation rates between blacks and whites. The third section describes our research design, data, and methods. The fourth and fifth sections present the results, and the final section concludes the article.

Literature Review and Hypotheses

We now summarize what other research has demonstrated about the racial gap and develop hypotheses to test later. The presentation is organized in three parts. First, we review literature that points to a disparity between black and white invalidation rates and speculates on its causes. Second, we consider how features of common voting equipment could affect the magnitude of the disparity and hypothesize that DRE and lever machines should be associated with smaller black-white gaps than other technologies. Third, we identify the few parallel studies that discuss the relationship between race and invalidation for

[1]Only one other study uses methods for ecological inference, and it focuses entirely on Florida (USCCR 2001).

a wide range of voting technologies. Although valuable, these studies have limitations that our research overcomes.

A Black-White Gap in Voided Ballots Exists

A growing body of evidence suggests that blacks cast invalid ballots more often than whites. In one of the earliest contributions, Hansen (2000) used OLS to analyze precinct-level data from Palm Beach County, Florida, and uncovered a positive relationship between overvoting and the black percentage of registered voters. Others extended this finding to Broward, Duval, and Miami-Dade counties (Herron and Sekhon 2001; USCCR 2001) and reinforced it with cross-county regressions for the entire state (Jewett 2001; Klinkner 2001; Rasmusen 2001; USCCR 2001). Media organizations contributed their own evidence: using data the National Opinion Research Center collected on 175,010 Florida ballots that did not register a valid vote for president, they concluded that rejected ballots were more common in heavily black precincts than in heavily white ones (Fessenden 2001; Keating and Mintz 2001). Multi-state investigations detected a similar relationship between blacks and uncounted votes (Brady et al. 2001, 26, 40; GAO 2001; HCGR 2001; Knack and Kropf 2001).

Researchers have proposed various explanations for these racial patterns. Many contend that blacks are more prone than whites to commit voting errors on average, due to low education, illiteracy, and other socioeconomic factors (e.g., Posner 2001, 80–81). This is undoubtedly part of the explanation, although regression analyses that control for socioeconomic variables still find a positive though attenuated correlation between race and invalidation (e.g., Brady et al. 2001; Fessenden 2001; GAO 2001; Rasmusen 2001; USCCR 2001). Racial patterns of invalidation could also arise if black voters are less experienced than white ones. In Florida, for example, the NAACP sponsored a massive drive that brought many African Americans to the polls for the first time, perhaps contributing to the black-white gap. Research on the link between inexperience and invalidation has yielded mixed results, however (Holt and Berens 2001; Jewett 2001; Klinkner 2001; Knack and Kropf 2001). Finally, some add that blacks may be less likely than whites to seek or receive assistance in using machines and correcting mistakes, especially where there is racial intimidation or a history of disenfranchisement (e.g., USCCR 2001).

If such fundamental factors underlie the black-white gap, one might assume the problem would be difficult to address. On the contrary, we maintain that certain kinds of voting equipment can go a long way toward minimizing the racial difference in voided ballots. We now develop this hypothesis.

How Voting Equipment Could Affect the Gap

Americans use a wide range of voting machines, some with features that could weaken the ability of socioeconomic differences, relative inexperience, or racial antagonism to produce a black-white gap in voided ballots. We systematically describe the major types of voting equipment and explain why lever and DRE machines should be associated with smaller racial disparities in invalidation rates than other widely used technologies.

Studies have shown that punch-card ballots are highly vulnerable to human error. The voter inserts the punch card into a template with candidate names and uses a stylus to dislodge prescored rectangles called chads. A voter may fail to punch the card cleanly, leading to a pregnant, dimpled, or hanging chad that prevents the counting machine from registering a valid vote (Saltman 1988), or could overvote by punching more than one hole for a particular office. Punch-card ballots do not indicate which chad corresponds to which candidate, making it difficult to verify a vote after removing the ballot from the template. Moreover, a voter who spots an error might need to complete a new ballot, an inconvenient step the voter might not take. Probably for these reasons, punch-card machines have the highest rate of invalid presidential ballots in the country, 2.6 percent (Brady et al. 2001, 29).

Optically scanned ballots overcome some of these problems but do not eliminate the possibility of overvoting or undervoting. With optical ballots the voter must use a special writing instrument to fill in bubbles or arrows. It is possible to spoil an optical ballot by failing to blacken the bubble or arrow completely, by using the wrong kind of pencil or pen, or by marking spots next to multiple candidates. Optical technology is more user-friendly than punch cards, though, since the names of candidates appear on the ballot itself, allowing people to verify votes if they take the time. Many counties tabulate optical ballots at central locations, but some have the ability to scan ballots at the precinct and alert people who overvoted or undervoted. Counties do not always activate this warning feature, but where they do voters can decide whether to correct mistakes. In November 2000 only 1.4 percent of ballots in counties with optical scanners contained no valid presidential vote (Brady et al. 2001, 29); the rate was probably higher for centrally counted ballots and lower where ballots were tabulated at the precinct (GAO 2001; Keating and Mintz 2001; Posner 2001; USCCR 2001).

Alternative technologies exist that preclude overvoting and make it easier to verify choices and correct mistakes. With lever machines the names of candidates appear on cards next to levers that the voter pulls to make a se-

lection. These machines usually prevent the voter from switching too many levers for a particular office. Moreover, the voter is free to review and change any choices before finalizing votes. DREs offer a modern analog to levers. These machines display a list of candidates for a particular office, and the voter pushes a button or toggle switch next to the preferred candidate's name. Doing this highlights the name of the candidate so the voter can verify the choice. On DREs a person can change a vote only by deselecting the currently highlighted candidate and then choosing a new one, a two-step procedure that prevents overvoting. As with lever machines, the voter can review all choices before registering them officially. The machines can also signal when someone has not voted for a particular office, thereby reducing the risk of unintentional undervoting.

Thus, both lever and DRE machines prevent overvoting, allow citizens to review their selections, and enable them to correct errors without requesting a new ballot. We might therefore expect these machines to minimize invalid ballots. In fact only 1.7 percent of presidential ballots cast on either type of machine in November 2000 were not counted. Judging from the raw averages, these machines beat punch cards but did slightly worse than optical scanners. Recent studies with demographic and political control variables, however, show that lever and DRE machines perform as well as or better than optical scanners (Brady et al. 2001, 29, 34; GAO 2001; cf. Caltech/MIT 2001). We contribute to the debate not by testing which machines minimize voided ballots overall, but by assessing the differential impact of voting technologies on African Americans and whites.

Given the machine features summarized in Table 1, we advance the following hypothesis: *the black-white gap should be narrower with lever and DRE equipment than with common alternatives*. Our expectation follows

TABLE 1   Relative Strengths and Weaknesses of Voting Technologies

	Potential for accidental overvote	Potential for accidental undervote	Cost of correcting a mistake
Punch Card	High	High	High
Central Optical Scan	Medium	Medium	Medium
Precinct Optical Scan	Low/Medium	Low/Medium	Medium
Lever	Low	Medium	Low
DRE	Low	Low	Low

directly from the logic in this and the previous section. By preventing overvoting and making undervoting more transparent and correctable, lever and DRE machines reduce the influence of fundamental factors—socioeconomic disadvantages, relative inexperience, and racial antagonism—that might lead blacks to make mistakes or fail to correct them more often than whites. Precinct-counted optical ballots share many of these advantages when warning systems are active. Before introducing our own tests of this proposition, we review the few pertinent studies that exist.

