From Punchcards to Touchscreens: Some Evidence from Pasco County, Florida on the Effects of Changing Voting Technology[1]

Michael C. Herron[2]　　Jeffrey B. Lewis[3]

April 24, 2006

[1]An earlier draft of this paper was presented at the 2006 Annual Meeting of the Midwest Political Science Association, Chicago, IL. The authors thank Tina Catania for research assistance, Shirley Laughery in the office of the Pasco Supervisor of Elections for data and clarifications, and seminar participants at the University of Michigan at Ann Arbor, George Washington University, and Emory University for helpful comments.

[2]Department of Government, Dartmouth College. 6108 Silsby Hall, Hanover, NH 03755-3547 (Michael.C.Herron@Dartmouth.edu).

[3]Department of Political Science, University of California, Los Angeles. 4289 Bunche Hall, Los Angeles, CA 90095-1472 (jblewis@ucla.edu).

Abstract

Many counties across the United States are in the process of replacing old voting technologies with electronic touchscreen systems. Little is known, however, about the effects of touchscreens on voter candidate choices and dropoff decisions. With this in mind we study Pasco County, a Gulf Coast county in Florida that after the 2000 general election changed its voting technology from Votomatic punchcards to iVotronic touchscreens. We treat Pasco County's technology shift as a natural experiment and using a collection of several hundred thousand ballot images show that touchscreens facilitated the casting of valid votes by those voters who wanted to cast them; that touchscreens produced ballots with relatively high degrees of internal predictability in terms of dropoffs and vote choices across races; and, that voters' proclivities to dropoff from voting were correlated with precinct demographics in 2000 yet were essentially uncorrelated in 2004. These results in conjuction with others imply that touchscreens diminish voter error rates and hence constitute an improvement over punchcards. Nonetheless, we also show that touchscreen machines appear to produce larger numbers of affirmative votes in judicial retention races, and this should raise a cautionary flag about effects that these machines have on encouraging voter participation in low information races.

1 Introduction

One of the more striking developments in American election administration is the growth of touchscreen voting. In the 2000 general election approximately 9.9% of United States counties used touchscreens. This figure jumped to 17.57% in 2002 and then to 21.48% in 2004, and slightly more than a third of United States counties will use touchscreen voting in November, 2006.[1]

The impetus for the ongoing abandonment of ostensibly problematic voting technologies in favor of touchscreens can be laid largely at the feet of the 2000 general election and its extended Florida aftermath (e.g., Merzer 2001, Posner 2001, The Washington Post 2001). If nothing else, the now infamous pictures of southern Florida election judges peering at "pregnant" and "hanging" chads on punchcard ballots made it clear that at least one prominent voting technology was prone to producing invalid (e.g., Mebane 2003) and perhaps even incorrect (e.g., Sinclair, Mark, Moore, Lavis & Soldat 2000, Wand, Shotts, Sekhon, Mebane, Jr., Herron & Brady 2001) votes. Worse yet, problems associated with punchcards appeared to be disproportionately pronounced in precincts and counties with large minority populations (United States Commission on Civil Rights 2001, Ansolabehere 2002, Tomz & van Houweling 2003). Long taken for granted by academics and politicians alike, deficiencies in the fundamental mechanics of U.S. electoral administration threatened the legitimacy of the 2000 presidential election and became a source of national embarrassment. The aftermath of the 2000 debacle has witnessed considerable funds being dedicated to modernizing the apparatus by which voters' choices are recorded, and in many of the counties in which punchcards were employed one now finds touchscreen voting machines.

Despite the recent surge in support for touchscreen voting, almost nothing is known about the effects of this technology on the way that voters make election participation decisions ("Should I vote in the sheriff race?") and choose candidates conditional on participating ("Which sheriff candidate should I support?"). In light of this, we offer an individual-level, voting record analysis of what happens when a county transitions from punchcard to touchscreen voting. Our focus is Pasco

[1] See the publicly available reports authored by Election Data Services, Inc. The 2004 report can be found at http://www.electiondataservices.com/EDSInc_VEstudy2004.pdf and the 2006 report at http://www.electiondataservices.com/EDSInc_VEStudy2006.pdf.

County, a Gulf Coast county in Florida that in the 2000 general election used Votomatic punchcard technology but by the time of the 2004 general election had adopted iVotronic voting machines. The iVotronic is a Direct Record Electronic (DRE) machine manufactured by Election Systems & Software. Put simply, we treat Pasco County's technology change as a natural experiment and consider the question, what was the effect on voters of Pasco County's changing from punchcards to touchscreens?

Using over 300,000 electronic images of ballots cast in Pasco County in the general elections of 2000 and 2004 we show that, compared to punchcards, touchscreens (1) increased voter propensities to participate in down-ballot races, particularly in low salience contests such as judicial retention elections; (2) produced patterns in down-ballot participation levels that were more predictable within individual ballots and less correlated with aggregate voter demographics; and (3) produced greater partisan coherence within ballots. These findings suggest that touchscreen machines cause fewer voter errors and induce voters to participate in some races in which they would otherwise have intentionally abstained. Nonetheless, we also show that touchscreens appear to be responsible for causing excessive numbers of "yes" (i.e., retain) votes in judicial retention elections. This latter finding should raise a cautionary flag about potential effects of touchscreens on election outcomes.

Our findings have a number of normative implications. A transition from punchcards to touchscreens that leads to few voter errors is an obvious good. Indeed, it would be hard to argue that increasing the fidelity with which votes are recorded would do anything other than improve democratic processes. On the other hand, the normative value of other potential touchscreen effects is not clearcut. To the extent that touchscreens cue some voters to participate in contests about which they are uninformed and if the choices that such voters make in those races exhibit bias in favor of one candidate or alternative (for example, the "no" alternative on a ballot initiative or the first candidate listed in a non-partisan race), touchscreens could by some standards yield vote totals that are "worse" than those that would have obtained under punchcard voting. While a full treatment of the normative issues in voting technology is beyond our scope, we highlight normative considerations where we believe our empirical evidence sheds light on them.

2 Voting Technology: The State of the Literature

The voting technologies currently in use in the United States consist of paper ballots, lever machines, punchcards, optical scan ballots, and electronic touchscreens, often called Direct Record Electronic (DRE) machines. Knack & Kropf (2002) show that the distribution of inferior machines is not skewed toward any particular demographic group, and Ansolabehere & Stewart III (2005) note that voting technology usage is not uniform within states.

Since the aftermath of the 2000 presidential election the primary criterion used to evaluate voting technology has been the residual vote rate. The residual vote rate in a given contest of an election is the fraction of ballots cast in the election that do not contain valid votes in the contest. A general election in a county, then, has many residual votes rates, one for the presidential race, one for a U.S. Senate race (if applicable), one for a set of Congressional races (if applicable), and so forth.

Residual vote rates are known to vary by voter race (Tomz & van Houweling 2003, Herron & Sekhon 2005) and, most importantly from our perspective, by technology. Punchcard voting is widely considered to be associated with relatively high residual vote rates, and in contrast touchscreen voting technology is thought to produce relatively low residual vote rates (Brady, Buchler, Jarvis & McNulty 2001, Stewart III 2004, Stewart III 2005, Ansolabehere & Stewart III 2005). Residual vote studies typically focus on top electoral contests like presidential races, but touchscreens tend to have low residual vote rates in down-ballot contests as well (Nichols 1998, Nichols & Strizek 1995).

Residual votes are closely associated with voter rolloff, the somewhat vaguely characterized phenomenon where voters participate in top-ballot races and then, due ostensibly to fatigue, boredom, indifference, or irritation, do not participate in down-ballot races. Notwithstanding some definitional ambiguity about whether rolloff can be followed by rollon, reducing rolloff rates is often seen as a goal on par with that of reducing residual vote rates.

Beyond the association of residual vote rates and voting technology, Hanmer & Traugott (2004) study the vote-by-mail (VBM) system implemented statewide in Oregon in 2000. Like punchcards and touchscreens, VBM can be thought of as a technology that translates voter preferences into political outcomes. Hanmer & Traugott show that the introduction of VBM across Oregon did not disrupt patterns in voter behaviors evident prior to 2000. And, Herrnson, Bederson, Niemi, Conrad, Hanmer & Traugott (2006) employ a unique experimental design to assess the prevalence and correlates of voter errors. They show that voters' selecting an incorrect candidate is the most common form of voter error. Herrnson et al. appear to be the first to study directly voter errors, and as we show later our results have implications for voter error rates as well.

Much of the existing literature on voting technology suffers from two limitations. First, and as noted by others (e.g., Stewart III 2004), residual vote rates by construction aggregate voter behaviors, i.e., deliberate abstention, and deliberate ballot spoilage, that are very different from one another. And, residual vote rates tell us nothing about whether certain technologies produce voting errors that lead to valid votes, albeit valid votes for unintended candidates.

Second, most of the literature on voting technology relies on vote returns aggregated to either the precinct or county level. Consequently, for example, low residual vote rates in down-ballot contests are consistent with attention being paid to down-ballot races by voters who cast valid top-ballot votes yet are also consistent with attention being paid to down-ballot races by voters who did not cast top-ballot votes. Aggregate returns, that is, do not provide much leverage on the question of whether touchscreen voting reduces rolloff.

