Counting Ballots and the 2000 Election: What Went Wrong?

R. Michael Alvarez†
D.E. "Betsy" Sinclair‡
Catherine H. Wilson⋆

February 9, 2004

† Associate Professor of Political Science, Division of Humanities and Social Sciences, California Institute of Technology and corresponding author (rma@hss.caltech.edu); ‡ Research Assistant, Division of Humanities and Social Sciences, California Institute of Technology and Department of Mathematics, University of Redlands (betsy@hss.caltech.edu). ⋆ Ph.D. Candidate, Division of Humanities and Social Sciences, California Institute of Technology (chw@hss.caltech.edu). We thank Tara Butterfield for their assistance with the data used in this paper; Mary King Sikora for her administrative help; Conny McCormack and John Mott Smith for insights and information; and Steve Ansolabehere and Melanie Goodrich for their comments. This research was supported by the Caltech/MIT Voting Technology Project with a research grant from the Carnegie Corporation; however the conclusions reached here reflect the views of the authors and not necessarily the Caltech/MIT Voting Technology Project. This research was also supported by a grant to Alvarez from the IBM Corporation through the University Matching Grants Program.

1 Introduction

Imagine a high school, on the day that a comprehensive test is administered to assess the school's performance relative to statewide standards. Some students answer their tests using pencil and paper, some using a standard "fill in the bubble" form, some with a computer card where they "punch out" their answers, and others using computer terminals. After the test, the school principal announces that because of errors made by the students and the tabulation machines, some of the student tests were not tabulated. Specifically, she announces that while 3% of the "punch out" tests were not tabulated, only 1.5% of the other tests were not scored because of errors. That some of the student tests were not tabulated seems clearly unfair to students, but because twice as many tests were not tabulated due to the particular way in which the students took the test school administrators, parents, students, and people throughout the community would be deeply troubled about the fairness of this testing process.

But imagine also a scenario where the students using each type of answer sheet are descriptively different. That is, many more high school students with purple hair use the "punch out" answer system, while students with orange hair use the other test answering systems. Or, in another scenario many of the students with purple hair had never taken a "punch out" style test before, or for other reasons found the test style confusing, complicated, and difficult to use; each of these problems producing many more uncounted tests for students with purple hair than students with orange hair, even when students with orange hair used the "punch out" system. These scenarios would be quite troubling, because clear allegations that students with purple hair were disadvantaged in this testing situation would be difficult to discount.

To many observers, the deadlocked 2000 presidential election seems strikingly similar to these scenarios. Many analyses of the 2000 presidential election have shown that certain voting systems are associated with higher rates of uncounted ballots than other voting systems.[1] In particular, studies conducted by Caltech/MIT (2001), the GAO (2001),

[1]In this study we use the term "voting system" to refer to the method used by each county in the polling place for voters to mark their ballots. In the next section we discuss in more detail the various types of

the U.S. House Committee on Government Reform (2001), and various scholars (see especially Ansolabehere [2001], Knack and Kropf [2001] and Posner [2001]) have all demonstrated with a variety of databases and statistical methodologies that punchcard voting systems clearly have high rates of uncounted ballots, typically higher than the other technologies.[2] This recent work, moreover, essentially reaffirms the conclusions of a large body of earlier studies, focusing on transitions from paper ballots to punchcards or other types of voting systems decades ago (White 1960; Mather 1964; Thomas 1968; Fraser 1985), as well as more recent research focusing on the transition to electronic voting machines (Shocket, Heighberger and Brown 1992; Nichols and Strizek 1995).

Furthermore, many observers have also worried that some voters might be more likely than others to cast ballots that are not counted. In this regard, most of these same studies agree: uncounted ballots are cast more in areas with high concentrations of nonwhite, poor, and poorly educated residents (GAO 2001; U.S. House Committee on Government Reform 2001; Herron and Sekhon 2001; Knack and Kropf 2001; Posner 2001). This same general pattern was confirmed in media analyses of two large, and largely nonwhite, counties (Cook County, Illinois and Fulton County, Georgia), as well as in a prominent media consortium's study of the uncounted ballots in the Florida 2000 presidential eleciton.[3] The same patterns have been found in earlier academic studies, from earlier elections and many different geographic locations (Walker 1963; Walker 1966; Clubb and Traugott 1972; Vanderleeuw and Engstrom 1987; Darcy and Schneider 1989; Bullock and Dunn 1996; Nichols and Strizek 1996).

However, researchers disagree about whether nonwhite, poor, and poorly educated voters have higher uncounted ballot rates when they are in counties employing voting

voting systems and their useage in both the United States and California.

[2]Disagreement exists, though, over whether punchcard voting systems have the highest rate of uncounted ballots than all other systems. For example, the Caltech/MIT report found that electronic voting equipment had roughly similar rates of residual votes as punchcard systems; the GAO analysis of similar data did not find that electronic voting systems had such high levels of uncounted ballots. All these recent studies agree, though, with the basic point that punchcard voting systems are associated with high rates of uncounted ballots.

[3]"A Racial Gap in Voided Votes", *The Washington Post*, December 27, 2000.; "Bush Still Had Votes to Win in a Recount, Study Finds", *The Los Angeles Times*, November 12, 2001 (http://www.latimes.com/news/politics/la-111201recount.story).

systems with higher rates of uncounted ballots — especially punchcard voting systems. Studies here tend to find little support for higher rates of uncounted ballots for nonwhite voters using punchcard voting systems (Ansolebehere 2001). Among the earlier studies, Montgomery (1985) reports null results; more recently the GAO (2001) report on voting systems and uncounted ballots found no relationship between voting systems, race, and uncounted ballots.

The statistical analyses that have shown higher rates of uncounted ballots where punchcard voting systems are used, especially the Votomatic system, have led to calls for the elimination or phasing out of punchcard voting. The more troubling allegations that indicate the possibility that nonwhite voters might be more likely to both use punchcard voting systems and may be more likely to cast uncounted ballots when using a punchcard system have spawned the filing of a series of cases arguing that inequities in voting systems violate Section 2 of the Voting Rights Act and the Fourteenth Amendment to the United States Constitution. Cases making these voting rights claims have been filed in Georgia (*Andrews v. Cox*), Florida (*NAACP v. Harris*), Illinois (*Black v. McGuffage*), and California (*Common Cause v. Jones*).