Parallel Research on the Effects of Voting Equipment

Three studies, conducted concurrently with ours, discuss the relationship between race and invalidation for a wide range of voting technologies (GAO 2001; HCGR 2001; Knack and Kropf 2001).[2] Although they offer valuable insights, these studies do not isolate and compare invalidation rates for blacks and whites, as would be necessary to evaluate our hypothesis. Moreover, they rely on proxy variables for racial turnout, conflate absentee and in-person voters, and conduct analysis at a high level of aggregation. Finally, for reasons not entirely clear, they reach inconsistent conclusions. As we review these studies we emphasize the challenges of measurement and analysis that make our hypothesis difficult to test. This provides a foundation for the third section, which discusses how our special data and methods help overcome the limitations of the existing literature.

In a brief study the House Committee on Government Reform (HCGR 2001) compared 40 congressional districts, half with high poverty and large minority populations and half with the opposite conditions. It found that poor minority districts had higher percentages of invalid ballots, but "modern voting technologies"—especially DREs and precinct-level optical scanners—appeared to narrow the disparity. A more comprehensive study by Knack and Kropf (2001) examined county-level data from 39 states in 1996. For each type of voting machine the authors regressed invalid presidential ballots on 15 census and structural variables, including the black share of the county population. They detected a positive relationship between African Americans and invalidation, except in

counties where voting equipment could be "programmed to eliminate overvoting." The General Accounting Office (2001) employed similar data and methods but reached a different conclusion. In its regression analysis of county-level observations from 43 states in 2000, demographic "characteristics of voters did not appear to interact with voting equipment to affect the percentage of uncounted votes" (GAO 2001, 19). These conflicting studies do not provide equally detailed accounts of procedures and results, making it hard to say why they point in different directions.[3]

The studies not only disagree, but also do not speak directly to our hypothesis. Simply put, they do not estimate the proportions of blacks and whites that cast invalid ballots under each technology. With those statistics researchers could calculate the black-white gap by machine type; without them it would be impossible to test the central proposition of this article. Regrettably, one cannot infer black and white invalidation rates from the studies cited above. The coefficient on "percent black" in their regressions measures the *marginal* effect of race after controlling for education, poverty, voting experience, and many other variables. Although interesting in its own right, this is quite different from estimating how blacks and whites fare, on average, *given* the characteristics they actually have.[4] We posit that certain technologies can narrow the invalidation gap by weakening the impact of fundamental disparities that exist in the real world. For our purposes, then, it is crucial not to impose artificial equality between blacks and whites through the liberal use of control variables. Instead, we need a statistical model tailor-made to infer, from aggregate election data, how often members of each racial group cast invalid ballots.

Three data problems further limit the ability of existing research to address our hypothesis. First, nearly all work relies on proxy variables such as the black proportion of registered voters or the general population to characterize those who actually went to the polls. When the assumptions for unbiased ecological regression are otherwise satisfied, the use of such proxies introduces

---

[2] Some researchers have looked at a narrower range by comparing punch cards and optical scanners in Florida. The correlation between race and invalidation is substantially weaker where voters use precinct-counted optical ballots than where they use either punch cards or centrally counted optical ballots (Fessenden 2001; Keating and Mintz 2001; USCCR 2001).

[3] The divergent conclusions could have arisen because the studies used different datasets from different years, included different controls, and employed interaction variables in different ways. One cannot be sure, though, without access to their data and computer code.

[4] Also valuable but further removed from our interest in racial invalidation rates, Brady et al. (2001) present a county-level regression in which low education increases the proportion of voided ballots, but not in areas with DREs. Kimball, Owens, and McAndrew (2001) add that, other factors equal, the correlation between race and invalidation is weaker when ballots include a straight-party option.

nonrandom measurement error that depresses estimates of African American and white invalidation, as well as the difference between the two rates.[5] Second, except for a few analyses of Florida, research fails to distinguish between in-person voters, who use the machines in question on election day, and absentee voters, who do not. Special information collected for this study underscores the importance of the distinction. In Louisiana, for example, absentees accounted for only 4 percent of turnout but were responsible for 31 percent of voided presidential ballots in 2000.[6] Finally, work outside Florida involves data at a fairly high level of aggregation. The units of analysis, counties or congressional districts, do not differ as widely in racial turnout or ballot invalidation as smaller geographical units (e.g., precincts). Thus, the regression estimates are less precise than they could be. All three studies summarized above exhibit these data limitations, prompting the authors of one to emphasize that its "results should be interpreted with caution" (GAO 2001, 16). In the next section we explain how our unique data and methods overcome these problems.

Data and Analytic Methods

Our research design explicitly addresses the challenges of measurement and analysis that were highlighted in the previous section. We constructed a dataset that includes the racial breakdown of voter turnout by race, distinguishes between in-person and absentee voters, measures key variables at the precinct level, and covers a wide range of voting technologies. We then employed an extended form of ecological regression to estimate overall invalidation rates for black and white voters under each major type of voting equipment. This allowed us to test our main hypothesis: the black-white gap in voided ballots should be narrower with DRE and lever machines than with punch cards and optical ballots. We now provide details about our data and methods.

The Data

Our dataset contains precinct-level observations from South Carolina and Louisiana, the only states that of-

ficially supply the necessary information. To build the dataset, we merged millions of individual voter history records with demographic data from registration files. This allowed us to determine the race and precinct of origin for every person who cast a ballot in the 2000 election. We also commissioned programmers in the election offices of both states to generate special absentee databases that identified who voted at the precinct polling station and who cast ballots by other means. We then collapsed the individual-level data (1.4 million cases in South Carolina, 1.8 million in Louisiana) to calculate how many whites and nonwhites used voting machines in each precinct on election day. In our two states, 98 to 99 percent of individuals who voted in the 2000 election were either black or white, so our measure of nonwhite turnout essentially reflected the participation of African Americans.

We next computed the proportion of invalid ballots by precinct. The formula for invalidation in each precinct i was $I_i = 1 - P_i / V_i$, where P is the number of valid presidential ballots and V is our tally of voters who appeared in person on election day.[7] After dropping extreme cases of invalidation (below the 1st or above the 99th percentiles) in each state, we matched our precinct-level measures with information about the machines voters used. South Carolina and Louisiana together employ the four most prevalent voting technologies in the United States: punch cards, optical scanners, DREs, and lever machines. Thus, data from these states can shed considerable light on our hypothesis.

It is worth emphasizing that our treatment of absentee voters not only improves measures of racial turnout, but also corrects a common error in invalidation statistics. In publicly available records for Louisiana, South Carolina, and other states an absentee voter is credited as having turned out in the precinct where she resides, but her votes are tabulated in a fictitious county-wide "absentee precinct" that bears no relationship to the voter's precinct of origin. Thus, public data overstate the number of voided ballots in proportion to the number of absentees in each precinct. Moreover, if whites are more likely than minorities to vote absentee, using official turnout statistics would lead researchers to overestimate invalidation by whites. We avoid these problems by reconstructing the data from individual records.

Absentee lists for Louisiana are reliable, since an absentee voter cannot receive a ballot until her name is entered into a computerized poll list. Thus, we were able to strip all 76,765 absentee voters from the precinct-level turnout in Louisiana with a high degree of accuracy. It

---

[5] Our analysis of this phenomenon is available at http://www. stanford.edu/∼tomz/pubs/pubs.shtml. The statistical literature on errors in variables is extensive. See, e.g., the recent survey by Wansbeek and Meijer (2000), as well as more more specific studies on aggregate data by Ansolabehere and Rivers (1995), Grofman, Handley, and Niemi (1992), and King (1997).