3 Research Design

Our research design does not rely on aggregated residual vote rates, and it is motivated by the fact that a key difference between punchcard and touchscreen technology is that casting a vote when using the latter is relatively *easier*. In this section we first explain the potential consequences of such ease of use and then show how our design will ferret out such consequences if they indeed exist.

3.1 Ease of Use and the Direct and Indirect Effects of Touchscreens

It seems obvious that casting a valid vote with a touchscreen machine is easier than doing so with punchcard technology. Punchcard voting involves the use of a pre-scored, cardboard ballot and a stylus, but touchscreens instead use a computer screen designed to be pressed by fingers much like omnipresent automated teller machines. Unlike punchcards, touchscreens can prohibit certain types of invalid votes (for example, overvotes) and can be programmed to display candidate choices in a voter's choice of language without the voter's requesting a special ballot. Moreover, iVotronic machines have a voter review screen which prompts (or gives, depending on one's perspective) a voter who has finished voting the opportunity to consider his or her vote choices and/or abstentions. Herrnson, Abbe, Francia, Bederson, Lee, Sherman, Conrad, Niemi & Traugott's (2005) exit poll of approximately 1,200 touchscreens users shows that most such individuals found touchscreens relatively easy to use.

The implications of this greater ease of use turn in large part on two potential underlying effects, what we call a direct effect and an indirect effect. The direct effect of touchscreen voting is an increase in recording fidelity; we call this effect direct because, presumably, the ultimate objective of a good voting technology is high fidelity. If there is increased recording fidelity under touchscreens, then votes recorded with these machines more faithfully reflect the choices that voters truly intended to make compared to the choices that would have been recorded had they been executed with punchcards.

The indirect effect of touchscreens is an increase in the number of races that a given voter participates in across a ballot's worth of choices. If this effect obtains, then what would have been abstentions under punchcards are transformed into valid votes. We expect an indirect effect because touchscreens impose fewer costs on voters and hence drive them away less readily, and because they prompt (at least in the case of the machines used in Pasco County) voters to participate in races in which they initially abstain.

If the direct effect of touchscreens dominates the indirect, then following Pasco County's transition from punchcard to touchscreen voting we expect to observe (1) contest dropoff choices by voters that are more predictable as a function of what sort of offices are being contested and less predictable as a function of voter characteristics like socioeconomic status; and, (2) greater predictability of candidate choices within ballots—because, say, Democratic voters are not accidentally voting for Republican candidates.

On the other hand, if the indirect effect of touchscreens dominates the direct, then we expect to observe (1) increased levels of down-ballot participation without a corresponding change in the coherence of abstention patterns; and, (2) a reduction in the predictability of candidate choices within ballots. The first consequence follows immediately from the definition of the indirect effect which *is* the reduction in intentional dropoff induced by touchscreen machines. The latter consequence follows if the marginal voter who would have abstained on a given race under punchcard voting participates in the contest under touchscreen voting yet is less informed about the contest than the average voter participating in the contest under punchcard voting, i.e., if touchscreens induce poorly informed voters to participate in races that in reality they want to skip.

3.2 Ballot Images from Pasco County

To determine the strength of the direct and indirect effects of touchscreens we turn to a collection of 303,758 electronic ballot images produced in Pasco in the general elections of 2000 and 2004. Ballot images differentiate this study from Stewart III (2004), who examines the state of Georgia's transition from an amalgam of voting technologies to touchscreens throughout. Our images, one per voter per election year, give us the most complete picture possible of what voters actually did when casting ballots in Pasco County in November, 2000 and in November, 2004.

In the 2000 general election there were two types of voters in Pasco County, election day and absentee. Both types of voters used Votomatic punchcards and we have images from all these voters (146,648 total images). In 2004, though, there were three types of voters in Pasco: election day, absentee, and early. Election day and early voters in Pasco used iVotronic machines and we have images from such voters (157,110 images); absentee voters, however, used optical scan technology instead for which no images are available. Since our interest lies in comparing election day voters in Pasco in 2000 with election day voters in Pasco in 2004—this minimizes confounding our results with time of vote concerns—this is not a meaningful loss.

The addition of early voting in 2004 potentially confounds our analysis. If, say, early voting drew certain types of voters from a pool of registered voters who normally voted on election day, as opposed to absentee, then our comparison of election day voters in 2000 with election day voters in 2004 could potentially be problematic. Fortunately for us, Berinsky (2005) notes that there is little evidence that election day voters and early voters are systematically different in substantial ways.

Ballot images have two key features insofar as being used to study voting technology. First, these images include the complete records of how individuals voted. This is noteworthy because the extent to which individuals rolloff and subsequently rollon ballots cannot be determined from aggregate voting returns. The same is true of the extent to which voters cast ballots that are internally predictable. Second, ballot images are not subject to the types of response biases that conceivably affect exit polls and other forms of surveys. Our results, then, are free from concerns about sampling frames, voter forgetfulness, and whether voters tend to lie when asked by pollsters whether and how they voted.

Most importantly, ballot images allow us to search for evidence of a change in the structure of votes cast by each voter. Assuming that the voting population of Pasco County and the electoral choices faced by Pasco voters were broadly similar in 2000 and 2004, changes in the patterns of recorded votes can reasonably be inferred to result from the county's adoption of touchscreen voting. By analyzing the full string of votes cast by each voter in 2000 and in 2004, we can detect the effects of the change in technology much in the same way that having access to all the answers that a group of students give on a standardized test can yield insight into the prevalence of cheating (e.g.. Jacob & Levitt 2003).

3.3 Why Pasco County?

In the interest of being forthright about our research design's limitations, our use of ballot images to avoid the pitfalls of aggregated returns, opinion polls, and self-reported votes is a function of studying a single county in the United States. Indeed, no one would argue that studying ballot images from only one county is preferable to studying images from many counties. Nonetheless, we focus solely on Pasco County, Florida because it alone, among all United States counties, is known by us to satisfy two key criteria. First, Pasco County has made available ballot images for the 2000 and 2004 general elections. Second, the collections of 2000 and 2004 ballot images from Pasco perfectly match the county's certified vote totals for these two election years (except for 2004 absentee voters for whom, as noted already, we have no images).

This latter criterion might seem practically trivial, but in fact it is not. For instance, 2000 general election ballot images from ten Florida counties can be found in an archive maintained by the National Election Study (NES).[2] The NES ballot archive is beyond the scope of this paper, but we note that many ballot images that should be in the archive are not present. In some instances audit files appear to have replaced actual ballot images; in other cases there are no obvious explanations for discrepancies between certified county vote totals and the vote totals one generates from the archive. Regardless, Pasco County does not suffer from any missing data problems (except for 2004 absentee voters), and we can thus analyze it without considering whether ballot missingness is random or is correlated with ballot or precinct features.

Nonetheless, a natural concern about our focus on a single county is the extent to which Pasco County is representative of counties in general. In fact, on a number of dimensions Pasco is rather typical.

*** Figure 1 about here ***

Figure 1 shows how Pasco compares to United States counties across four dimensions. With respect to economic status (upper left corner), Pasco County's median family income was $39,568 in 1999 dollars compared with an average value of $42,143.34 ($\hat{\sigma} = 9,885$) among United States counties tracked by the census. This is certainly within range of representativeness. With respect to the fraction of adults in the county with low levels of education, Pasco is also plausibly representative: 4.4% of its adults have low education compared to an average 5.9% ($\hat{\sigma} = 3.4\%$). With respect to political predispositions, the lower left corner of Figure 1 shows that Pasco is more Democratic

[2] See `http://www.umich.edu/~nes/florida2000/data/ballotimage.htm`.

than the average United States county.

Our primary concern about Pasco's representativeness is the county's racial breakdown. As shown in the bottom-right of Figure 1, Pasco County is less African-American than the average United States county. In particular, it is only 2.2% African-American while the average United States counties is 8.9% ($\hat{\sigma} = 15\%$) African-American.

Insofar as voting technology limitations are thought to affect less educated and politically unsophisticated voters more than highly educated voters who can navigate with ease all types of voting equipment (e.g., Herrnson et al. 2006), and given that in the contemporary United States there is an education gap between whites and other racial minority groups, it follows that our results on the effects of Pasco's transition from punchcards to touchscreens should be considered conservative. Anything we find would almost certainly be magnified in a more racially heterogeneous context that transitioned, like Pasco, from punchcards to touchscreens.

Similarly, our results are conditioned on the fact that Pasco's election administration system is well run; were it not we would not have access to Pasco ballot images. Well run elections presumably reflect the presence of highly educated voters who are high-demanders of efficient county government. Nonetheless, as we focus on voting idiosyncrasies, errors, and other anomalous behaviors, conditioning our results on the existence of a good election administration system will make our results conservative.

3.4 Analyzing Ballot Images, Part I: Simple Counts

Simple counts based on ballots with given characteristics, i.e., ballots with no presidential votes, can be very informative. Thus, our results focus initially on undervote rates by race and by ballot, voter rolloff rates, the extent to which voters voted for top-listed candidates, and blank ballot rates.