We begin this chapter with an examination of the problems that were uncovered in studies of the 2000 election. We focus on the concept of "lost votes", and we discuss the different places where voters went missing in this most recent presidential election. We then turn to a case study of elections and election administration data from the 2000 presidential election in the State of California. California was selected as the case for this analysis for four reasons. First, elections data, including over and under vote information, as well as elections administration data, are easily available for California's 58 counties. Second, as we discuss in the next section, California counties use punchcard, optical scan, and touchscreen voting systems — and there is thus a wide diversity in the voting systems currently in use in California. Third, California's political, social, and economic diversity allows for important testing of hypotheses relating to the determinants of uncounted ballot rates across California counties. Fourth, we are interested in studying this question in California due to the ongoing debate about the use of punchcard voting systems and their eventual decertification in California. Our analysis of the California elections data highlights places where significant election reform are necessary, and we discuss this in our conclusion.

2 The 2000 Presidential Election

The 2000 presidential election spawned a focus on voting systems that was unprecedented in American history; never before had the public, press, nor politicians focused on the wide variety of ways in which Americans cast their ballot. In general, there are six basic types of voting systems that are currently used in the United States. The oldest voting system is the simple paper ballot — voters simply make their mark on a piece of paper, which is counted by hand. Lever machines, large mechanical devices that let voters indicate their preferences by throwing levers and that record votes in a mechanical fashion, are the second type of voting system now used in America. Punchcards, a voting system introduced about 30 years ago, where voters indicate their preferences by punching out "chads" in computer-readable cards, are another voting system used today in the United States. Optical scanning, introduced in the past two decades, allows voters to connect broken lines or fill in ovals to indicate their vote; importantly, optical scanning machines can be used either in a central location to tabulate votes or they can be used in each precinct to scan ballots for errors and tabulate locally. Electronics, either in the form of "direct-recording electronic" or the newer "touchscreen" devices, have been used in the past decade or so. Last, some election jurisdictions use a combination of these voting systems, and hence, are called "mixed" voting systems.

Immediately in the aftermath of the Florida vote recount, many groups initiated studies of the 2000 presidential election and issued many policy proposals. One of the most unique research groups was the Caltech/MIT Voting Technology Project, a collaborative effort between the social science and technology faculties on both university campuses. The Caltech/MIT project was given the immediate challenge by their respective university presidents to identify the technological challenges associated with counting ballots in American elections, and to determine ways to fix whatever problems were discovered.

Thus, we began by gathering as much data on voting systems, their use, and measures of the rate at which each system rejected or had ballots that were not counted. We issued a report on March 30, 2001, titled "Residual Votes Attributable to Technology: An Assessment of the Reliability of Existing Voting Equipment." This report generated enormous publicity and widespread interest, largely due to the fact that it was the first systematic study of voting systems conducted in the wake of the 2000 elections — and because this study produced some very shocking conclusions.

Our report focused solely on what we called the "residual vote rate"; which is the total number of uncounted ballots. Ballots may not be counted for at least three reasons; first, a voter may not indicate a preference on the ballot (an undervote); second, a voter may make more indications of a preference than are allowed (an overvote); third, the voter may mark their ballot in a way that is uncountable. We focused on this measure because it is the only yardstick we could devise to compare voting systems across counties, systems, and elections.

What we found in the data we collected ranging from 1988 through 2000 was two clusters of voting systems. One cluster, paper ballots, lever machines, and optically scanned ballots, were shown to have the lowest rate of residual votes throughout this period. On both average and median residual vote rate, these three voting systems had lower residual vote rates than punchcard and electronic voting equipment. The first cluster averaged residual vote rates of around 23

We were not terribly surprised to find that punchcard systems had high residual vote rates; initial reports from Florida, and our examination of some of the previous academic research on this topic led us to hypothesize that punchcard systems would have high residual vote rates. What was surprising was that electronic voting systems, the so-called "direct recording devices" and newer ATM-style voting systems, had residual vote rates roughly comparable to those of punchcard systems. These differences, moreover, were robust to multivariate statistical analysis.

Why did the electronic systems fare so poorly in our study? There are a number of reasons. First, many of the electronic systems that have been used since 1988 have poor user interfaces and bad ballot designs. Second, as electronic systems have only been used in recent elections, there is likely to be a technology learning curve and a corresponding reduction in residual vote rates over time (and there was some evidence of such a technology learning curve in the data we provide in that report). Third, there might be a similar voter learning curve, as they adapt to the new voting technology. Fourth, it simply might be true that electronic voting systems require more administration in the polling places, and thus are potentially more error-prone than other voting systems. Fifth, the electronic technology might be more sensitive to maintenance and reliability problems than the other mechanical technology. And sixth, there might be simple differences in how humans interact with electronic versus mechanical technologies.

While the heterogeneity in voting system use — and the vast differences that clearly can be attributed to different voting systems in the rates of uncounted ballots across counties and states — was the original focus of the Caltech/MIT Voting Technology Project, as we dug deeper into the problems surrounding elections in the United States we found that this was only a piece of the puzzle. As we held workshops, attended public forums and conferences, interviewed elections administrators, and researched the entire process of election administration, we became aware of the fact that in any election, many votes go uncounted for reasons other than voting system problems.

These uncounted votes we termed "lost votes", and the concept of "lost votes" figured prominently in the report we issued in July 2001, *Voting: What Is, What Could Be.* Based on research conducted using a variety of sources, we found that at least 4 to 6 million votes were lost in the 2000 presidential election. Approximately 1.5 to 2 million voters were lost due to voting system problems: faulty voting equipment, confusing ballots, or other problems with voting technology. But more votes, between 1.5 and 3 million, were lost due to problems in the voter registration system: voters who showed up to vote, but who were denied the opportunity to cast a ballot because of some type of error in their voter registration. Last we found that up to a million voters were lost due to problems in polling places, in particular, mix ups of polling place locations and long lines on election day.

Thus, while the Caltech/MIT Voting Technology Project was initiated to study and fix problems associated with voting systems technology, we quickly learned that voting systems technologies were only part of the problem with elections in the United States. Fixing the problems with voting systems technologies, while obviously important, also would require fixing voter registration systems, polling place practices, and various issues associated with election administration practices. The next phase of this project is devoted to more detailed, and in-depth, analyses of voting systems problems, including the nalysis we turn to in the next section of this chapter. How does one large, and politically powerful state (California) compare fare in terms of uncounted ballots in the 2000 presidential election? If there are differences in uncounted ballot rates across technologies and counties, do these differences impact some voters more and other voters less? In particular, how do minority voters interact with the more error-prone technology, punchcard voting systems?

3 Voting System Use and Residual Rates

We begin our empirical study by discussing the relative use of different voting systems in the United States, and in California. This sets the stage for our later analysis of the relative rates of uncounted ballots for each voting system and how those are correlated with the racial complexion of each county in California.