[6] As the next section explains, we could not be as precise about turnout and invalidation by absentees in South Carolina.

[7] We calculated P from official election returns compiled by Lublin and Voss (2001).

TABLE 2   Descriptive Statistics for South Carolina and Louisiana

	South Carolina			Louisiana	
	Punch card	Optical	DRE	Lever	DRE
Number of counties	12	13	21	52	12
Number of precincts	658	354	859	2,163	1,640
Invalidation rate (%)	5.6	5.3	3.4	0.5	0.6
Nonwhite turnout (%)	22	33	31	26	39
Poverty rate (%)	14	21	19	25	25
High school education (%)	68	62	69	62	71
Median income ($000s)	31	25	28	23	27

was more difficult to implement the correction in South Carolina, where the list of absentee voters was incomplete. As evidence of this problem, the number of valid absentee votes for president exceeded the number of certified absentee voters in nine of 46 counties in November 2000. Nevertheless, we removed all known absentees (102,878, a figure larger than the 98,125 valid absentee ballots statewide), and thus significantly reduced the bias in official counts of in-person turnout.

The descriptive statistics for our dataset appear in Table 2. The key variables for our analysis are invalidation and nonwhite turnout, measured at the precinct level. Compared with county-level statistics, our dataset contains substantially more variation on the racial variable. The standard deviation around the nonwhite proportion of voters in our dataset was 0.33 in Louisiana and 0.27 in South Carolina, whereas the statistic for U.S. counties (using nonwhite population as a proxy for nonwhite turnout) was only 0.16. According to 1996 Census figures, nonwhites make up more than 95 percent of the population in only one of 3,138 U.S. counties. In contrast, nonwhites represent more than 95 percent of voter turnout in 326 Louisiana precincts (8 percent of the total) and 43 South Carolina precincts (2 percent of the total). This confirms our expectation about the advantages of precinct-level data. Table 2 also includes measures of poverty, education, and income, which pertain to the total population rather than actual voters and are measured at the county rather than the precinct level (U.S. Census Bureau 1998).

The table reveals a diversity of voting equipment and socioeconomic characteristics in South Carolina. In the November 2000 election, 21 South Carolina counties employed DREs, 12 relied on punch card voting systems, and 13 used optically scanned ballots—10 counted at a central location and three counted at the polling station. The average share of voided ballots under the first two systems (5.6 percent and 5.3 percent) exceeded the share

under DREs (3.4 percent) by a noticeable margin. The table also shows that counties with optical scanners had more poverty, lower education and income, and larger minority populations than the rest of the state. DREs occupied an intermediate position on most of these dimensions. Overall, minority voters in South Carolina are not disproportionately exposed to inferior technologies.

Table 2 also summarizes the Louisiana data. Only 12 of 64 parishes employed DREs, but the list includes the five most populous parishes, making the distribution by precinct somewhat less lopsided. Residents of parishes with DREs had higher levels of income and education, but they were also more likely to be African Americans. The most striking feature in the Louisiana dataset is the miniscule proportion of invalid presidential ballots, compared not only with South Carolina but also with the national average of 2 percent in November 2000. In fact, Louisiana can claim one of the lowest recorded invalidation rates in the country, whereas South Carolina shows one of the highest. To some extent this difference arises from the way the two states track voter turnout.[8]

---

[8]In constructing the data, we discovered that states often undercount the number of in-person voters. On election day, Louisiana voters sign poll lists to document that they have turned out. After the election, clerks key in the data but tend not to enter the turnout of all voters, a problem Louisiana officials acknowledge. Consequently, the number of valid in-person votes for president exceeded the number of recorded voters in nearly 11 percent of the precincts in November 2000. Invalidation may be abnormally low in Louisiana precisely because official data understate turnout.

We eliminated the most extreme instances of negative invalidation by dropping cases below the 1st percentile. Nevertheless, the Louisiana data still contained 378 precincts with more votes than voters. We retained these districts to avoid selecting on the dependent variable, but our conclusions did not change when we dropped them from the dataset temporarily. South Carolina keeps better track of in-person voters. Before the voter enters the booth, poll workers fill in bubbles on a computer-generated, machine-scannable list. Even this procedure is not error free, though: 5.5 percent of South Carolina precincts reported fewer voters than presidential votes in November 2000. As in Louisiana, we preserved

Ecological Regression Methods

Having constructed a new dataset, we used an extended form of ecological regression to estimate invalidation rates for white and nonwhite voters. The fraction of invalid presidential ballots in any particular precinct is given by the accounting identity $I = \beta_n T_n + \beta_w (1 - T_n)$, where I is invalid ballots as a proportion of all ballots, β_n is the nonwhite invalidation rate, β_w is the white invalidation rate, T_n is the nonwhite share of total turnout and $(1 - T_n)$ is the white share. Recall that nearly all nonwhites in South Carolina and Louisiana are black, so in our analysis β_n essentially represents the invalidation rate of African Americans. Rearranging these terms gives $I = \beta_w + (\beta_n - \beta_w) T_n$. If the parameters β_n and β_w are constant across precincts, we can estimate them via a Goodman's regression of I on T_n and a constant term (Goodman 1953).

It seems unrealistic, however, to assume that β_n and β_w do not change with the racial composition of districts. In predominantly white precincts, voters of all races tend to be wealthier and better educated than those who live in minority areas. Consequently, β_n and β_w should be somewhat lower in white areas than in minority ones. The standard Goodman's regression does not allow this variation and thus is vulnerable to problems of aggregation bias, which can occur when the parameters of interest (in this case the invalidation rates) are correlated with the regressors.

Our statistical approach rests on work of Achen and Shively (1995), who derive an extended version of Goodman's regression and prove that it helps mitigate the effects of aggregation bias. We relax the assumption that the parameters are mean-independent of the regressors and instead model the invalidation rates in each district i as linear functions of T_{ni}. Let $\beta_{ni} = b_1 + b_2 T_{ni} + e_{ni}$ and $\beta_{wi} = b_3 + b_4 T_{ni} + e_{wi}$. By substitution into Goodman's regression we obtain

$$I_i = \beta_0 + \beta_1 T_{ni} + \beta_2 T_{ni}^2 + \varepsilon_i,$$

where $\beta_0 = b_3$, $\beta_1 = b_1 - b_3 + b_4$, and $\beta_2 = b_2 - b_4$. The coefficients of this quadratic equation can be estimated by regressing I_i on T_{ni}, T_{ni}^2, and a constant term.[9]

To convert the estimates of β_0, β_1, and β_2 into measures of b_1, b_2, b_3, and b_4, we needed to make an assumption. Only b_3, the intercept for whites, is identified by the quadratic regression, but we calculate values for the other bs by assuming that either b_2 or b_4 is zero. This assumption makes our model less restrictive than the basic Goodman's regression, which constrains *both* b_2 and b_4 to be zero and thus does not allow the invalidation rate of either group to vary with the composition of the precinct. In cases like ours, the added flexibility from letting even one invalidation rate change with T_n reduces aggregation error relative to the basic specification (Achen and Shively 1995).

Our criteria for identification are straightforward. Note that β_2 represents the difference between b_2 and b_4. Suppose that the proportion of uncounted ballots increases—or at least does not decrease—for both racial groups as precincts become less white. If the estimate of β_2 is positive, we conclude that $b_2 > b_4$, which means the invalidation rate among nonwhites increases more quickly than the invalidation rate among whites as T_{ni} rises. Consequently, we set the smaller of these values (b_4) to zero, allowing us to solve for b_1 and b_2. In this example, white invalidation is held constant across precincts, while nonwhite invalidation is allowed to vary with the racial composition of turnout. If instead the estimate of β_2 is negative, implying that $b_4 > b_2$, we identify the system of equations by setting $b_2 = 0$, thereby fixing nonwhite invalidation at a constant level while letting the white rate fluctuate from one precinct to the next.