Entirely blank ballots deserve special attention because they are difficult to rationalize: why would a voter make a trip to the polls only to abstain in every contest possible? The most likely explanation for entirely blank ballots is a failure of voting technology to record faithfully a voter's intentions in even a single contest in which he or she presumably came to the polls to participate in. While blank ballots are rare, they are a relatively unambiguous signature of voting technology failure. If touchscreens failed to produce fewer blank ballots, one would certainly have to question whether touchscreen were more accurate in recording voters' intentions than punchcards.[3]

3.5 Analyzing Ballot Images, Part II: Scaling and Within-Ballot Predictability

Simple non-parametric counts cannot efficiently isolate the direct and indirect effects of touchscreens. For example, rolloff counts describe the extent to which voter rolloff is a function of the physical layout of a ballot. In fact, the order in which voter interest and participation wanes might be a function of contest visibility or another race-level feature that does not correspond to physical ballot location.

To address limitations inherent in non-parametric counts we adopt methods from the literature on test-taking and roll call voting (Poole & Rosenthal 1997) to infer from our ballot images underlying propensities for dropping off from individual races and for voting for particular candidates. Much as item response theory (IRT) applied to test questions is used to capture student aptitude levels (i.e., the ability of a student to answer a given test question correctly), our analysis provides estimates of voters' latent dropoff propensities (i.e., the willingness to not vote in a given race) and latent voter partisanships (i.e., the willingness to vote for Democratic over Republican candidates). And, as IRT applied to tests estimates the ability of test questions to discriminate high aptitude from low aptitude students, our ballot analysis provides estimates of so-called discrimination measures that reveal the degree to which race-level decisions to dropoff (alternatively, to support a particular Democratic candidate) are a function of a voter's underlying propensity to dropoff (alternatively, a voter's partisanship). The more a dropoff choice (candidate choice) in a specific contest is predicted by an underlying summary measure of dropoff (an underlying summary measure of partisanship), the more *predictable* are dropoff decisions (candidate decisions) across the entire ballot.

The methodological details of our application of IRT to Pasco County ballots can be found in Lewis (2001), Gerber & Lewis (2004), and Herron & Lewis (2005). With respect to dropoff,

[3] Or, alternatively, one would have to question the notion that residual vote rates indicate a failure of voting machinery to count votes that voters intended to cast.

the decision to dropoff in each race is written as a binary logit function of voter-specific underlying propensities to participate and contest-specific constant and slope parameters. For candidate choice, the choice of which candidate to support is modeled as a logit function (multinomial if there are more than two candidates) of an underlying unidimensional voter-specific partisanship and candidate-specific constant and slope parameters.[4] We pool voters by Pasco precinct and thus estimate the race-level coefficients of logistic regressions that predict decisions to dropoff from races (support candidates in each race) and average propensities to dropoff (average partisanships) of voters in each precinct.[5]

If controlling for office type race-level slope coefficients in our logistic regressions—these slopes are called discrimination parameters—are larger in magnitude in 2004 than in 2000, then we conclude that patterns of dropping off (candidate choice) are more predictable in 2004 than in 2000.[6] Such greater predictability would result from a reduction in the number of ostensibly random errors in the recording of vote intentions made using touchscreen machines as opposed to punchcard machines. We might also expect that discrimination in our partisanship model would fall between 2000 and 2004 as less informed voters are induced by the ease or prompting of touchscreen machines to participate in races in which they would have previously abstained.

Because individual propensities to dropoff and partisanship are latent measures with no natural scale, we must make identifying restrictions to ensure the comparability of our estimates across the 2000 and 2004 general elections. For our candidate choice or partisanship analysis, we assume that the mean partisanship of voters in the 26 Pasco County precincts that had the same geographic boundaries in 2000 and 2004 remained constant.[7] For the dropoff analysis, it is not reasonable

[4]The partisanship of each voter is determined only by the pattern of votes that he or she casts. We find *ex post* that Democratic candidates are generally favored by voters at one end of our ostensible partisanship scale and Republican candidates are favored by voters at the other end.

[5]We assume that, conditional on each precinct, abstentions represent vote choices that are are missing at random. And, our dropoff and partisanship scalings ignore all judicial retention races except for the first such race in 2000 (Justice Lewis) and in 2004 (Justice Bell). Given the large number of judicial retention votes particularly in 2000, were we not to drop all but Lewis and Bell then our extracted dropoff and partisanship dimensions would be a judicial dimensions. Votes across judicial retention races are highly correlated, as we show later, so we are not losing much information by scaling only the Lewis and Bell races.

[6]One might be concerned that the propensity to dropoff and partisanship are correlated. Our model of participation and vote choices assumes that these qualities are independent among voters within precincts though they can be correlated across precincts.

[7]There were 132 election day Pasco precincts in 2000 (one precinct reported zero votes and we ignore it) and 152

to suppose that dropoff propensities generally remained fixed between 2000 and 2004. Indeed, the indirect effect of touchscreen technology is exactly a reduction in the propensity to dropoff. Still we must make some identifying restriction so that 2000 and 2004 dropoff propensities are on the same scale. In what follows, we find substantially lower propensities to dropoff in high income precincts; this finding is consistent with literature reviewed earlier which suggests that low income, low education, and minority communities tend to have the highest residual vote rates under punchcard balloting. In light of this, we expect that the switch to touchscreens had relatively large effects on low income precincts and relatively small effects on higher income precincts. We thus identify our dropoff model by assuming that the average voter-level propensity to dropoff remained unchanged in the four highest income precincts in Pasco County among the 26 precincts that did not have boundary changes between 2000 and 2004. If propensity to dropoff also increased in these four precincts, then we will somewhat *underestimate* the change that took place in the others.

4 Results

In accordance with the discussion above, we begin our analysis by describing patterns in abstention and then turn to within-ballot predictability. Lastly we consider judicial retention contests because, as we show, they speak to the indirect effects that we identify in our predictability analysis.

4.1 Abstention, Rolloff, and Blank Ballots

Figure 2 describes Pasco County undervote rates by race for 2000 and 2004 election day voters.[8] Several patterns are evident this figure, which does not depend on data from ballot images of course. First, undervote rates are at their lowest, albeit are strictly positive, at the top of the 2000 and 2004 ballots; they fluctuate non-monotonically between U.S. House races and other county-wide races;

such precincts in 2004. The constant boundary precincts, which had the same numbers in 2000 and in 2004, are 2, 3, 7, 9, 20, 21, 29, 33, 35, 40, 42, 44, 52, 53, 55, 60, 65, 82, 93, 95, 97, 114, 118, 120, 121, 122. We determined which Pasco precincts had constant boundaries between 2000 and 2004 by overlaying electronic shapefiles of the precincts.

[8] Undervote here is strictly defined: 2000 election voters who overvoted are not considered to have contributed an undervote.

they jump at the level of judicial retention votes; and finally, they drop at the amendment level. Second, and perhaps not surprisingly, uncontested races tend to have relatively high undervote rates when compared to similar races. And third, undervote rates by virtue of their evident non-monotonicities are not consistent with standard notions of voter rolloff. If Pasco voters rolled off their ballots at various points and then never rolled on again, undervote rates would never decrease as one moved down a ballot.

*** Figure 2 about here ***

With respect to punchcard versus touchscreen voting technology, i.e., 2000 undervote rate patterns versus those from 2004, the two panels of Figure 2 do not appear appreciably different except that undervote rates are usually but not always lower than corresponding rates in 2004 than in 2000. Table 1 reports some basic statistics, which show that presidential undervote rates were significantly smaller in 2004 (approximately 0.456%) than in 2000 (approximately 1.11%). U.S. Senate undervote rates, though, do not follow this pattern.

*** Table 1 about here ***

Table 1 has two key implications. First, holding race type constant, touchscreens do not always lead to more voting. One can see this in the Pasco County sheriff race comparison. However, at this point we do not have leverage over whether this might reflect the fact that there were more races on the 2004 general election ballot than on the 2000 ballot.

Second, Table 1 shows that overvoting matters in the context of residual votes. Note that the two U.S. Senate races described in the table are both second on their respective ballots. The fact, therefore, that the 2004 U.S. Senate undervote rate was significantly greater than the 2000 rate cannot be attributed to ballot location differences. However, once overvotes are accounted for, we see that the residual vote rate in the 2004 U.S. Senate contest was lower than the corresponding rate in 2000. While this is only a single example, it nonetheless highlights the possibility that a key mechanism behind the low residual vote rates observed under touchscreen machines is the prohibition of overvotes.

Moreover, if 2000 U.S. Senate overvotes were protest votes as opposed to accidental non-votes, then we would expect to see undervotes replacing such protest votes in 2004. That we do not suggests that a non-trivial portion of the overvotes cast in the 2000 U.S. Senate race reflects accidental non-votes.

More evidence of relatively low undervote rates with touchscreen voting can be found in Figure 3, which displays two histograms of per ballot undervote rates. A ballot has an undervote rate of zero if the voter who cast it voted in all applicable races, and so forth.

As is evident in Figure 3, the zero (approximately) spike is considerably greater in 2004 than in 2000; this reflects the fact that, in 2004, a relatively large number of voters had very low or even zero undervote rates. Relatedly, the average undervote rate in 2000 was 0.168 ($\hat{\sigma} = 0.000675$) and in 2004 it was 0.098 ($\hat{\sigma} = 0.000543$). Confidence intervals for the difference between these two averages do not, even at very high levels of confidence, contain zero.