The relative use of the major voting systems in both the United States and California is given in Table 1. There we provide in the first two columns the distribution of voting systems across the United States, first expressed in terms of the percentages of counties using each system and then in terms of the percentage of the U.S. population using each system.[4] The remaining three columns in Table 1 give similar information for California.[5]

Table 1 Goes Here

[4] These figures were compiled and reported in the Caltech-MIT Voting Technology Project's report (2001).

[5] This data was obtained from the California Secretary of State's Office, http://www.ss.ca.gov/elections/elections_w.htm.

In terms of counties, Table 1 shows that the most prevalent voting system in the United States is optical scan (unfortunately, the national data obtained from the Caltech-MIT report do not break down optical scan systems into precinct- or centrally-based systems), as just over 40% of counties in 2000 use some type of optical scan system. Next in use are punch card systems, with 19% of counties using punchcards; most of these counties are using the Votomatic system according to this data (17.5%). Almost 15% of U.S. counties use lever machines, about 13% use paper ballots, while 9% use direct-recording electronic or touchscreens.

The picture changes at the national level, however, when we turn to the percentages of the national population that use each type of voting system. In population terms the most widely used voting system is punchcards, used by 34% of the U.S. population. Optical scans are second, used by almost 28% of the population. Lever machines are the system used by almost 18% of the population, with direct-recording electronic or touchscreen systems being used by almost 11% of the population. In population terms, paper ballots are hardly used, as only 1.3% of the U.S. population uses paper ballots.

The last three columns of Table 1 give the same distribution statistics for California. Punchcards are more prevalent in California than in the United States, both in terms of the percentage of counties that employ punchcard voting systems (52%) and in terms of the population in California using punchcards (74%). In both county and population terms, punchcard use is over twice as extensive in California than in the United States. In California, furthermore, the Datavote punchcard system is most widespread in county terms, while the Votomatic system is the most widespread in population terms.

Optical scanning voting systems are currently employed by almost 47% of California counties, and by 21% of the state's population. Almost 28% of counties (16 counties in total) use a precinct-based optical scan system, while 19% of counties (11 in total) use a centrally-based optical scanning system. In population terms, precinct-based scanning is more prevalent than centrally-scanned ballots in California.

One county in California — Riverside — uses an electronic, touchscreen voting system. Thus, 4.6% of California's population uses this type of system. Unlike the rest of the

United States, paper, lever and mixed voting systems are not employed in California.

Next we turn to an analysis of the apparent differences across voting systems in their relative rates of uncounted ballots. One way to examine the uncounted ballot rate is to study the "residual" vote rate, which is simply the difference between the number of ballots cast (usually for top-of-the-ticket races) and the number of ballots counted; this for example is the measure used in the Caltech/MIT report (2001). The second way to study uncounted ballots is to look at the rate of ballot spoilage (Alvarez, Butterfield and Wilson 2001); unfortunately, ballot spoilage rates are not widely available. Third, one can decompose the uncounted ballot rates into the over- and undervote rates, where overvotes are ballots that have more marks than allowed for a particular race and undervotes are ballots that have no discernable mark for a particular race.[6] The under- and overvote rates, unfortunately, are not universally available; we have the under- and overvote rates for 43 of the 58 California counties.[7]

Table 2 Goes Here

The first two columns of Table 2 give baseline data for the United States, from the Caltech/MIT report (2001). First we provide the national average residual vote, by the major voting systems, from 1988 through 2000. As noted by the Caltech/MIT report, there are two clusters of voting systems, ones with high residual vote rates (punchcards

[6]There exists substantial debate in the popular discussions of uncounted ballots over the extent to which they are the result of intentional voter behavior or are the result of mistakes or voting machine failure. There is little academic work on the question of the intentionality of uncounted ballots. In a recent unpublished study, Knack and Kropf (2001b) focus on exit poll data and find that about 0.75% of exit poll respondents report not voting at the top of the ticket; thus, they conclude that only about one third of uncounted ballots are due to intentional behavior. In an early study of West Germany, Steifbold (1965) discusses intentional voiding of ballots in West German elections as a result of political protest. Steifbold gives examples of elections where intentional voiding seems to have occurred in West German elections, but he does not attempt to produce a universal estimate of intentional voiding of ballots in West German elections. The extent of intentional under- and overvoting is clearly a question for future research.

[7]We do not have over- and undervote data from Alpine, Imperial, Inyo, Mendocino, Napa, Orange, Placer, San Benito, San Francisco, Santa Barbara, Shasta, Sierra, Sutter, Trinity, or Tuolumne Counties. We do have the overall uncounted ballot rate for all 58 counties.

and electronic), and ones with lower residual vote rates (optical scan, paper and lever). When we turn to the same national data, but presented for only the 2000 election, the same basic clustering again is present. Only in 2000 we see that the level of residual vote rates fell for every voting system.

The residual vote rate, undervotes and overvotes are reported for California counties in the remaining three columns of Table 2. In 2000, California counties using punchcard systems recorded higher rates of residual votes (1.9% for the three punchcard systems in use in California) than counties using optical scan or electronic, touchscreen systems. In particular, the Votomatic system stands out in the residual vote rates, with a statewide 2.4% residual vote rate, compared to 1.8% for Pollstar and 0.8% for Datavote. Overall, counties using optical scanning systems had residual vote rates at least half that of the counties using punchcard systems; also, precinct-based optical scanning counties had slightly lower residual vote rates than centrally-based optical scanning counties. These rates are consistent with the national figures from the Caltech/MIT report.

The last two columns of Table 2 give under- and overvote rates, by voting systems, in California. Undervote rates are high for punchcard systems (2.3%), but this is driven primarily by Votomatic (3.5%) and Pollstar (1.5%). Optical scanning has much lower undervote rates, averaging 0.4% across the two implementations of optical scanning. Precinct-based optical scanning has about half the rate of undervoting as central optical scanning (0.3% relative to 0.6%, respectively). Touchscreen voting, in Riverside county, had a low rate of undervotes in 2000, with only 0.4%.[8] Overvote rates are much lower than undervote rates, and punchcard systems have slightly higher overvote rates (0.3%) than optical scanning systems (0.2%). Within the set of punchcard systems, though, again Votomatic and Pollstar have high overvote rates, and Datavote has a much lower overvote rate. Also, precinct-based optical scanning has very low overvote rates (0.1%), with centrally-scanned ballots having a higher overvote rate (0.2%).[9] Of course, the touchscreen system

[8]The difference between the residual vote rate of 0.9% for Riverside County, and the 0.4% undervote rate, is that the residual vote is computed for all ballots counted, which in Riverside includes optically scanned absentee ballots. The undervotes are solely for precinct voters, and in Riverside County, precinct voters all used touchscreen systems.