To see whether the racial gap in voided ballots depends on voting equipment, we divided the sample by state and machine type, resulting in five datasets (three for South Carolina, two for Louisiana) with precinct-level observations. We then ran a quadratic regression for each dataset and used it to estimate invalidation rates for blacks and whites. The next section describes our results.

Results

In line with our hypothesis, the black-white gap in voided ballots was substantially lower with DRE and lever machines than with punch cards and optical scanners. We present the results in various ways. First, we offer a table of ecological regression coefficients and standard errors. Second, we graph the estimated relationships between nonwhites as a proportion of total turnout and

---

the negative-spoilage precincts to avoid selection on the dependent variable. (Results were qualitatively similar when we deleted the negative-spoilage precincts, though the regression lines flattened slightly for punch cards and optical scanners, precisely as we would expect after selecting on the dependent variable.) Overall, South Carolina accounts for most in-person voters, which helps explain why the reported invalidation is higher there than in Louisiana.

[9] The heteroskedasticity in the error term $\varepsilon_i = e_{wi} + (e_{ni} - e_{wi}) T_{ni} + v_i$ does not bias the point estimates, though it does affect

their variance, an issue we address later in the article by computing heteroskedastic-consistent standard errors.

TABLE 3   Regression Results (with standard errors) for South Carolina and Louisiana

|	South Carolina			Louisiana	
	Punch card	Optical	DRE	Lever	DRE
Proportion Nonwhite	.138	.147	.064	.0060	.0007
	(.014)	(.029)	(.011)	(.0022)	(.0022)
Proportion Nonwhite^2	−.086	−.069	−.051	.0020	.0067
	(.017)	(.033)	(.013)	(.0024)	(.0022)
Constant	.035	.017	.023	.0029	.0036
	(.002)	(.005)	(.002)	(.0003)	(.0003)
N	657	354	858	2162	1639
R^2	.28	.24	.06	.07	.12

invalidation as a proportion of all ballots. Third, we use the regression results and computer simulation to calculate the gap (with confidence intervals) between black and white invalidation rates under each type of voting equipment. This not only conveys the substantive meaning of our regression results but also offers the most direct evidence concerning our hypothesis. Finally, we consider alternative specifications and find that they do not change our conclusions.

Regression results appear in Table 3, with standard errors in parentheses. These coefficients in themselves are not the primary quantities of interest, but they do suggest two patterns. First, the quadratic term is statistically distinguishable from zero (coefficient more than twice its standard error) in all but one column, giving strong reason to prefer the extended model over the standard ecological regression, which omits this term. Second, the

R^2 statistics are much larger for punch card and optically scanned ballots, suggesting that race explains a larger share of ballot invalidation in areas that use those technologies. Far from a cause for concern, the low R^2 values for DRE and lever machines accord with our hypothesis that racial patterns of invalidation should be very weak with such equipment. To better understand the substantive importance of these results, we turn to graphical analysis and computer simulation.

Figure 1 contains regression lines with confidence intervals for the three types of voting equipment in South Carolina. The horizontal axes measure nonwhites as a proportion of all in-person voters and the vertical axes give invalidation rates. The plots show that the average invalidation rate increases with the nonwhite share of total turnout in punch card and optical scanner precincts, but there is little relationship between these variables in precincts that

FIGURE 1   Regression Lines and 95% Confidence Intervals for South Carolina
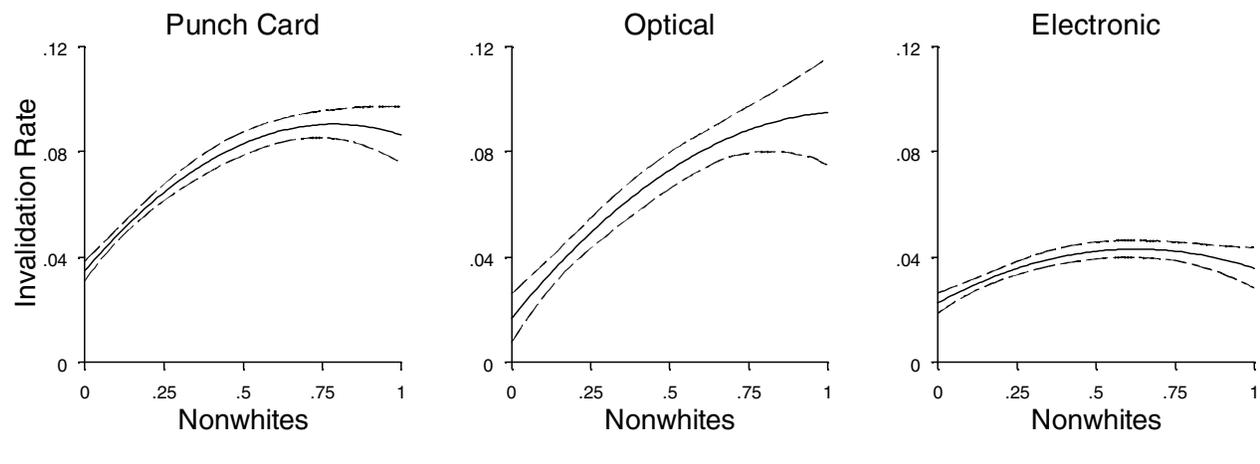
use DREs.[10] We omit the graphs for Louisiana, where the regression lines have only a slight upward slope.

Finally, we employed the following algorithm to obtain our key quantities of interest: overall estimates of nonwhite and white invalidation rates, as well as the difference between the two.

1. Randomly draw one set of parameters $\tilde{\beta}_0, \tilde{\beta}_1, \tilde{\beta}_2$ from their asymptotic sampling distribution, $N(\hat{\beta}, V(\hat{\beta}))$. The tilde over the parameters indicates that they were simulated.
2. Use the simulated parameters and the identifying assumption discussed in the third section to compute values for $\tilde{b}_1, \tilde{b}_2, \tilde{b}_3,$ and $\tilde{b}_4$. (Recall that $\tilde{b}_1 = \tilde{\beta}_1 + \tilde{b}_3 - \tilde{b}_4, \tilde{b}_2 = \tilde{\beta}_2 + \tilde{b}_4, \tilde{b}_3 = \tilde{\beta}_0,$ and $\tilde{b}_4 = \tilde{b}_2 - \tilde{\beta}_2$. We identify the system by assuming a value of zero for either b_2 or b_4.)
3. Given those values and the observed level of T_{ni}, calculate $\tilde{\beta}_{ni} = \tilde{b}_1 + \tilde{b}_2 T_{ni}$ and $\tilde{\beta}_{wi} = \tilde{b}_3 + \tilde{b}_4 T_{ni}$ for each precinct i.
4. Compute a turnout-weighted average of the precinct-level estimates, following the procedure in (Achen and Shively 1995, 121–23). The formulas in this case are $\tilde{\beta}_n = \sum_i \tilde{\beta}_{ni} V_{ni} / V_n$ and $\tilde{\beta}_w = \sum_i \tilde{\beta}_{wi} V_{wi} / V_w$, where V_{ni} and V_{wi} are the number of nonwhite and white voters who turned out in precinct i, and V_n and V_w measure the number of nonwhite and white voters in the entire sample. This produces one simulation of β_n and one simulation of β_w.
5. Calculate $\tilde{\beta}_n - \tilde{\beta}_w$, the racial gap in uncounted ballots.