*** Figure 3 about here ***

Figure 3 is notable because lower undervote rates in 2004 compared to 2000 (as seen in Figure 2) do not provide much information on changes in voter-level undervote rates. It is thus clear from Figure 3 that the average Pasco voter in 2004 voted more frequently than in 2000, that the number of voters who participated in almost all races increased between these two years, and that the fraction of voters who participated in very few races decreased.

How can we compare voters in 2000 who cast undervotes to those who behaved similarly in 2004? One way is to consider ballot locations at which a first undervote was cast—see Figure 4. With respect to the 2000 election, while a small number of voters did not cast a valid vote in the 2000 presidential race (these voters have an initial undervote in the presidential race), initial undervotes were most common in legislative races, particular uncontested legislative races. This is evident in the spikes around U.S. house races as well as Florida house races. There is also spike in the first judicial retention vote, that for Florida Supreme Court Justice Lewis.

*** Figure 4 about here ***

Similar patterns characterize first undervote locations in the 2004 election although there does not appear to be a Florida house effect in 2004. The presence of a judicial retention vote effect—

note the Justice Bell spike in the lower panel of Figure 4—suggests that voters' disinterest in participating in judicial retention election is not ameliorated with technology. In fact, the similarities between the two plots in Figure 4 suggest that punchcard voting did not make voting such a task so as to encourage early undervotes.

If one effect of touchscreens is the lowering on average of undervote rates, a natural question is whether voter rolloff rates are similarly affected. The extent to which Pasco voters rolled off in 2000 and 2004 is described in Figure 5, which has a layout similar to earlier figures. In this figure a voter is said to rolloff at a given race if she voted validly in all races prior to the race of interest, abstained in that race, and then abstained in all applicable races afterward.

*** Figure 5 about here ***

Figure 5, which has voter counts on its vertical axes, has two notable features. The first is that, in fact, voter rolloff defined strictly is uncommon; this holds in 2000 and 2004 despite that fact that voting technology differed significantly across these two years.

Second, the difference between 2000 and 2004 rolloff patterns is similar to that between 2000 and 2004 undervote patterns: conditional on rolling off, voter rolloff rates were relatively high at the top of the ballot and at the judge retention level. That is, some election day voters, more in 2000 than in 2004, voted only for two federal offices (president and U.S. Senate) and some voted only up to the judicial retention level (Lewis in 2000 and Bell in 2004). There are also small rolloff spikes at the amendment level. Regardless, absolute rolloff numbers were low in both 2000 and 2004, and this suggests that ballot fatigue may not be as serious as problem as conventional wisdom holds.

We can loosen the definition of rolloff by characterizing it as occurring on a given ballot if there is some race after which no valid votes were cast yet before which there was at least one valid vote. Then (figure not shown), voter rolloff rates loosely defined were higher in 2000 then in 2004; there are areas on both 2000 and 2004 ballots that seem to be natural points of voter departure: the U.S. Senate race, the first post-U.S. House race, the Florida House race, the first post-Florida House race, the first judicial retention race, and the first amendment; and, voter rolloff rates vary in 2000 in ways that are not easily rationalizable.

Lower and more easily rationalizable rolloff patterns suggest that touchscreen machines decrease inadvertent undervotes and facilitate more coherent vote-casting. This is consistent with what we have called the direct effect of touchscreens. However, perhaps indifferent touchscreen voters are not rolling off precisely because they are punching the first available option for each race in order to complete voting as soon as possible. Were this the case then we would observe high rates of top candidate votes, what we define as top candidate straightlining. As depicted in Figure 6, though, there is no evidence of this.

*** Figure 6 about here ***

In fact, straightlining reared its head more frequently in 2000 than it did in 2004. This is evident in the spike that appears at the point of the 2000 Lewis retention vote in Figure 6: this spike reflects a group of voters who from Justice Lewis onward chose the top candidate (or option) in every following race. There are no comparable spikes in the 2004 panel of Figure 6, and straightlining is not a prominent feature of 2004 ballots until the point at which amendments start.

If, as this section implies, direct effects were strong, then we should see evidence of this in diminished blank ballot rates. Accordingly, Figure 7 reports blank ballot rates for election day voters through the Florida state legislature (i.e., a ballot is blank if it has no valid votes in all races prior to a Florida Legislature contest), up to the judicial retention level, and across the entire ballot.

*** Figure 7 about here ***

Chi-squared tests for differences in blank ballot rates across 2000 and 2004 are all extremely significant.[9] Moreover, the patterns in Figure 7's blank ballots rates are obvious: there were fewer blank ballots cast in 2004 than in 2000. This is consistent with the idea that touchscreens impose fewer costs on voters than punchcards.

4.2 Within-Ballot Predictability

Thus far our evidence supports the existence of indirect touchscreen effects (more voting in 2004 than in 2000) with some support of direct effects (more predictable patterns in abstention in

[9] The three chi-squared values for the blank ballot rates in Figure 7 are approximately 160, 205, and 230. With one degree of freedom these values are easily statistically significant.

the former). Nonetheless, the existence of more votes in 2004 is not synonymous with the presence of fewer voting errors or an increase in the number of meaningful votes. To explore these latter two subjects we turn now to an application of IRT to Pasco ballots cast in 2000 and 2004.

IRT as noted earlier consists of the simultaneous estimation of voter-level latent traits (i.e., ballot-level propensities to dropoff and ballot-level partisanships) and race-level discrimination parameters that are part of logistic forms that model choices. A given race is said to discriminate if voters with high dropoff propensities (alternatively, large partisanship scores) tend to vote in it (tend to vote for a specific type of candidate in it). A key question, then, is whether logit slope parameters—recall, these are known as discrimination parameters—are zero, meaning that an associated latent trait has no effect on the choice being studied. We are interested less in whether discrimination parameters are positive or negative (i.e., whether support for a candidate is increasing or decreasing in the underlying partisanship dimension) than in whether these parameters are zero or non-zero (i.e., whether the underlying dimension is predictive of a particular choice or not).

For each race in 2000 and 2004 we define a race-level dropoff discrimination score as the absolute value of the non-zero discrimination parameter—in a two-choice logistic regression one slope coefficient is normalized to zero. Similarly, for partisanship discrimination in two-candidate races we define a race-level partisanship score as the absolute value of the non-zero discrimination parameter. In contests involving more than two candidates the partisanship discrimination score is calculated as the difference between the maximum discrimination parameter and the minimum discrimination parameter across all candidates in a given race.

*** Figure 8 about here ***

Partisanship discrimination scores are displayed in Figure 8, and this figure shows that partisan discrimination is largest for top-ballot races and is essentially negligible for non-partisan amendments, judicial retention votes, and local referenda. This is logical, we would surmise, as these latter sorts are races are often not explicitly partisan in a Democratic versus Republican sense.[10] Within-precinct standard deviations of underlying voter partisanships are normalized to unity, and

[10]The race in Florida House District 62 is anomalous because of the tiny number of Pasco County voters (156) who turned out to vote in this district. In contrast, most races that intersected Pasco County had thousands of voters.

calculations (not shown) imply that moving the partisanship of a voter who is indifferent between Bush and Gore one standard deviation to the Democratic side of the partisanship line increases her probability of voting for Gore to over 0.99. In contrast, increasing by one standard deviation the Democratic propensity of a voter who is indifferent between 9th District US House candidates in 2000 increases her probability of supporting the Democratic candidate to 0.83. Alternatively, increasing by one standard deviation an indifferent voter's partisanship only increases support for State Amendment 7 in 2004 from 0.50 to 0.57.

For our partisanship and dropoff scales we regress race-level discrimination scores on a number of race-level variables plus an indicator variable for 2004. Results are in Table 2, and the variables in the table are largely self-explanatory. For example, the independent variable "Federal" is set to one if and only if the race in question is a federal race.

*** Table 2 about here ***

The estimates in Table 2 show that federal races have high partisanship discrimination; this is evident in the positive (3.10) and significant ($\hat{\sigma} = 0.40$) estimate in the middle column of Table 2. Thus, a voter's latent partisanship strongly predicts how he or she voted in federal races in both 2000 and in 2004. State-level races sometimes discriminate on partisanship as do county races. However, and consistent with Figure 8, judicial races do not discriminate based on partisanship. That is, even if one knows a hypothetical voter's partisanship level, one could not predict how he or she would vote in judicial retention races.

The key results in Table 2 are those pertaining to the 2004 indicator variable. For partisanship, the relevant indicator estimate is positive and significant with a one-tailed hypothesis test at the 0.06 level ($t = 1.59$). However, in the dropoff propensity regression the 2004 indicator slope is easily significantly positive ($t = 3.22$).

Turning first to the latter, the implication of the significant year 2004 indicator in the dropoff model is that there was more dropoff discrimination in general across all races in 2004 compared to 2000. That is, a given voter's dropoff propensity was more predictive of whether or not he or she skipped races in 2004. We would expect to see this if the voting technology used in 2000

(Votomatic punchcards) caused random problems in vote recording, that is, if punchcards made it relatively more difficult for voters to cast valid votes. Table 2 thus presents ballot-level evidence of direct effects: touchscreens facilitate valid voting by voters who want to cast them.

This is more than saying that touchscreens facilitate or perhaps even coerce valid voting for voters writ large. Rather, the fact that there is more dropoff discrimination in 2004 means that touchscreens in Pasco County facilitated valid votes from voters who wanted to cast them and facilitated abstentions from voters who sought to abstain.