[9]As most of the counties that employ precinct-based optical scannning employ error-checking in the

in Riverside County does not allow overvoting, and thus the overvote rate in Riverside County is zero.

Next, we turn to the question of race and uncounted ballots in California. We begin by examining the percentages of nonwhites, Blacks, Asians, and Hispanics in California in the counties using the various voting systems. The race data we use here, and in the subsequent analyses, was obtained from the U.S. Census Bureau, and measures the number of persons of each racial category.[10] We present in Table 3 the percentages of non-white, Black, Asian and Hispanic Californians in the counties using each voting system in Table 3.

Table 3 Goes Here

The first column of data in Table 3 gives the percentages of the non-white population in California counties using the various voting systems. Counties that use punchcard systems have greater populations of non-white residents (43%) than counties that use optical scanning systems (34%). Also, Riverside County's non-white population is almost identical to the figure for optical scanning counties, 34%. Furthermore, notice that counties using the Votomatic system, which according to Table 2 produces the highest residual vote, and the highest rate of under- and overvoting in California, have high non-white populations, almost 48%. Counties using the other punchcard system with relatively high rates

precincts, note that overvoting in precinct-based optical scanning counties might indicate intentional overvoting. Of the counties who respondent to our query for information about how their precinct-based optical scanning systems were configured in the 2000 presidential election, two (Humboldt and Amador) did not enable error-checking in their precinct scanning systems. Placer County stated that their system rejected only undervotes, or blank ballots. Of the remaining counties that provided us information (only San Luis Obispo and San Francisco did not respond to our requests for information), their systems indicated overvoting on ballots in the precincts.

[10]This data comes from the "Profiles of General Demographic Characteristics, California" publication of the U.S. Census Bureau. The county-level measures of race we use are for the entire population, not the voting-age nor the population of registered voters. The latter quantities are not measured by the Census, and could only be estimated from these data. We prefer to focus on the known population statistics, rather than use estimated values for voter registration or voter turnout by racial category. To produce estimates of voter registration or turnout for multiple racial categories would necessitate the use of multi-stage ecological inference estimation, which would be difficult if not impossible, in this context.

of uncounted ballots, Pollstar, also have relatively high populations of non-whites (42%). Interestingly, when we look at the differences within the optical scanning category, we see that counties using precinct-based optical scanning have greater non-white populations than counties that use centrally-located optical scanning systems.

The next three columns of Table 3 provide detailed breakdowns of the non-white populations into Blacks, Asians and Hispanics. Beginning with Blacks, we see a high population of Blacks in counties using the various punchcard systems: 7.3% of the population in those counties is Black. By way of comparison, the Black population in counties using optical scanning systems is 4.5%, and electronic touchscreen systems 6.2%. We also see that within both punchcard and optical scanning categories, the populations of counties using the systems with higher rates of uncounted votes have greater Black populations: 9.5% in counties using Votomatic and 7.0% in counties using Pollstar; likewise, 5.1% in centrally-scanned counties and 4.1% in precinct-scanned counties.

We also see in Table 3 that Asians and Hispanics are also strongly represented in counties using voting systems with higher rates of uncounted ballots. 12% of the population of punchcard counties is Asian, with higher concentrations of Asians in counties using Votomotic and Pollstar. 10% of the population of counties using optical scanning is Asian, but here we see that the Asian population of counties using precinct-based scanning is greater than the Asian population in counties using central scanning. A low percentage of Asians are located in California's one county using electronic voting equipment (almost 4% in Riverside County).

Last, 34% of the population in counties using punchcard systems is Hispanic, while 25% of the population in counties using optical scanning is Hispanic. The Hispanic population is high in all of the counties using punchcard systems, 37% for Votomatic counties, 33% for Datavote counties, and 28% for Pollstar counties. As for Asians, we see a slightly greater representation of Hispanics in counties using precinct-based scanning rather than central scanning. And last, 36% of Riverside county's population is Hispanic; thus the one county in California where touchscreen voting is used has a strong Hispanic population in 2000.

Next, we look directly at the relationship between the racial composition of a county and the rate of uncounted ballots, over- and undervoting. We begin by a simple examination of the bivariate correlation between the percentage nonwhite population in each county and the residual vote, the overvote rate, and the undervote rate.[11] Then we turn to a multivariate statistical analysis of our data, to provide separate confirmation of the relationship between racial composition of a county and uncounted ballots in the county.

In Table 4 we give the bivariate correlation between the nonwhite population and the residual, over- and undervotes, broken down by the voting system used by the county. The correlation coefficient ranges from one to negative one; a positive correlation close to one indicates a strong positive relationship between the percent nonwhite and the uncounted ballot rate, while a correlation close to zero indicates no real relationship between the two variables. Thus, a strong positive correlation for a particular voting system indicates that as the nonwhite population within counties using that voting system increases, so do the uncounted ballot rates.

Table 4 Goes Here

The first column of Table 4 provides the correlations between the nonwhite population and the residual vote rate, for the different voting systems in use in the 2000 presidential election in California. First note the correlations between nonwhite population and residual vote rates for punchcard counties relative to optically scanned counties: moderately strong and positive in the case of the punchcard counties (0.44) while weakly negative in optically scanned counties (-0.15). When we look at the specific types of punchcard

[11] Our focus in the statistical analyses that follow is on nonwhites, rather than the more specific categories of Blacks, non-white Hispanics, and Asians. First, the broader category of nonwhites is traditionally the category that has been considered the legally protected group. Second, in these data, where we only have a limited number of observations, statistical analyses may become very sensitive to specifications that many highly correlated measures like the relative sizes of these different racial groups. However, to the extent that, for example, Blacks and non-white Hispanics to not have similar patterns of political behavior, our decision to include both in one measure of the nonwhite population could bias our estimates for the impact of the nonwhite population on uncounted ballots.

systems, we see positive but modest correlations across the three different punchcard systems. The correlations for precinct- and centrally-scanned counties are negative; but it is important to point out that the correlation is much closer to zero in the case of precinct-based optical scanning than centrally-based optical scanning.

The second column of Table 4 provides the correlations for overvotes. Here we see correlations between the nonwhite population and overvotes that are very close to zero for all punchcard and all optically scanned counties; however the positive correlation for punchcard counties and the negative correlation for optically scanned counties is in the same direction as the residual vote rate. But the correlations between nonwhite population and overvoting are exceptionally strong in the case of Votomatic and Pollstar, but weak for Datavote. Oddly, we see a positive correlation between nonwhite population and overvoting in precinct-based optical scan counties, but a negative correlation in centrally-scanned counties.