By repeating this algorithm 1000 times, we obtained 1000 simulated values of $\tilde{\beta}_n - \tilde{\beta}_w$, thereby approximating the sampling distribution of the racial difference in invalidation rates. One could also compute the standard error around the estimated gap analytically, but simulation is more convenient and no less accurate in this case (King, Tomz, and Wittenberg 2000).

We now provide a numerical example, for punch cards in South Carolina, of how this algorithm generates estimates of the black-white gap. The regression estimated an intercept of $\hat{\beta}_o = .035$, a coefficient on non-white turnout of $\hat{\beta}_1 = .138$, and a coefficient on the quadratic term of $\hat{\beta}_2 = -.086$. These estimates have a 3x3 variance matrix with diagonal elements equal to the standard errors in Table 3. In step 1 we draw one set

of betas from a multivariate normal distribution with mean equal to the vector of estimated coefficients and variance equal to their variance-covariance matrix. Suppose, by chance, we draw exactly the coefficients we estimated, i.e. $\tilde{\beta}_0 = .035, \tilde{\beta}_1 = .138, \tilde{\beta}_2 = -.086$. In step 2 we note that the quadratic coefficient is negative and identify the model by assuming $\tilde{b}_2 = 0$. This allows us to calculate $\tilde{b}_1 = .138 + .035 - .086 = .087, \tilde{b}_3 = .035,$ and $\tilde{b}_4 = .086$. Step 3 requires the nonwhite share of voter turnout in precinct i. Consider the average punch-card precinct, where nonwhites made up 22 percent of those who went to the polls. For this precinct we estimate a nonwhite invalidation rate of $\tilde{\beta}_{ni} = .087 + 0 \times .22 = .087$ and $\tilde{\beta}_{wi} = .035 + .086 \times .22 = .054$, for a racial gap of $.087 - .054 = .033$ or 3.3 percentage points. Our sample contains many precincts, though, whose values for nonwhite turnout may be higher or lower than the mean. To complete step 3 we compute $\tilde{\beta}_{ni}$ and $\tilde{\beta}_{wi}$ for each of the 657 punch-card precincts, using their actual levels of nonwhite turnout. In step 4 we calculate a weighted average of the 657 values for nonwhite invalidation, with weights proportional to the number of nonwhite voters in the precinct, and make a parallel calculation for white voters. Given the distribution of nonwhite turnout in our data, the weighted averages are $\tilde{\beta}_n = .087$ and $\tilde{\beta}_w = .045$, implying in step 5 an overall black-white gap of $8.7 - 4.5 = 4.2$ percentage points. This estimate is uncertain because it relies on regression coefficients that are themselves uncertain. The simulation procedure, which repeatedly draws from the sampling distribution of the coefficients, accounts for the uncertainty and allows us to compute confidence intervals around our point estimate for the racial gap.

Table 4 summarizes our estimates of the racial gap in voided ballots for each type of voting equipment. The columns give invalidation rates (with 95 percent confidence intervals) for nonwhites and whites, as well as the difference between the two. In South Carolina, the estimated racial gap is 4.2 percentage points in precincts that use punch cards and 6.2 percentage points in those with optical scanning devices. Confidence intervals around these estimates (2.9–5.4 and 3.0–9.3) are far from zero, suggesting that the racial gap for these two machines probably did not arise by chance. On the other hand, we estimate a difference of only 0.3 percentage points (confidence interval −0.7–1.4) between nonwhite and white invalidation rates in precincts with DREs. For Louisiana, the regressions suggest only a mild relationship between invalidation and race. The values in Table 4 reflect this. We find a racial gap of only 0.7 percentage points for lever machines and 0.5 percentage points for DREs. Neither confidence interval (0.5–0.9, 0.4–0.7) includes zero

[10] For each type of machine in South Carolina, our estimate of the quadratic term was negative and highly certain. Consequently, we identified the system of equations by treating the minority invalidation rate as constant across districts, while allowing the white invalidation rate to change. In Louisiana the coefficients on the quadratic terms were positive, leading us to hold white invalidation constant.

TABLE 4 The Estimated Racial Gap in Voided Ballots

	Nonwhite		White		Gap	
	Mean	95% CI	Mean	95% CI	Mean	95% CI
South Carolina						
Punch card	8.7	7.6 to 9.7	4.5	4.2 to 4.9	4.2	2.9 to 5.4
Optical scan	9.5	7.3 to 11.6	3.3	2.1 to 4.4	6.2	3.0 to 9.3
DRE	3.6	2.8 to 4.4	3.2	2.9 to 3.6	0.3	−0.7 to 1.4
Louisiana						
Lever	1.0	0.8 to 1.2	0.3	0.2 to 0.3	0.7	0.5 to 0.9
DRE	0.9	0.8 to 1.0	0.4	0.3 to 0.4	0.5	0.4 to 0.7

FIGURE 2 The Racial Gap: Estimates, Interquartile Ranges, and 95% Confidence Intervals
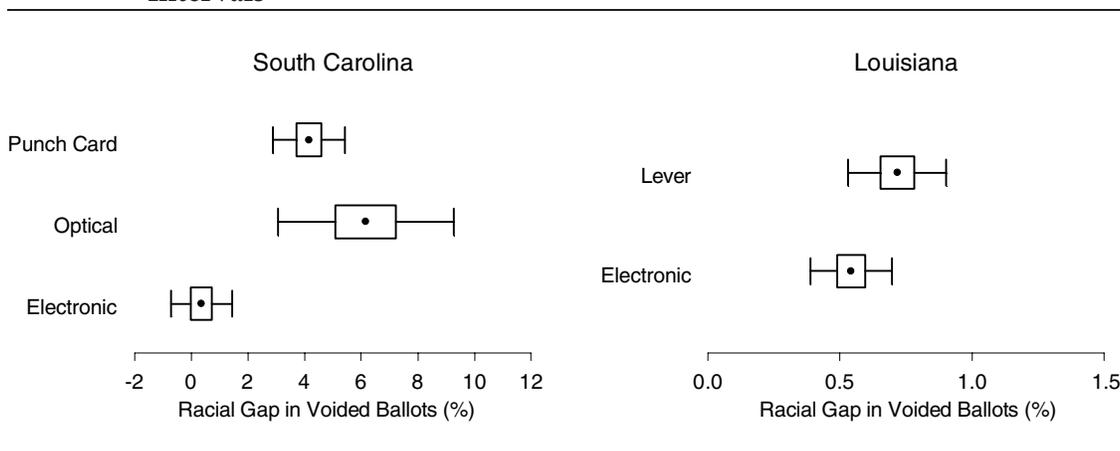
but the two intervals overlap, so we cannot conclude that either of these machine types leads to a smaller racial difference than the other.