The 2004 indicator variable in the partisanship discrimination regression, as we noted, is not significant at conventional confidence levels with a two-tailed hypothesis test though it is nearly significant at the 0.05 level with a one tailed ($p = 0.056$). If the 2004 partisanship indicator is treated as compelling, then it follows that there was more partisanship discrimination in 2004 than in 2000. This means that there was more predictability vis-a-vis voter partisanship, something that is consistent with touchscreens in 2004 having high recording fidelity and causing relatively few voter errors compared to punchcards. Recall that Herrnson et al. (2006) show that the most common type of voter error is voting for the wrong candidate; to the extent that voters are themselves partisan and desire to cast ballots that reflect this, then it follows from Table 2 that touchscreens decrease the most common voter errors.

It is important to note what Table 2 does not imply. Namely, and consistent with 2004 county returns (Card & Moretti 2005), it does not imply that touchscreens have partisan biases themselves and that they lead to excessive votes for Democratic or Republican candidates. Indeed, if they did so, then there would be less partisan discrimination in 2004 compared to 2000 because races that overwhelmingly have Democratic (or Republican) votes by definition discriminate hardly at all.

*** Figure 9 about here ***

One can see additional evidence of 2000-2004 dropoff differences in Figure 9. This figure displays smoothed histograms of precinct-level average dropoff propensities from our IRT dropoff model. The implication of the histograms in the figure is straightforward: dropoff propensities in 2000 were greater than in 2004 and in addition they were also more varied. Indeed, the adoption of touchscreens appears to have had a leveling affect on dropoff propensities across precincts bringing high dropoff propensity areas in 2000 closer to the low dropoff propensity areas in 2004. Because in principle the direct effects of a change to touchscreens should be picked up in race-specific slopes and not in underlying precinct propensities, we interpret this convergence be the result of indirect effects. That is, the direct effect reflects changes in the fidelity with which underlying propensities to drop off are mapped into individual vote decisions (captured by logistic regression coefficients) while the indirect effect is a change in propensity that touchscreens induce.

If there was greater variance in 2000 dropoff propensities, what might be responsible for this? To answer this question we consider linear regressions in which estimates of dropoff propensities and partisanships are regressed on precinct-level demographics. Table 3 contains results.

*** Table 3 about here ***

Our earlier results (in particular, Figure 8 and Table 2) required that 2000 and 2004 dropoff and partisanship scales plus associated discrimination parameters be comparable across election years; as noted, this necessitated an identifying assumption based on 26 Pasco precincts that had constant boundaries between 2000 and 2004. Table 3's results, however, do not depend on comparability across elections, and this latter table is based on dropoff and partisanship scales re-estimated without any comparability constraints. This means that 2000 estimates in Table 3 and 2004 estimates cannot be directly compared.

We see from the two partisanship models in Table 3 that precinct-level partisan and income demographics are highly correlated with precinct partisanships.[11] Note the relatively high R^2 values and significant F statistics. These are useful consistency checks on our results: it would be surprising if, say, the use of touchscreen machines in 2004 eliminated correlations between voter demographics and candidate preferences. Moreover, the signs of the Democratic and income coeffi-

[11]Precinct-level income figures were estimated by overlaying shapefiles of precincts and census block groups. We allocated counts of households by income category in each block group provided in the Census's Summary File 3 (STF3) from the 2000 census to precincts in proportion to the fraction of the area of each block group that was in each precinct which that block group intersected. We then approximated the median income in each precinct using Pareto interpolation (the method of interpolation used by the Census for block group level medians reported in STF3). In total we allocated 160 block groups to 132 precincts. The median number of precincts overlapped by a block group was three; 24 block groups intersected a single precinct, only 12 intersected more than four precincts, and no block group intersected more than seven precincts.

cient estimates in the table are intuitive: they imply that precincts with many Democratic registrants are associated with Democratic voting behavior and that wealthier precincts are associated with Republican voting records.

The dropoff models in Table 3, however, tell a different story. In 2000, precinct dropoff propensities were correlated with precinct demographics, in particular Democratic registration status and income. More Democratic precincts were associated with greater dropoff rates and wealthier precincts with lower dropoff rates. This type of result is consistent with existing literature on the relationship between voter demographics and residual votes.

In contrast, the number of significant dropoff slope coefficients becomes zero 2004, and the 2004 dropoff model F statistic becomes barely statistically insignificant using a conventional 0.05 standard. That is, with touchscreen voting, partisan registration statistics do not explain dropoff rates and precinct wealth likewise has no effect. We conclude, then, that the introduction of touchscreen voting interfered with the relationship that existed in 2000 between voter characteristics and dropoff rates. Insofar as there is no reason to think that, say, relatively poor voters desire to abstain more than wealthy ones, this is evidence of the direct effect of touchscreens.

*** Table 4 about here ***

In Table 4 we decompose the direct and indirect effects of the adoption of touchscreen machines. This table shows predicted participation levels in the retention elections of Justices Lewis in 2000 and Bell in 2004 for the average voter in the precinct with the upper quartile propensity to dropoff in 2000 and 2004. Reading across the diagonal reveals the total change in the rate of predicted participation by our hypothetical voter across this pair of judge elections and is equal to about eight percentage points. Reading down the columns reveals the direct effect of the change to touchscreens. That is, holding dropoff propensities fixed, we find four and six percentage point increases in participation due to changes in the way that touchscreens map underlying propensities into recorded votes. Moving across the columns reveals indirect effects: holding fixed the mapping between dropoff propensity and the decision to participate in these two races and allowing the dropoff propensity to shift, we find two and three percentage point increases in participation.

Thus, it appears that about two-thirds of the shift of the total increase in participation stems from direct effects (changes in the mapping from dropoff propensity into recorded voting choices) and one-third stems from indirect effects (changes in dropoff propensity).

4.3 Neutrality in Low Information Races

Thus far we have presented evidence of better and more voting under touchscreens, that is, of both direct and indirect effects. We now conclude our analysis with a study of Florida Supreme Court and Appellate Court judge retention votes.[12]

Judicial retention races can be of great use to voting technology studies because these races do not feature distinctive campaigns and voters are thought to have little information about them. The low information status of retention races implies that returns in them are not contaminated by candidate and campaign effects. Indeed, there is no expectation that the various Florida judicial retention candidates in 2000 and in 2004 exhibited any special affinities. This is easily contrasted with the presidential candidacies of Al Gore and John Kerry. It is thus safe to analyze judicial retention votes under the assumption that 2000 and 2004 retention votes were structurally equivalent.

*** Figure 10 about here ***

Consider Figure 10, which provides for Pasco County election day voters two different perspectives on the Florida judicial retention votes in 2000 (ten such votes) and in 2004 (seven votes). Voters who participated in, say, the Lewis (2000) race could choose to either support Lewis or not, and the top panel of the figure displays judge candidate support rates. For each 2000 and 2004 judicial retention vote the “Yes” option to retain a particular Florida judge was listed above the “No” option for rejecting.

The key feature of the top panel of Figure 10 is that the minimum judicial support rate (0.595) in 2004 was greater than the maximum support rate (0.571) in 2000; the average judge support rate in 2004 (0.619) was greater than than in 2000 (0.529); and finally, the support rate (0.632) for Bell,

[12]Florida Supreme Court justices and judges on District Courts of Appeal are appointed by the Florida governor and face initial retention votes in the next general election at least one year after appointment. Thereafter they face retention votes every six years.

the first judicial candidate in 2004, was greater than the support rate (0.557) for Lewis in 2000.

We are not aware of any reasons to think that the Florida justices and judges up for retention in 2004 were uniformly more popular or considered more deserving of reappointment than those who faced voters in 2000. Might voting technology, then, have played a role in the swing toward supporting judges? The bottom panel of Figure 10 suggests one mechanism. This panel shows that most Pasco voters either voted for all judges facing retention or voted for none. Note the spike in the two rightmost bars in the figure: in 2004 more voters voted in favor of all judges than in 2000. We have seen that voter participation rates under touchscreens are higher than those observed when voters use punchcards. If a reasonably high percentage of voters support all judges conditional on supporting any, then the tendency of touchscreen technology to encourage voting will also lead to increases in rates by which judges are retained.

*** Table 5 about here ***

Table 5 shows that, conditional on participating in all judicial retention races, the most likely vote outcome is supporting all judges. Furthermore, the relationship between participating in all such races and supporting retention for all judge candidates is stronger in 2000 than in 2004: in the former year's general election, approximately 55% of all voters who participated in all judicial races voted for in favor of all judges, and the comparable figure for 2004 is approximately 62%. Given the large numbers of voters involved, the difference between these two fractions is easily statistically significant at conventional confidence levels.

A potential concern about this finding is that there were different numbers of judges up for retention in 2000 than in 2004. Perhaps voting for all candidates is easier and hence more likely when there are fewer retention possibilities to consider in the first place. If this is true, then it will always be the case that fewer judges up for retention translates to higher judge support rates.

To see that this concern cannot explain all of Table 5 we turn to some across-Florida evidence about changes in voting technology and support rates for the first two judicial retention candidates in 2000 and in 2004. We focus on these two candidates in particular because we cannot recreate Table 5 for other Florida counties; this would require ballot images which we do not have.