The third column of Table 4 presents the same correlation analysis, but now for undervotes. Comparing the correlations for all punchcard and all optically scanned counties, we see a strong positive correlation in the case of punchcard counties, but a weakly positive one for optically scanned counties. We see a strong positive correlation for Votomatic, a weakly positive one for Pollstar, and a weakly negative one for Datavote. While the overall correlation between the nonwhite population and undervotes is weakly positive across all optically scanned counties, we do see negative correlations when we break the two optical scan systems into central- and precinct-based optical scanning.

Thus, this simple data analysis, examining the bivariate correlations between the nonwhite population, the rate of uncounted ballots, the rate of overvoting and undervoting, has revealed what we see as three important conclusions. First, we uniformly see a positive correlation between the nonwhite population and each measure of uncounted ballots across punchcard counties. Second, we see that in particular the Votomatic punchcard system, and to some extent the Pollstar system, have uniformly positive and usually strong correlations between the nonwhite population and each measure of uncounted ballots. Third we see a general pattern for optically scanning that indicates a negative correlation between nonwhite population and uncounted ballot rates for counties using both types of optical scanning system.

However, the results in Table 4 are simple bivariate correlations. While the results are indicative of important trends in the data, they are by no means conclusive. That is because they suffer from a series of questionable assumptions. Most importantly, we have only examined bivariate correlations so far, and have made no attempt to control for any important and potentially intervening variables, most importantly income and education. In the next section we utilize a multivariate statistical analysis to examine the relationship between the nonwhite population and uncounted ballots, with a special emphasis on punchcard voting systems.

4 Multivariate Analysis

Unfortunately, the bivariate analysis did not allow for the estimation of possibly intervening effects.[12] For example, it is plausible to hypothesize that what appeared to be a correlation between race and uncounted ballots really is the result of education or income differences between whites and nonwhites in California. To allow for the possibility of intervening variables, we turn to a multivariate statistical analysis.[13]

We gathered data on other demographic and economic attributes of each California county. In particular, in addition to the data from the U.S. Census Bureau on race from the 2000 Census, we obtained measures of the median age of each county, the amount of money each county spent on primary and secondary public education, and the county's unemployment rate. These data came from the California Department of Finance (2001). Including these variables in a multivariate statistical analysis with measures for white and nonwhite populations, and variables for whether the county was a punchcard county or not, and whether the county employed local precinct counting of ballots or not, should

[12]Furthermore, there is also the question of aggregation bias and whether we are making incorrect ecological inferences (King 1997). In other work using this same data, we conducted an ecological analysis of these data. We do not find results that are qualitatively different from those reported here.

[13]Multivariate statistical analysis uses several descriptive variables in a statistical model to explain the variation in the dependent variable. Multivariate analysis is particularily useful because it allows the user to study the impact of one descriptive or independent variable on the dependent variable, holding all of the other descriptive variables constant. For more discussion, see Tufte (1974).

give us a strong test for the impact of race on uncounted ballot rates, controlling for voting system, age, education, and unemployment.

Table 5 Goes Here

We utilized two different multivariate statistical models to produce our estimates. First, we use simple linear regression. However, linear regression in this application might not be appropriate, because of the nature of the dependent variable (uncounted ballot rates). Linear regression works best when the dependent variable is continuous and the relationships between independent variables and the dependent variable are linear; in a situation where the dependent variable is bounded (like a percentage) neither property is necessarily present.[14] So we also estimate the multivariate models using grouped logit, a statistical technique that is well-suited for aggregated and bounded data.[15]

Table 5 gives the linear regression and grouped logit results for residual votes, overvotes, and undervotes. Entries that have a single star are statistically significant at the $p < .10$ level, while those with two stars are significant at the $p < .05$ level. Beginning with the linear regression models, we find that two variables have significant effects on residual vote rates, the nonwhite population and the median age. Both have positive signs, indicating that as the nonwhite population and the median age of a county increase, so do their residual vote rates, controlling for the other variables in the model. When we turn to the over- and undervote linear regression results in Table 5 we again see that the nonwhite population variable has a positive and significant sign in each model.

[14] One important problem that arises with linear regression models in a situation like the one we are studying is that the linear regression estimates can often produce predictions that are outside the 0 to 100% interval. For example, two of the linear regression specifications we report in Table 6 below produce predictions for uncounted ballots for certain ranges of the independent variables that are below zero. This is an important indication that the linearity assumption necessary for the linear regression model to produce accurate estimates of the effects of the independent variables is a poor assumption to make in this context.

[15] For more discussion of the grouped logit model see Greene (2000). For an application of the grouped logit model that is very similar to that we present here, see Alvarez and Nagler (2001). In this application, we used the weighted least squares approach for estimating the grouped logit model, rather than the maximum likelihood approach.

We see that the median age variable is positive in both models as well, but is significant in only the overvote model. The unemployment rate has a positive and significant effect in the overvote model as well.

In the linear regression models reported in Table 5 we find that whether a county has a punchcard voting system does not have a statistically significant effect on residual vote rates, nor on over- or undervotes, controlling for other attributes of the counties. We do see a consistently negative sign on the impact of having a precinct error correction system, though, in each of the three linear regression models. The estimate is significant, moreover, in the overvote model. These results indicate that once we control for other attributes of counties, that whether a county has a punchcard system or not has no statistically discernable effect on the rate of uncounted ballots; however, the use of a precinct-based error checking system leads to lower uncounted ballot rates, especially in the case of overvotes.

In the right-hand panel of Table 5 we have the grouped logit results. In many ways these estimates parallel the linear regression results. In particular, we see positive and significant estimates for the nonwhite population, and for median age in the overvote grouped logit analysis. Unemployment rates again have a positive and significant impact on overvote rates, but in the grouped logit analysis have a significant and positive effect on undervotes as well.

The biggest difference between the linear regression and grouped logit specifications, though, arises in the case of voting system effects. In the grouped logit specification we find that whether a county employs a punchcard system has a positive and statistically significant impact on each measure of uncounted ballots that we employ. Furthermore, the precinct-counting coefficients are all again negative, but are not statistically significant in any of the three grouped logit models.

What conclusions can we draw from Table 5? First, both statistical analyses agree that three demographic or economic attributes of counties influences the rate of uncounted ballots, controlling for the type of voting system employed in the county: the size of the nonwhite population, the median age, and the unemployment rate. The most consistent results are for the nonwhite population, as in each of the six analyses we present in Table 5 we obtain positive and significant results. Median age and unemployment rates influence the overvote rate in both types of statistical model as well.