We also convey these results graphically in Figure 2. The vertical axis indicates the type of voting equipment and the horizontal axis measures the *difference* in invalidation rates between nonwhite and white voters. In each customized boxplot, the central dot represents our best estimate of the racial difference, the box covers the interquartile range of simulated values, and the "wings" span a 95 percent confidence interval around the estimate. The plots show that the estimated gap is much smaller with DREs and lever machines than with punch cards and optical ballots. They also convey that the difference is statistically significant: the confidence intervals for the two underperforming technologies do not overlap either the point estimates or the confidence intervals for DREs and levers.[11]

Both technologies in Louisiana were associated with small racial gaps. We attribute this to the equipment voters used, rather than other factors peculiar to that state. To confirm this interpretation we analyzed the invalidation rates of Louisiana's absentee voters, who used punch cards. For reasons discussed earlier invalidation rates for absentees are available only at the parish level, but with 64 parishes we had enough observations for a preliminary analysis.[12] A quadratic regression finds that 10.3 percent of nonwhites and 5.3 percent of whites invalidated their absentee ballots, for a racial gap of 5 percentage points. One should not place too much weight on this point estimate, which has a standard error of 5.9 due partly to the small number of observations. Nevertheless, it is

(95 percent confidence interval 2.5–9.1) with optical scanners and was 3.9 percentage points larger (95 percent confidence interval 2.1–5.5) with punch cards. Thus, the difference in machine performance almost certainly did not arise by chance.

[12]One of the 64 parishes reported more valid absentee ballots than absentee voters, implying a spoilage rate of −18 percent. We dropped this precinct from the analysis.

[11]Consider, for example, the plot for South Carolina. Compared with DREs, the estimated racial gap was 5.9 percentage points larger

consistent with precinct-level results for punch cards in South Carolina, which increases confidence that voting equipment is responsible for the narrow racial gaps among in-person voters in Louisiana.

To shed additional light on the performance of optical technologies, we considered whether the racial gap depended on where officials counted the ballots: at the polling station or at a central location. Of the South Carolina counties that used optical equipment in November 2000, the vast majority employed central counting but three scanned at the polling place. The estimated gap of 7.1 points (confidence interval 2.2–11.9) in those three counties did not differ substantially from the gap of 7.9 points (confidence interval 1.2–14.5) in areas that tabulated centrally. Further analysis revealed heterogeneity among the precinct-tabulating counties, though. Orangeburg County exhibited a large racial gap, but in Beaufort and Laurens there was no meaningful difference between our estimates of nonwhite and white invalidation rates. The inferior preformance in Orangeburg may have arisen because it employed a somewhat different warning procedure and ballot design.[13] Overall, our data do not point to a clear conclusion about the effects of precinct counting. Other studies with different data find that precinct counting weakens the relationship between African Americans and invalidation (HCGR 2001; Keating and Mintz 2001; Knack and Kropf 2001; USCCR 2001). More research is needed to evaluate the impact of this procedure on racial groups.

The results in this article are fairly stable across alternative specifications. Our conclusions did not change when we included a cubic term T_{ni}^3, and they became a bit stronger but harder to interpret when we implemented log and logit transformations of the dependent variable.[14] To deal with potential heteroskedasticity in the quadratic regression, we reestimated the models using White's standard errors. This widened the confidence intervals around the racial gap for punch cards and optical ballots (2.2–6.2 and 2.2–10.1) but had no effect on the intervals for lever machines and DREs, and thus did not alter our substantive conclusions.

Finally, we estimated a model that allowed invalidation rates to vary with socioeconomic conditions, rather than the percentage of nonwhite voters in the precinct. The model used socioeconomic variables for a different purpose than other studies. Regressions by Knack and Kropf (2001) and the GAO (2001) estimate the *marginal* effect of race after holding poverty, education, and other factors constant. The second section explained why this approach does not allow direct inferences about the invalidation rates of blacks and whites, our key quantities of interest. We recognize, however, that the invalidation rates of these groups may vary across precincts and are probably higher in minority areas than elsewhere. To address this possibility our primary analysis let invalidation rates change with the nonwhite share of voters in the precinct. Here we explore an alternative: let invalidation rates of blacks and whites depend on average levels of poverty and education. This is simply another way to address the potential for aggregation bias, which could arise if black and white inhabitants of the same precinct have socioeconomic commonalities. Unlike other studies, which use controls to equalize (in a statistical sense) the socioeconomic characteristics of areas that people inhabit, we emphasize the racial implications of those socioeconomic conditions and impound them into estimates of invalidation.

To do this, we defined $\beta_{ni} = b_1 + b_2\text{Poverty}_i + b_3\text{Education}_i + e_{ni}$ and constructed an analogous equation for β_{wi}. This led us to regress I_i on a constant, T_{ni}, poverty, education, and interactions between the socioeconomic measures and T_{ni}. Our measures of poverty and education contain error—they refer to the overall population rather than people who actually voted, and they are recorded at the county rather than the precinct level—but are the best available data. As before, we ran the regression and drew the parameters from their sampling distribution. We then used *actual values* of T_{ni}, poverty, and education to obtain simulated values of $\tilde{\beta}_{ni}$ and $\tilde{\beta}_{wi}$ for each precinct i. Finally we computed a weighted average of the precinct-level estimates and calculated the racial gap.

Consistent with our previous results, the difference was much wider with punch cards and optical ballots than with lever machines and DREs. In South Carolina the estimated racial gaps, with 95 percent confidence intervals in parentheses, were 6.6 (5.7–7.6) percentage points for punch cards and 6.9 (5.4–8.3) for optical ballots,

---

[13] Machines in Orangeburg warned people that the ballot contained an overvote, but equipment in Beaufort and Laurens went further by identifying the specific race where the mistake occurred. The ballot in Orangeburg required voters to connect arrows; in the other two counties voters filled bubbles on the ballot.

[14] We also tried dropping the quadratic term, thereby estimating a standard Goodman's regression. This led to higher estimates of the black-white gap on all types of equipment, which is consistent with our expectation that, without confronting aggregation bias, the gap may appear larger than it actually is. These results underscore the value of the more flexible quadratic regression, which is less vulnerable to aggregation bias. With the Goodman model, the point estimates for the black-white gap in South Carolina were 6.7 percentage points for punch cards, 8.9 for optical scanners, and 2.1 for DREs. In Louisiana they were 0.8 and 0.7 for levers and DREs, respectively. A bivariate local regression of invalidation on the nonwhite share of total turnout produced a curve that very closely matched our quadratic model.

compared with only 2.1 (1.4–2.7) for DREs. Although the estimate for DREs exceeds what we found with the quadratic model, it remains consistent with our principal conclusion. In Louisiana the estimates were 0.8 (0.6–0.9) and 0.8 (0.7–0.9) for lever machines and DREs, respectively.[15] Across a variety of specifications, then, the estimated racial gap was narrower with lever machines and DREs than with the alternatives.

Is the Remaining Gap Intentional?

In the previous section we showed that DRE and lever machines reduce the black-white gap to a fraction of a percentage point. In this section we deepen the analysis by exploring whether a gap of that magnitude could be due to racial patterns of *intentional* undervoting, rather than human error. Intentional undervoting occurs when a citizen goes to the polls but opts not to select a presidential candidate. Using opinion surveys and exit polls, we find that intentional undervoting is more prevalent among African Americans than among whites. The difference in rates of intentional undervoting is a few tenths of a percentage point, the order of magnitude of the estimated racial gap with lever machines and DREs. This implies that even the best voting equipment cannot eliminate the racial gap in uncounted ballots, but DRE and lever machines can eliminate most of the gap that arises from human error.

Our analysis begins with the American National Election Study, a nationwide survey of the U.S. electorate in presidential and midterm election years. In every presidential election year since 1964, the National Election Study has asked respondents whether they voted and, if so, whether they voted for a candidate for president.[16] We pooled the data from all but one of the NES surveys, producing a sample of 11,521 voters across nearly four decades.[17] All but 309 of these individuals were either white or African American, so we restricted our analysis to those two racial groups.

[15]Results were very similar when we substituted income for poverty.

[16]Of people who claimed to vote, the interviewer asked "How about the election for President? Did you vote for a candidate for President?" The NES recorded whether the respondent answered Yes, No, or DK.