*** Table 6 about here ***

Table 6 classifies 65 of the 67 Florida counties into one of five categories based on voting technology in 2000 and voting technology in 2004.[13] In particular, the voting technologies used in Florida in 2000 consisted of punchcards, precinct-count optical scan, and central-count optical scan. The difference between the latter is that, when using precinct-count optical technology, ballots are analyzed within precincts and voters thus have opportunities to fix mistakes. This is called "second chance voting." By November, 2004 Florida counties used either precinct-count optical scan or touchscreen technology.

The counties summarized in Table 6 can be divided into two groups, those that saw ostensibly improved voting technology between 2000 and 2004 and those that did not. The only counties that fall into the latter category are the 22 that used precinct-count optical scan voting technology both in 2000 and 2004. We aggregate these counties, and the z test statistics in Table 6 correspond to difference in proportions tests for the fraction of 2000 voters who supported Lewis and that which supported Bell.

For the 22 counties that used precinct-count optical scan in 2000 and 2004, we cannot reject the hypothesis that support for Lewis and Bell was the same in 2000 and in 2004. In contrast we see for all the Florida county types that had easier to use voting technologies in 2004 than in 2000 statistically significant increases in Bell support over Lewis support. We know that this does not reflect a general Florida-wide surge in support for the former; if this were the case then all rows in Table 6 would contain Bell over Lewis increases.

What exactly is going here on we cannot say. Are the judicial retention effects a function of the fact that the "Yes" option is listed before the "No" option? We do not have leverage on this because all of our judge votes have the same "Yes" and "No" orderings. Similarly, would the effects be as strong if judicial retention races appeared in a different ballot location? Again we do not know.

Regardless, we believe that the judicial effects described here warrant additional scrutiny by both academics and by policymakers, some of whom are rushing to embrace new voting technologies. If the dominance of indirect touchscreen effects over direct touchscreen effects in judicial

[13] Martin County, which used lever voting machines in 2000, and Union which used paper ballots in 2000, are ignored.

retention races is systematic, then it would be hard to imagine that there are not other ballot locations or race types that feature this as well. Normatively, as we noted in the introduction, situations in which indirect effects are stronger than direct effects are troubling. The modal judicial retention race in Florida is not close, but other types of races, some of which are low information contests, are regularly decided by only a few percentage points.

5 Conclusion

We motivated our research on touchscreen voting by considering two separate effects of this type of technology. One, which we called the direct effect, is manifested in high recording fidelity. The other, which we called an indirect effect, reflects increased voter participation in races that ordinarily would have been skipped. These effects are generic and, we believe, are useful for conceptualizing voting technology changes in general.

We have found evidence of both direct and indirect effects in our study of Pasco County, Florida's transition from punchcard voting in 2000 to touchscreen voting in 2004. Namely, we show that touchscreens are associated with higher participation rates in both top and down-ballot races, and this finding appears across the many ways that we have characterized voter participation.

Moreover, we have uncovered evidence of direct effects, and we believe that ours is the first observational study to have shown that touchscreens appear to reduce voter errors. Notably, among Pasco County voters within-ballot predictability of voter dropoff decisions and candidate choices was higher under touchscreens in 2004 than under punchcards in 2000. This suggests that the former are higher fidelity vote recorders than are the latter.

One corollary of our results is that residual vote rates, the cornerstone of much of the extant literature on voting technology, simply do not adequately capture the way that voters interact with voting equipment. Residual vote rates cannot tell us anything about voter error rates, within-ballot predictability, and so forth. To the extent that measuring, with an eye toward lowering, voter error rates is a key feature of voting technology research, then studies need to move away from a reliance on residual vote rates to the measures described here and in Herrnson et al. (2006).

The results we have offered are based on a study of a single county, and this should give rise to two caveats. One is that the county we have studied is excessively white compared to other United States counties. Because racial and socioeconomic status are correlated in this country, and because voter errors are disproportionately common among low socioeconomic status voters (Herrnson et al. 2006), our results on the effects of touchscreens are conservative. That is, a representative American county that transitioned from punchcards to touchscreens would have observed greater effects than those we have documented. The fact that Pasco County is excessively white, then, is not a problem.

The second caveat to our results is based on the fact that Pasco's transition to touchscreens took place after the 2000 presidential election debacle. Could this timing exaggerate the effects we have uncovered? Certainly, but this is no more true of our results than of other observational studies of voting technology that use post-2000 election returns. Furthermore, our across-Florida results on judicial retention races shows that participation rates in these votes did not change among counties that did not incur voting technology changes between 2000 and 2004. There should be no doubt, then, that touchscreens have complicated effects on voting behaviors, effects that will become more clear with additional experimental work on voting technology and with the passage of the 2008 general election.

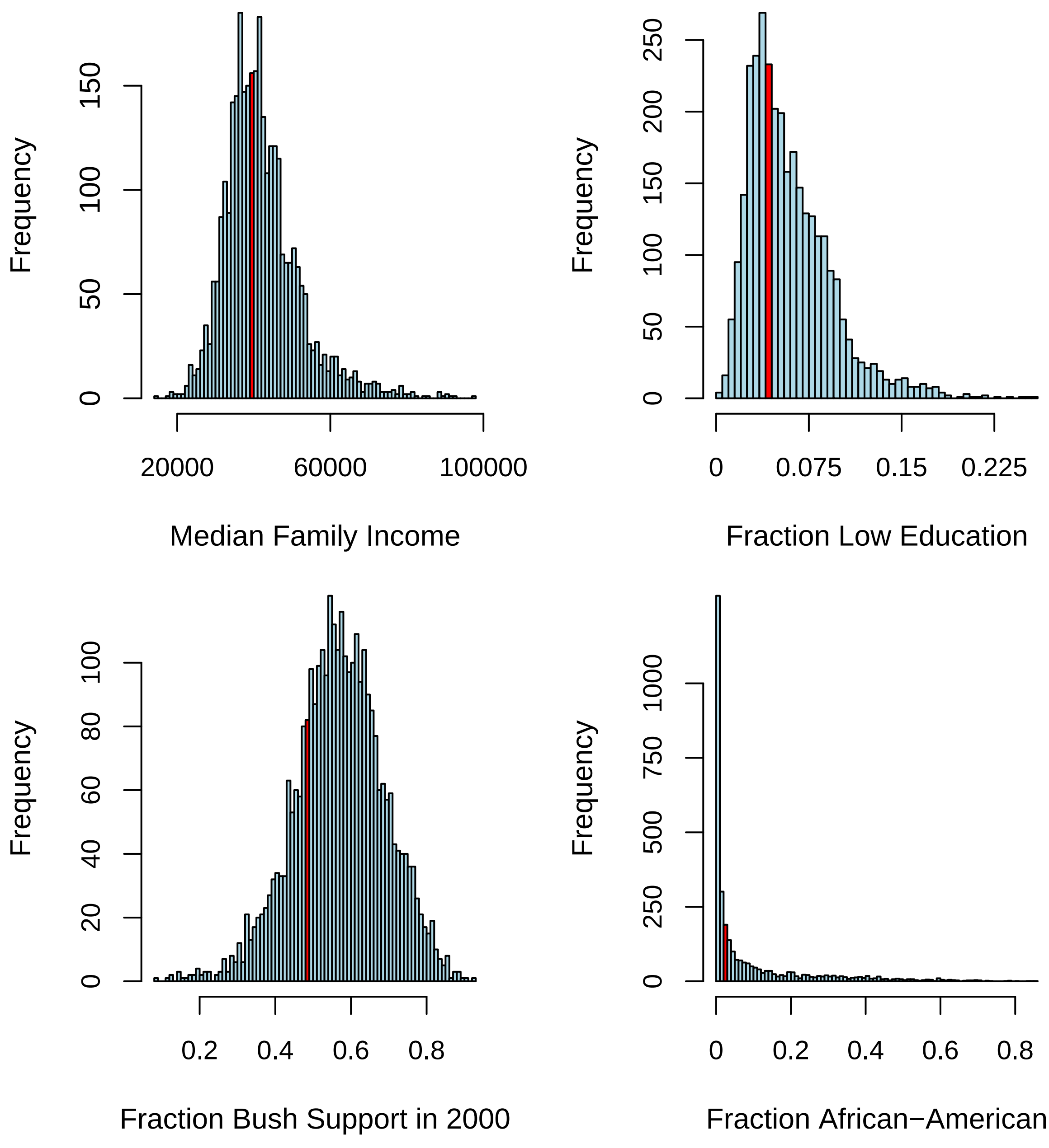
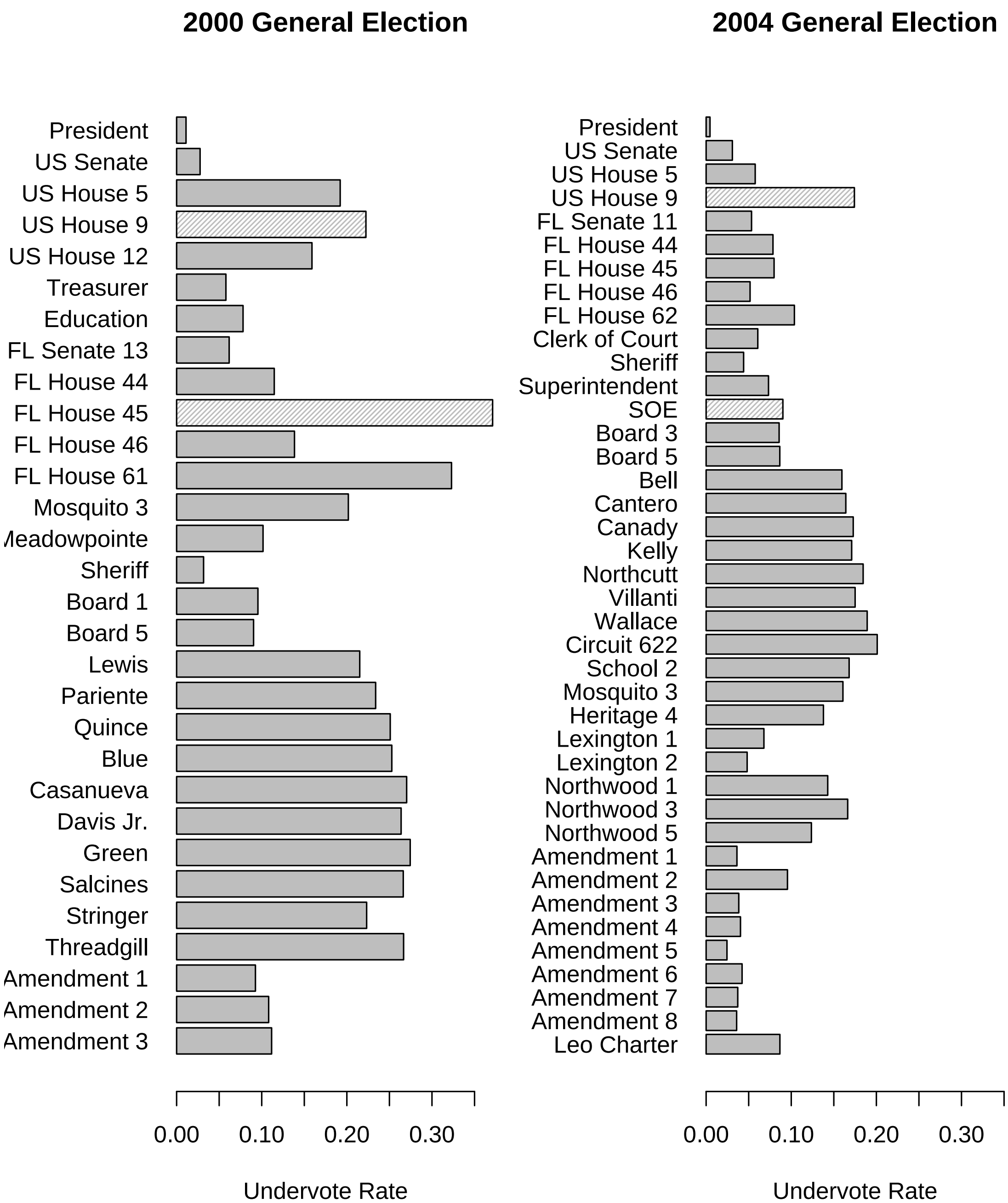