Second, the two statistical specifications produce complicated results for the impact of the voting system. Regarding residual and overvote rates, both models produce positively signed estimates for whether the county uses a punchcard system, and the estimates are statistically significant in the grouped logit models. Regarding the undervote rate, the models produce inconsistent results about the impact of a punchcard system on a county's undervote rate. With respect to the impact of a precinct-based error correction check, both models produce results that indicate such a system leads to lower uncounted ballot rates, but the estimate is only statistically reliable for overvotes in the linear regression specification.

Table 6 Goes Here

To delve deeper into the consistent result in Table 5 regarding the impact of nonwhite population on uncounted ballot rates across the model specifications, we utilized an interactive specification for nonwhite population and whether the county used a punchcard voting system or not. Thus, the basic model specification is the same: we use the residual vote rate, and the over- and undervote rates, as our dependent variables; we also use the same independent variables as above, including nonwhite population, median age, per capita primary and secondary education expenditures, the unemployment rate, and whether the county used a punchcard system or a precinct-based counting system. The only addition is an interaction term for nonwhite population and punchcard voting system use, to examine if the estimates in Table 5 are due to the higher patterns of uncounted ballots we observed primarily in punchcard counties for nonwhites earlier in this paper.

Table 6 reports the results for the interactive specifications, using both linear regression and grouped logit to estimate the model parameters. The introduction of the interaction term does lead to some important new interpretations of the effect of the size of the nonwhite population in a county on uncounted votes, especially if the county in question uses a punchcard system. Importantly, we see in Table 6 that three important changes in the results arise due to the interactive specification, no matter whether we use linear regression or grouped logit. First, with the inclusion of the interactive term in the multivariate model the nonwhite variable that had previously been a consistent predictor of uncounted ballots in the previous models (Table 5) is now statistically insignificant. Second, we see that in all three of the linear regression models, and in one of the three grouped logit models, the estimated impact of the punchcard system is statistically significant, and negative. Third, we see that in every specification reported in Table 6 the interaction term is positive and statistically significant.

How can these results be easily understood? Unfortunately, interactive regression models, especially nonlinear multivariate models like grouped logit, can be difficult to easily discuss and understand. To make the results of this interactive analysis more readily understandable, we translate the three grouped logit models into a graphical analysis, presented in three figures. Figure 1 provides an analysis for residual votes, figure 2 gives the results for overvotes, and figure 3 shows the same results for undervotes. Each figure depicts the rate of uncounted ballots for successive population sizes of nonwhites, depending on whether the county uses a punchcard system (graphed with circles), no punchcard and no local error checking devices in precincts (graphed with triangles), and no punchcard but precinct-based error checking (graphed with squares). Each of these graphs shows the estimated relationship between the size of the nonwhite population and uncounted ballot rates, for certain types of voting systems, holding the other variables in the model constant at their mean values.

Figures 1 through 3 go here

Beginning with Figure 1, we see that the estimated rate of uncounted ballots, here measured by the residual vote rate, is low for counties with small nonwhite populations, no matter what type of voting system employed by the county. But in counties with high nonwhite populations, for example, where 50% of the county's population is nonwhite, the story is quite different. The residual vote rate for counties that have a 50% nonwhite population is almost 2.5% in punchcard systems, whereas the residual vote rate is less than 1% for counties using other voting systems. The differences between punchcard and non-punchcard counties are stark, and the positive and strong relationship between the size of the nonwhite population in a county and the residual vote rate in punchcard counties is easily seen in Figure 1.

The basic conclusion is the same for overvotes, as shown in Figure 2. The rate of overvoting is quite low for counties with small nonwhite populations, but the divergence between counties with large nonwhite populations using punchcard voting systems, and counties with large nonwhite populations using other types of voting systems is stark. To continue the example of looking at counties with nonwhite populations of 50%, we see in Figure 2 that their overvote rate is twice as high in counties using punchcard systems relative to counties using other voting systems.

Figure 3 provides the same analysis, but for undervotes. We again see the same patterns as before. When there are small nonwhite populations, the undervote rate is low; when there are large nonwhite populations, the undervote rate is extremely high in punchcard counties but very low in non-punchcard counties. With specific regard to undervotes, the estimated difference in undervote rates for counties with 50% nonwhite populations is an almost eight-fold increase in undervotes for counties using punchcard relative to counties not using punchcard systems.

Thus, in this section we have carefully examined the relationship between race and uncounted ballots in California's 2000 Presidential election. Using a variety of statistical techniques, ranging from simple bivariate correlation analyses, to ecological inference estimates, and then to two different multivariate models, we have repeatedly found the same basic patterns. First, nonwhite voters have higher rates of uncounted ballots, over-, and undervotes, than white voters. Second, this same pattern is most apparent in counties using punchcard voting systems.

5 Conclusion

In a democratic society, the act of casting a ballot on election day is perhaps the most important type of political expression. The translation of a voter's opinions and preferences into the ballot they cast should occur without any error, and the technology used to register and tabulate the vote should transparently operate and not impact the vote in any way. Only when the translation of preferences to counted ballot is transparent, when the process is neutral and does not favor any particular party or candidate, or does not harm the voting process for certain classes of voters, is a voting system operating correctly.

We have presented evidence in this paper that this transparency of the voting system may not have existed in the 2000 presidential election in California. We provided substantial evidence, examined from a number of methodological perspectives, documenting that there are higher rates of uncounted ballots in counties with large nonwhite populations in the 2000 presidential election in California. We also showed that this effect is particularly noticeable in California counties that employ punchcard voting systems, especially counties with high populations of nonwhites.

Why nonwhites in California, and in other election jurisdictions, have higher rates of uncounted ballots than other voters is an open question. As many recent studies, including ours, show the correlation between nonwhite population and uncounted ballot rates while controlling for income and education differences (Posner 2001; GAO 2001), clearly these other factors cannot account for the observed correlations. Instead, the differences might be caused by differing levels of political information and involvement, differences in familarity with voting system technologies, higher levels of newly mobilized voters in nonwhite communities, or other factors. Our focus in this paper has been on documenting the differences, as nonwhite voters are a legally protected class. Future research must study the underlying question of why nonwhite voters have higher levels of uncounted ballots.