[17]The 1972 survey did not contain enough information to distinguish intentional undervoters from those who did not know or did not remember whether they voted for president. The final sample sizes for black and white voters were 1290 in 1964, 1109 in 1968, 1320 in 1976, 994 in 1980, 1432 in 1984, 1202 in 1988, 1644 in 1992, 1143 in 1996, and 1078 in 2000. Data came from National Election Studies (2000).

TABLE 5   Intentional Undervoting by Race (NES, 1964–2000)

	Black		White	
	%	N	%	N
Voted for president	99.18	1,328	99.42	9,816
Didn't vote for president	0.82	11	0.58	57

Table 5 presents a cross-tabulation of race and intentional undervoting. White respondents claimed to undervote intentionally at slightly lower rate than blacks: 0.58 percent (0.44–0.75) versus 0.82 percent (.41–1.47), with 95 percent binomial confidence intervals in parentheses. We used simulation to assess whether the difference between these two rates was statistically significant.[18] Given the NES data, the probability that blacks undervote at a higher rate than whites is 0.88 corresponding to a *p*-value of 0.12 for a one-sided significance test. Even with the relatively small NES sample, then, we can be fairly confident that blacks profess to undervote more frequently than whites.[19]

Exit polls lend more certainty to the conclusion that blacks choose to undervote at a somewhat higher rate than whites. In November 1992 the Voter News Service (VNS) asked a sample of voters, as they left the polls, which presidential candidate they had just voted for. The possible responses were Bill Clinton, George Bush, Ross Perot, Other, or "Didn't vote for president."[20] We combined Clinton, Bush, Perot, and Other into a new category

[18]For both racial groups the probability of intentional undervoting is close to zero. Under these conditions the normal approximation to the binomial distribution is poor, even with a large sample, making classical significance tests such as the chi-square inappropriate (Agresti 1992). We adopted a Bayesian approach. Specifically, drew random variates from two independent Beta distributions, $\pi_{black}|$ data $\sim$ beta(12, 1329) and $\pi_{white}|$data $\sim$ beta(58, 9817), the posterior densities for black and white undervoting rates given the NES data and a uniform prior (Johnson and Albert 1999, 11). Drawing two numbers at a time, one from the black posterior and one from the white, we computed the Bayesian *p*-value as the relative frequency with which the draw of π_{white} exceeded the draw of π_{black}.

[19]Research indicates that African Americans are more likey to overreport turnout than whites, controlling for other socioecomic differences (see, e.g., Deufel and Kedar 2000; Silver, Anderson, and Abramson 1986). Restricting attention to validated NES voters did not change our substantive conclusion: the gap was 0.77 − 0.53 = 0.24 percentage points, but cell sizes were very small. It is also possible that, even among genuine voters, African Americans disproportionately overreport having cast a ballot for president. If so, our estimates provide a lower bound on the racial gap in intentional undervoting.

[20]The original sample included 54,806 validated voters, extracted from Voter Research and Surveys (1993). We listwise-deleted 732

TABLE 6   Intentional Undervoting by Race (VNS Exit Poll, 1992)

	Black		White		Hispanic		Asian	
	%	N	%	N	%	N	%	N
Voted for president	99.60	6511	99.80	44,400	99.55	1,764	99.68	617
Didn't vote for president	0.40	26	0.20	91	0.45	8	0.32	2

called "voted for president." The results, tabulated by race, appear in Table 6. The VNS data suggest that blacks undervote intentionally at approximately twice the rate of whites—0.4 percent (0.31–0.58) versus 0.2 percent (0.16–0.25), again with 95 percent binomial confidence intervals in parentheses.[21] Although the absolute magnitude of the gap is small, the *p*-value for a one-sided significance test is 0.002, leaving us very certain that blacks report to undervote at a higher rate than whites.[22] Table 6 also suggests that, in 1992, Hispanic voters were most likely to refrain from voting for president, and that Asian respondents undervoted deliberately more often than whites but less frequently than blacks. The small samples for Hispanic and Asian voters prevent us from endorsing these conclusions with a high degree of certainty, however.

Given the infrequency of intentional undervoting and the limited sources of data about it, one should draw conclusions with care. But the best available information suggests that intentional undervoting adds approximately 0.2 (VNS) or 0.3 (NES) percentage points to the black-white gap in voided ballots. This difference is too small to explain discrepancies of several percentage points that we find under punch card and optical machines. Nevertheless, intentional undervoting could account for much of the racial gap in areas that use lever machines, DREs, and perhaps precinct-counted optical ballots. Put another way, those technologies do a remarkable job of eliminat-

ing the portion of the gap that arises from human error, rather than conscious choice.

Conclusion

The black-white gap in voided ballots depends crucially on the voting equipment that people use. Based on data from the 2000 presidential elections in South Carolina and Louisiana, we estimated a gap of around 4 percentage points in precincts that used punch cards and approximately 6 percentage points in areas that used optically scanned ballots. DRE and lever machines cut the discrepancy by a factor of ten, leaving a gap of only 0.3–0.7 percentage points.[23] We find that much of the remaining difference could be due to intentional undervoting, which African Americans profess to practice at a slightly higher rate than whites.

We arrived at these conclusions by constructing an original dataset and analyzing it with ecological regression techniques. The dataset, drawn from South Carolina and Louisiana, exhibits several features that make it especially appropriate for research on the racial consequences of voting equipment. It contains direct measures of turnout by race, distinguishes between absentee and in-person voters, measures variables at the precinct level, and covers a wide variety of voting technologies. We analyzed the data with quadratic ecological regression, leading to the first estimates of the difference in invalidation rates of blacks and whites across all major machine types.

Our focus on South Carolina and Louisiana offers clear benefits, but it also comes at a potential cost: racial patterns of invalidation in those two states may differ from patterns in other parts of the country. To some extent the generality of our results depends on the mechanism behind the black-white gap in

(1.3 percent) of these due to missing data on race or the presidential vote and further eliminated 655 cases in which respondents placed themselves in a racial category labeled "other." VNS asked similar questions after the 1996 and 2000 presidential elections, but their coding did not allow us to identify people who professed not to vote for president.

[21] Our estimates of intentional undervoting are smaller with VNS data than with NES data. We thank an anonymous reviewer for noting that the difference could be due to the very different ways in which the surveys are administered.

[22] As with the NES data, we drew random variates from two independent Beta distributions, in this case $\pi_{black}|\text{data} \sim \text{beta}(27, 6512)$ and $\pi_{white}|\text{data} \sim \text{beta}(92, 44441)$, and computed the Bayesian *p*-value as the relative frequency with which the draw of π_{white} exceeded the draw of π_{black}.

[23] Recall that the 95 percent confidence intervals around these estimates are 2.9–5.4 percentage points for punch cards, 3.0–9.3 for optically scanned ballots, 0.5–0.9 for lever machines, −0.7–1.4 for DREs in South Carolina, and 0.5–0.9 for DREs in Louisiana.

uncounted ballots.[24] If hostility to minority voting contributes to the racial gap, and if such attitudes are especially prevalent in South Carolina and Louisiana, then punch card and optical machines in other states may be associated with smaller black-white differences than we found. Likewise, if socioeconomic differences underlie the racial gap on punch cards and optical scanners, then we must consider the racial incidence of variables like poverty and education. Where African Americans and whites are more socioeconomically equal than in South Carolina and Louisiana, punch cards and optical scanners should pose less of a disadvantage for black voters. In fact, socioeconomic discrepancies between blacks and whites in South Carolina and Louisiana are similar to those in other southern states (including Florida) and certain northern states (e.g., Michigan, Illinois, Wisconsin) but wider than the national mean.[25] Finally, if the relative inexperience of black voters contributes to a racial gap in voided ballots, one must be sensitive to changes in experience across regions and elections. For example, a large percentage of blacks in Florida went to the polls for the first time in 2000, in response to an unprecedented get-out-the-vote drive by the NAACP. The atypical infusion of inexperienced black voters in Florida, where scholars have focused the most attention, makes it all the more informative to consider other states, as well, including the ones in this article.