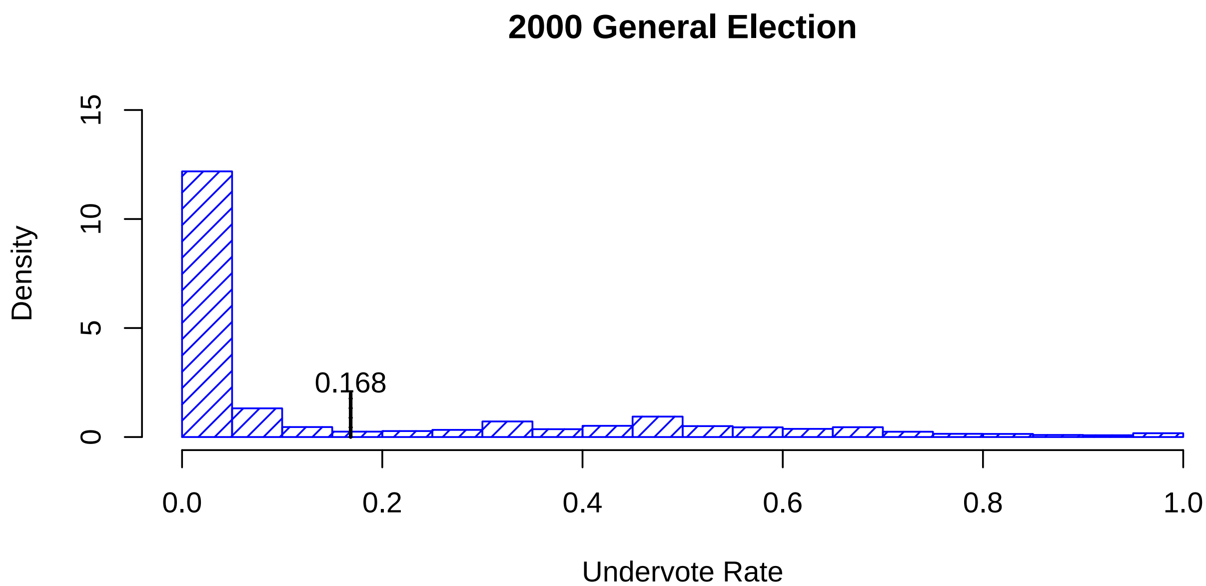
Figure 1: Distributions of Four Characteristics of United States Counties with Pasco County in red.

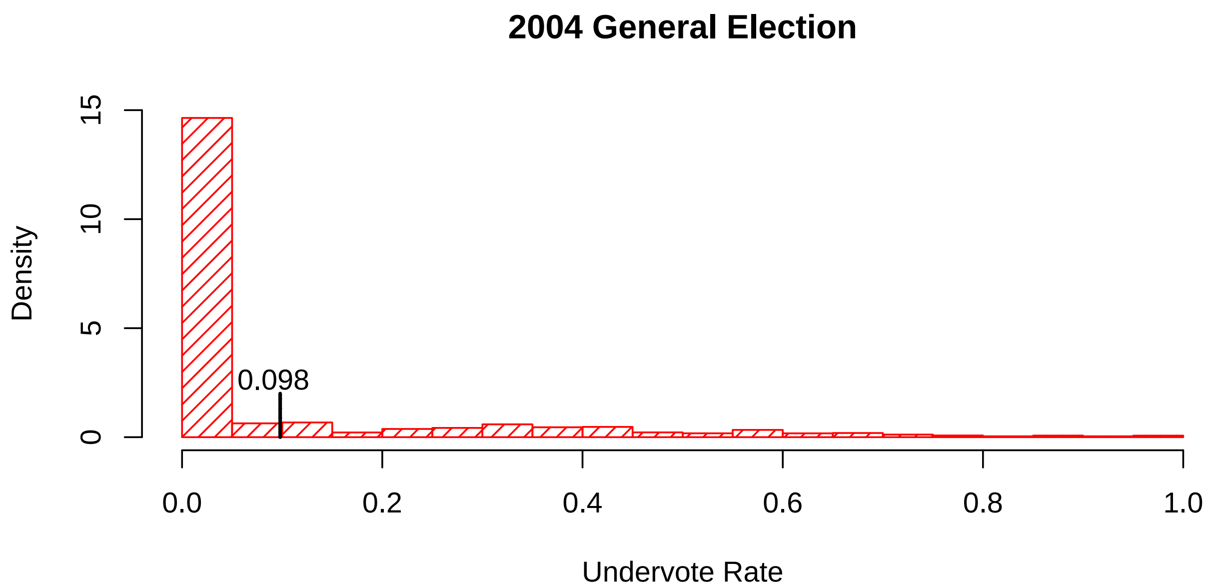
Figure 2: Undervote Rates by Race



Note: Based on 131,655 and 127,526 election day voters in 2000 and 2004, respectively. Shaded races were not contested by multiple, major party candidates. For 2000, Treasurer and Secretary of Education were Florida state offices in 2000; Mosquito 3 refers to a seat on a county-wide mosquito control board; and, County Board of Commissioners, Districts 1 and 5, are county seats as well. Names at the end of the 2000 ballot refer to Florida state judges facing retention votes. This holds for 2004, and in addition SOE denotes Supervisor of Elections. The Lexington, Northwood, and Leo votes in 2004 were for local voters only.