Our analysis is of course limited in scope to one presidential election and to only the

State of California. But as California counties employ a wide range of voting systems, and as California has a heterogeneous electorate, we argued that California is an ideal state for close study. Given the disagreement in contemporary public policy studies of this same problem at the national level, for example the disagreement between the GAO (2001), Ansolabehere's study (2001), and the U.S. House Committee on Government Reform Minority Staff report (2001), we believe that additional state-level analyses like ours and Posner's (2001) study of Florida are necessary. First, states and lower-level political jurisdictions are the appropriate focus for studies like these because they are the legally relevant juristictions for voting rights studies. Second, it is possible that nationally focused studies ignore important features of a state or local jurisdiction's political context. Third, it may be difficult if not impossible to develope finely grained data at the national level, as different states and local jurisdictions have different rules and procedures for generating election administration information.

No matter what the methodoligical implications of our work, it is clear that we have uncovered significant patterns in the 2000 presidential election data from California that have legal implications for California and other states. There is no doubt — and our work reinforces the previous studies looking at voting systems and uncounted ballot rates — that punchcard voting systems have higher rates of uncounted ballots than other voting systems. To the extent that voters within the same electoral jurisdiction are using different voting systems (the so-called "mixed" voting systems), the implication of our study and other similar studies is that there is the possibility of an equal protection issue in such electoral jurisdictions. Also, voters within a larger jurisdiction, like a state, might make the same claim if they reside in a county using one of the punchcard voting systems that have high rates of uncounted ballots, if other voters in other counties in the same state use different voting systems that produce lower rates of uncounted ballots.

Furthermore, our study indicates that in California counties using punchcard voting systems, the rate of uncounted ballots increases with the size of the nonwhite population. The same does not appear to be true in counties using other voting systems. This raises the possibility that Section 2 of the Voting Rights Act might be used as the basis for additional litigation in other election jurisdictions using punchcard voting systems, where similar patterns can be observed. There are many other states in which heavily nonwhite counties use punchcard voting systems, especially the Votomatic system. Unless significant action is taken to provide resources to states and local election officials to replace punchcard voting systems quickly, it is possible that a new area of voting rights litigation might arise regarding whether or not nonwhite voters have higher uncounted ballot rates where punchcard voting systems are employed.

6 References

Alvarez, R. Michael, Tara Butterfield, and Catherine Wilson. 2001. "Counting Ballots." Manuscript, California Institute of Technology.

Alvarez, R. Michael and Jonathan Nagler. 2001. "The Likely Consequences of Internet Voting for Political Representation." *Loyola Law Review* 34,3: 1115-1154.

Ansolabehere, Stephen. 2001. "Voting Machines, Race, and Equal Protection." *Election Law Journal*, 1: 61-70.

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

California Department of Finance, Economic Research. 2001. *California County Profiles*. Sacramento, CA: California Department of Finance.

Caltech/MIT Voting Technology Project. March 30, 2001. "Residual Votes Attributable to Technology: An Assessment of the Reliability of Existing Voting Equipment." Pasadena, CA: Manuscript, California Institute of Technology

Caltech/MIT Voting Technology Project. July, 2001. *Voting: What Is, What Could Be.* Pasadena, CA: California Institute of Technology.

Clubb, Jerome M. and Michael W. Traugott. 1972. "National Patterns of Referenda Voting: The 1968 Election." In Harlan Hahn, *People and Politics in Urban Society*, Beverly Hills, CA: Sage.

Darcy, Robert and Anne Schneider. 1989. "Confusing Ballots, Roll-off and the Black Vote." *Western Political Quarterly* 42: 347-64.

Fraser, Jeanette Lynn. 1985. "The Effects of Voting Systems on Voter Participation: Punch-card Voting Systems in Ohio." Ph.D. Dissertation, Ohio State University, Columbus, Ohio.

Greene, William H. 2000. *Econometric Analysis*, 4th Edition. New York: Macmillan Publishing Company.

Herron, Michael C. and Jasjeet S. Sekhon, 2001. "Overvoting and Representation: An Examination of Overvoted Presidential Ballots in Broward and Miami-Dade Counties." Manuscript, Harvard University.

King, Gary. 1997. *A Solution to the Ecological Inference Problem: Reconstructing Individual Behavior from Aggregate Data.* Princeton, NJ: Princeton University Press.

Knack, Stephen and Martha Kropf. 2001. "Invalidated Ballots in the 1996 Presidential Election: A County-Level Analysis." Manuscript, University of Missouri at Kansas City.

Knack, Stephen and Martha Kropf. 2001b. "Roll Off at the Top of the Ballot: Intentional Undervoting in American Presidential Elections." Manuscript, University of Missouri at

Kansas City.

Mather, George B. 1964. *Effects of The Use of Voting Machines on Total Votes Cast: Iowa — 1920-1960.* Iowa City, IA: Institute of Public Affairs, University of Iowa.

Minority Staff, Special Investigations Division, Committee on Government Reform, U.S. House of Representatives, July 2001. "Income and Racial Disparities in the Undercount in the 2000 Presidential Election." Washington, D.C.: U.S. House of Representatives.

Montgomery, Michael J. 1985. "Voting Systems and Disenfranchisement." *Election Politics* 2: 16-19.

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

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

Shocket, Peter A., Neil R. Heighberger, and Clyde Brown. 1992. "The Effect of Voting Technology on Voting Behavior in a Simulated Multi-Candidate City Council Election: A Political Experiment of Ballot Transparency." *Western Political Quarterly* 45,2: 521-537.

Stiefbold, Rodney P. 1965. "The Significance of Void Ballots in West German Elections." *American Political Science Review* 59,2: 391-407.

Thomas, Norman C. 1968. "Voting Machines and Voter Participation in Four Michigan Constitutional Revision Referenda." *Western Political Quarterly* 21: 409-19.

Tufte, Edward R. 1974. *Data Analysis for Politics and Policy*. New York: Prentice-Hall.

U.S. Census Bureau, May 2001. "Profiles of General Demographic Characteristics: 2000 Census of Population and Housing, California." Washington, D.C.: U.S. Department of Commerce.

U.S. General Accounting Office, October 2001. "Statistical Analysis of Factors that Affected Uncounted Votes in the 2000 Presidential Election." Washington, D.C: GAO-02-122.

Walker, Jack L. 1963. "Negro Voting in Atlanta: 1953-1961." *Phylon*, 379-387.

Walker, Jack L. 1966. "Ballot Forms and Voter Fatigue: An Analysis of the Office Block and Party Column Ballots." *Midwest Journal of Political Science* 10,4: 448-463.