Remarkably, we find that DRE and lever machines nearly eliminate the difference between black and white invalidation rates, even in two states where socioeconomic and attitudinal factors may be less favorable than average. If anything, South Carolina and Louisiana present hard cases for DREs and levers, yet those technologies go a considerable way toward leveling the playing field for black voters. Nevertheless, it is important to view our research in the context of parallel studies on other parts of the country. Our data and methods offer certain advantages, which we have enumerated, but they are also restricted in scope. Thus, it is encouraging to know that other research, focusing on Florida, points in a similar direction. Punch cards and optical scanners, the technologies that performed worst in our study, also generated wide gaps there, but counting ballots at the polling place reduced the gap, just as our reasoning would suggest. The combined weight of studies on Florida and across the nation reinforce our conclusion: the choice of voting technology affects the black-white gap in voided ballots.

Although we find that lever machines, DREs, and perhaps precinct-level optical scanners can narrow the gap in invalidation rates between African Americans and whites, such machines may not be best for democracy or most appropriate for the United States. Scholars and policymakers should weigh many factors—in addition to the differential effects of voting equipment on racial groups—as they work to upgrade the nation's voting systems. Cost may enter the calculation, since devices that narrow the racial disparity could be expensive. Moreover, one could explore the potential tradeoff between speed and accuracy: lever machines, DREs, and precinct-level scanners may create long lines and affect voter turnout, offsetting their advantages on other dimensions. For now, though, we know that the choice of voting equipment can have a significant impact on the racial gap in voided ballots.

---

[24]Scholars would need a different research design, with different data, to measure the relative importance of each possible cause of the black-white gap that we identify in the second section.

[25]In 1989, the black-white gap in poverty rates was 23 percentage points in SC and 32 in LA, compared with 20 when considering all individuals nationwide. The gap in high school degree attainment was 20 percentage points in SC, 21 in LA, and 15 for all individuals. Calculated from the 1990 U.S. Census.

References

Achen, Christopher H., and W. Phillips Shively. 1995. *Cross-Level Inference*. Chicago: University of Chicago Press.

Agresti, Alan. 1992. "A Survey of Exact Inference for Contingency Tables." *Statistical Science* 7(1):131–53.

Ansolabehere, Stephen, and Douglas Rivers. 1995. "Bias in Ecological Regression Estimates." Working Paper. Massachusetts Institute of Technology and Stanford University.

Brady, Henry E., Justin Buchler, Matt Jarvis, and John McNulty. 2001. *Counting All the Votes: The Performance of Voting Technology in the United States*. Berkeley: University of California.

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

Deufel, Benjamin J., and Orit Kedar. 2000. "But Do They Really Vote? Correcting for Overreporting of Turnout." Working Paper. Harvard University.

Fessenden, Ford. 2001. "Ballots Cast by Blacks and Older Voters Were Tossed in Far Greater Numbers." *New York Times*, November 12, A17.

GAO (United States General Accounting Office). 2001. *Statistical Analysis of Factors That Affected Uncounted Votes in the 2000 Presidential Election*. Washington, DC.

Goodman, Leo. 1953. "Ecological Regressions and the Behavior of Individuals." *American Sociological Review* 18(6):663–64.

Grofman, Bernard, Lisa Handley, and Richard G. Niemi. 1992. "Defining and Measuring Racially Polarized Voting and Other Elements of the Totality of the Circumstances." In *Minority Representation and the Quest for Voting Equality*, ed. Grofman et al. New York: Cambridge University Press.

Hansen, Bruce E. 2000. "A Precinct-Level Demographic Analysis of Double-Punching in the Palm Beach Presidential Vote." Working Paper. University of Wisconsin.

Herron, Michael C., and Jasjeet S. Sekhon. 2001. "Overvoting and Representation: An Examination of Overvoted Presidential Ballots in Broward and Miami-Dade Counties." *Electoral Studies*. Forthcoming.

Holt, Douglas, and Michael J. Berens. 2001. "State Worst in Ballot Errors," *Chicago Tribune*, April 29.

HCGR (House Committee on Government Reform). 2001. *Income and Racial Disparities in the Undercount in the 2000 Presidential Election*. Washington, DC.

Jewett, Aubrey. 2001. "Explaining Variation in Ballot Invalidation Among Florida Counties in the 2000 Election." Working Paper. University of Central Florida.

Johnson, Valen E., and James H. Albert. 1999. *Ordinal Data Modeling*. New York: Springer-Verlag.

Keating, Dan, and John Mintz. 2001. "Florida Black Ballots Affected Most in 2000; Uncounted Votes Common, Survey Finds." *Washington Post*, November 13, A3.

Kimball, David C., Chris Owens, and Katherine McAndrew. 2001. "Who's Afraid of an Undervote?" Working Paper. University of Missouri, St. Louis.

King, Gary. 1997. *A Solution to the Ecological Inference Problem*. Princeton: Princeton University Press.

King, Gary, Michael Tomz, and Jason Wittenberg. 2000. "Making the Most of Statistical Results: Improving Interpretation and Presentation." *American Journal of Political Science* 44(2):347–61.

Klinkner, Philip A. 2001. "Whose Votes Don't Count? An Analysis of Spoiled Ballots in the 2000 Florida Election." Working Paper. Hamilton College.

Knack, Stephen, and Martha Kropf. 2001. "Invalidated Ballots in the 1996 Presidential Election: A Cross-Level Analysis." Working Paper. University of Maryland and University of Missouri, Kansas City.

Lublin, David, and D. Stephen Voss. 2001. "Federal Elections Project." American University and the University of Kentucky.

National Election Studies, Center for Political Studies, University of Michigan. 2000. Electronic resources from the NES World Wide Web site (http://www.umich.edu/~nes). Ann Arbor, MI: University of Michigan, Center for Political Studies [producer and distributor], 1995–2000.

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

Rasmusen, Eric. 2001. "The Pattern of Spoiled Ballots by County in Florida." Working Paper. Indiana University.

Saltman, Roy. 1988. *Accuracy, Integrity and Security in Computerized Vote-Tallying*. NBS SP 500-158. Gaithersburg, MD: NIST.

Silver, Brian D., Barbara A. Anderson, and Paul R. Abramson. 1986. "Who Overreports Voting?" *American Political Science Review* 80(2):613–24.

USCCR (U.S. Commission on Civil Rights). 2001. *Voting Irregularities in Florida During the 2000 Presidential Election*. Washington, DC.

US Census Bureau. 1998. *USA Counties 1998: A Statistical Abstract Supplement*. Washington, DC: Census Bureau.

Voter Research and Surveys. 1993. *Voter Research and Surveys General Election Exit Polls, 1992* [Computer file]. New York City, NY: Voter Research and Surveys [producer], 1992. 2nd release. Ann Arbor, MI: Inter-University Consortium for Political and Social Research [distributor], 1993.

Wansbeek, Tom, and Erik Meijer. 2000. *Measurement Error and Latent Variables in Econometrics*. Amsterdam: Elsevier.