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

Table 1: US and California Use of Voting Systems

	Percentage of US Counties	Percentage of Population	Number of California Counties	Percentage of CA Counties	Percentage of CA Population
Paper	12.5%	1.3%	0	0	0
Lever	14.7%	17.8%	0	0	0
Punch Card	19.2%	34.4%	30	51.7%	74.0%
Datavote	1.7%	3.5%	21	36.2%	38.6%
Votomatic	17.5%	30.9%	5	8.6%	42.6%
Pollstar	NA	NA	4	6.9%	13.6%
Optical Scan	40.2%	27.5%	27	46.6%	21.4%
Central	NA	NA	11	19.0%	9.0%
Precinct	NA	NA	16	27.6%	12.4%
Mixed	4.4%	8.1%	0	0	0
Electronic	9.0%	10.7%	1	1.7%	4.6%

Note: Statistics for the United States are from Caltech/MIT (2001). Statistics for California were compiled by the authors from data provided by the California Secretary of State and the U.S. Census.

Table 2: Residual, Over, and Under Votes by Voting System

	US Average Residual Vote, 88-2000	US Average Residual Vote, 2000	CA Average Residual Vote, 2000	CA Percent Reported Undervote	CA Percent Reported Overvote
Paper	1.9%	1.3%	NA	NA	NA
Lever	1.9%	1.7%	NA	NA	NA
Punch Card			1.9%	2.3%	0.3%
Datavote	2.9%	1.0%	0.8%	0.2%	0.1%
Votomatic	3.0%	3.0%	2.4%	3.5%	0.4%
Pollstar			1.8%	1.5%	0.3%
Optical Scan	2.1%	1.2%	0.9%	0.4%	0.2%
Central			0.9%	0.6%	0.2%
Precinct			0.8%	0.3%	0.1%
Mixed	2.2%	2.7%	NA	NA	NA
Electronic	2.9%	1.6%	0.9%	0.4%	0

Note: Statistics for the United States are from Caltech/MIT (2001). Statistics for California were compiled by the authors from data provided by the California Secretary of State.

Table 3: Race and Voting System Use, California 2000

	Percent Non-white Using System	Percent Black Using System	Percent Asian Using System	Percent Hispanic Using System
Punch Card	43.2%	7.3%	11.6%	34.3%
Datavote	33.8%	2.4%	8.7%	32.7%
Votomatic	47.6%	9.5%	12.0%	37.2%
Pollstar	41.6%	7.0%	14.0%	27.5%
Optical Scan	33.7%	4.5%	10.1%	24.9%
Central	30.4%	5.1%	7.4%	22.7%
Precinct	36.0%	4.1%	12.0%	26.4%
Electronic	34.4%	6.2%	3.7%	36.2%

Note: Statistics for the United States are from Caltech/MIT (2001). Statistics for California were compiled by the authors from data provided by the California Secretary of State and the U.S. Census.

Table 4: Correlation Between Race, Residual, Over- and Undervotes

	Correlation Between Non-whites and Residual	Correlation Between Non-whites and Overvotes	Correlation Between Non-whites and Undervotes
Punch Card	0.44	0.06	0.53
Datavote	0.16	0.17	-0.05
Votomatic	0.17	0.70	0.43
Pollstar	0.23	0.52	0.14
Optical Scan	-0.15	-0.03	0.10
Central	-0.22	-0.26	-0.35
Precinct	-0.11	0.27	-0.28
Electronic	NA	NA	NA

Note: Table entries are bivariate correlation coefficents.

Table 5: Regression Analyses

Variables	Linear Regression			Grouped Logit		
	Residual	Over	Under	Residual	Over	Under
Constant	-1.65*	-1.99**	-1.59	-6.70**	-9.48**	-1.99
	1.27	1.02	2.65	1.43	1.45	4.18
% Non-white	0.02**	0.01*	0.03**	0.03**	0.02**	0.05**
	0.009	0.007	0.02	0.005	0.005	0.01
Median Age	0.05**	0.05**	0.04	0.02	0.08**	-0.04
	0.03	0.02	0.06	0.03	0.03	0.08
Education $	0.0002	-0.000004	0.0004	0.0004	-0.0002	-0.003**
	0.0003	0.0003	0.0008	0.0005	0.0007	0.002
Unemployment	0.02	0.06**	-0.03	-0.005	0.08**	0.14**
	0.02	0.02	0.05	0.03	0.03	0.07
Punchcard	0.13	0.11	-0.20	0.41**	0.50**	0.61*
	0.21	0.14	0.37	0.21	0.18	0.46
Local count	-0.21	-0.20*	-.026	-0.23	-0.23	-0.45
	0.21	0.15	0.38	0.22	0.22	0.51
Sample size	58	43	43	58	42	42
Adjusted R^2	0.13	0.26	0.004	0.57	0.45	0.67

Note: Entries give linear regression or grouped logit estimates (first estimate for every variable in each column) or the associated standard error (second estimate for every variable in each column). * indicates statistical significance at the $p < .10$ level, ** at the $p < .05$ level, one-tailed tests.

Table 6: Interactive Regression Analyses

	Linear Regression			Grouped Logit		
Variables	Residual	Over	Under	Residual	Over	Under
Constant	-1.54*	-2.01**	-1.63	-5.45**	-8.56**	2.77
	1.22	0.98	2.57	1.55	1.48	4.01
% Non-white	0.009	0.001	0.01	0.01	-0.001	-0.02
	0.01	0.008	0.02	0.01	0.01	0.03
Median Age	0.05**	0.05**	0.04	0.009	0.06**	-0.13**
	0.03	0.02	0.05	0.03	0.03	0.08
Education $	0.0004	0.0001	0.0007	0.0001	-0.0000	-0.002*
	0.0003	0.0003	0.0008	0.0005	0.0006	0.002
Unemployment	0.02	0.06**	-0.01	0.004	0.08**	0.11**
	0.02	0.02	0.05	0.03	0.03	0.07
Punchcard	-0.59*	-0.39*	-1.37**	-0.21	-0.21	-2.26**
	0.37	0.28	0.75	0.39	0.41	0.99
Local count	-0.22	-0.22*	-0.30	-0.12	-0.18	-0.44
	0.21	0.14	0.37	0.22	0.21	0.46
Race-Punch	0.02**	0.02**	0.04**	0.02**	0.02**	0.08**
Interaction	0.01	0.008	0.02	0.01	0.01	0.03
Sample size	58	43	43	58	42	42
Adjusted R^2	0.19	0.32	0.06	0.59	0.49	0.74

Note: Entries give linear regression or grouped logit estimates (first estimate for every variable in each column) or the associated standard error (second estimate for every variable in each column). * indicates statistical significance at the $p < .10$ level, ** at the $p < .05$ level, one-tailed tests.

Figure 1: Residual voting estimate from grouped logit analysis

Figure 2: Overvoting estimate from grouped logit analysis

Figure 3: Undervoting estimate from grouped logit analysis