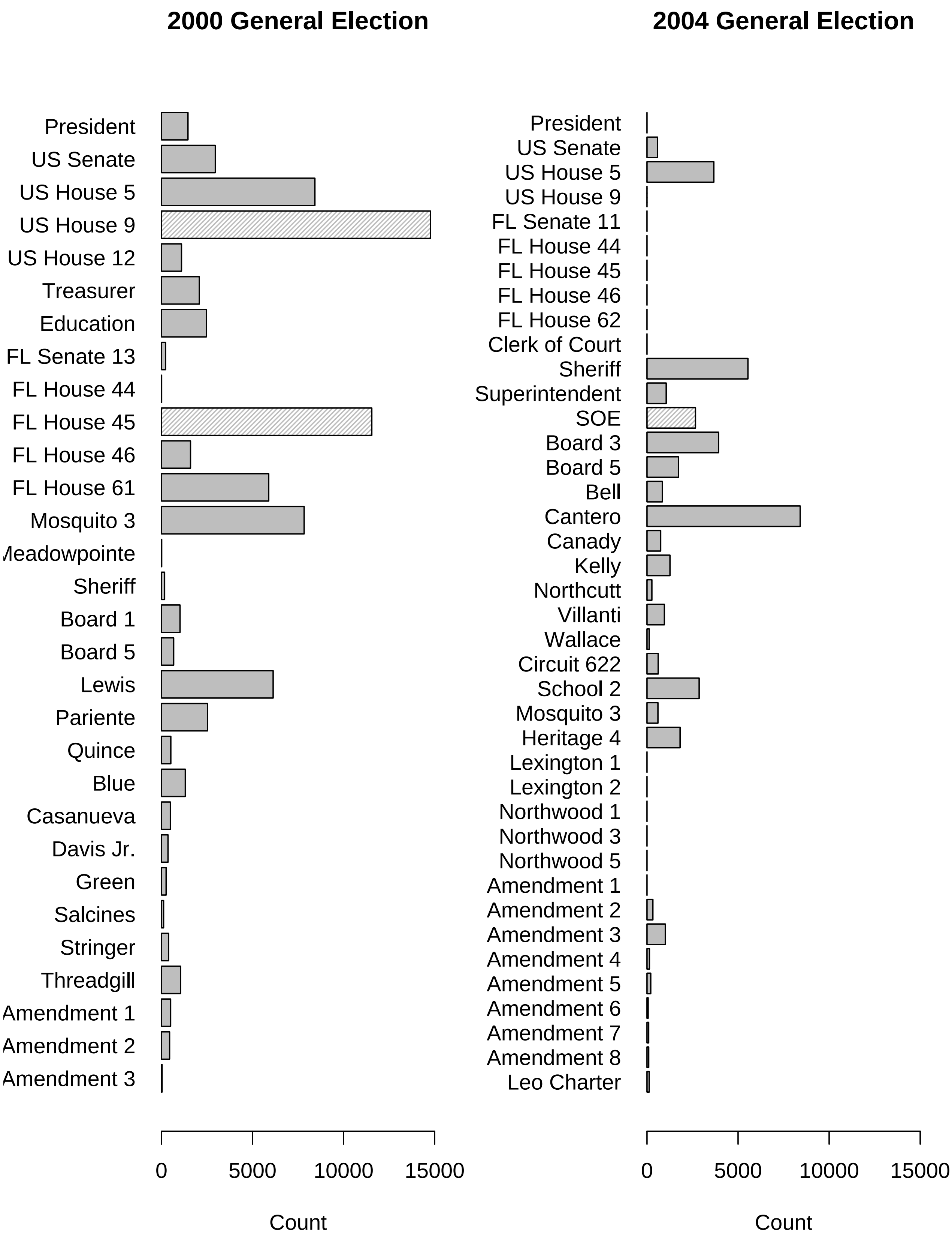
Figure 3: Histograms of Undervote Rates by Ballot





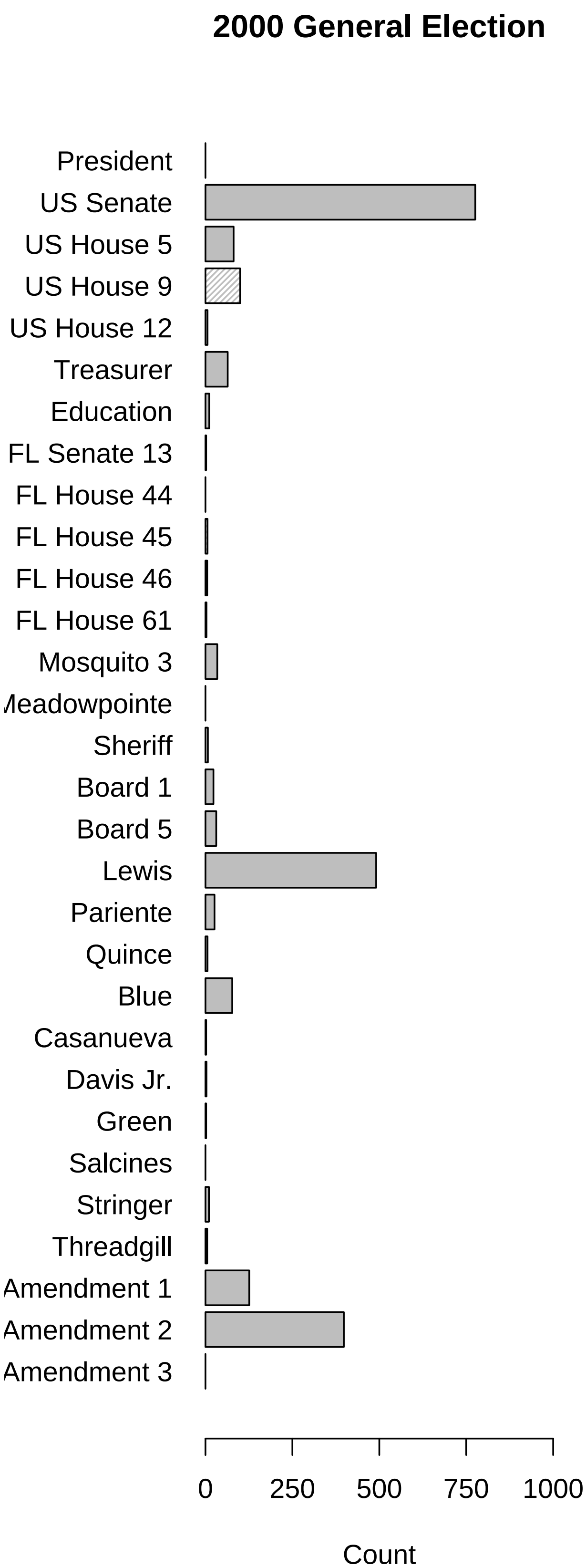
Note: Residual vote rates, based on 131,655 election day voters in 2000 and 127,526 election day voters in 2004, are normalized based on the number of possible valid votes that each voter could cast. The lines in each histograms denote sample means; sample standard errors are 0.000675 and 0.000543 in 2000 and 2004, respectively.

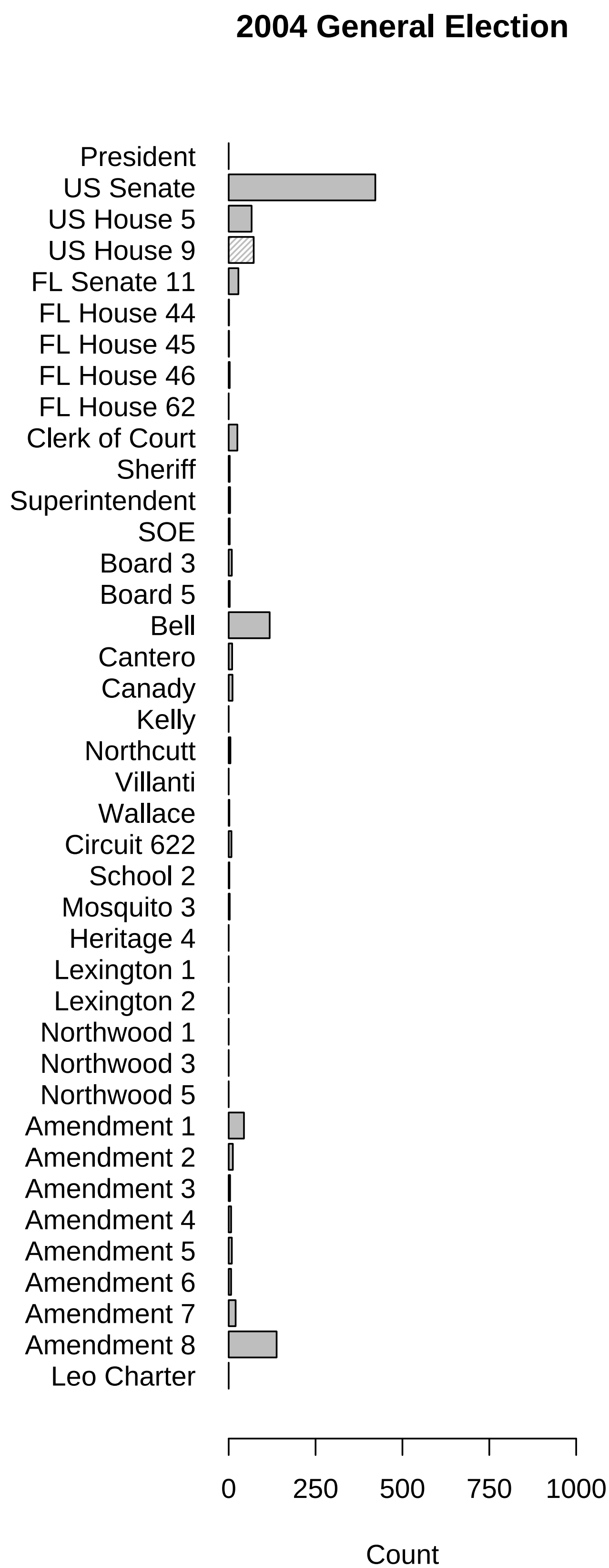
Figure 4: Location of First Undervote among Election Day Voters in 2000 and 2004



Note: Based on 131,655 election day voters in 2000 and 127,526 election day voters in 2004. Shaded races were not contested by multiple, major party candidates.

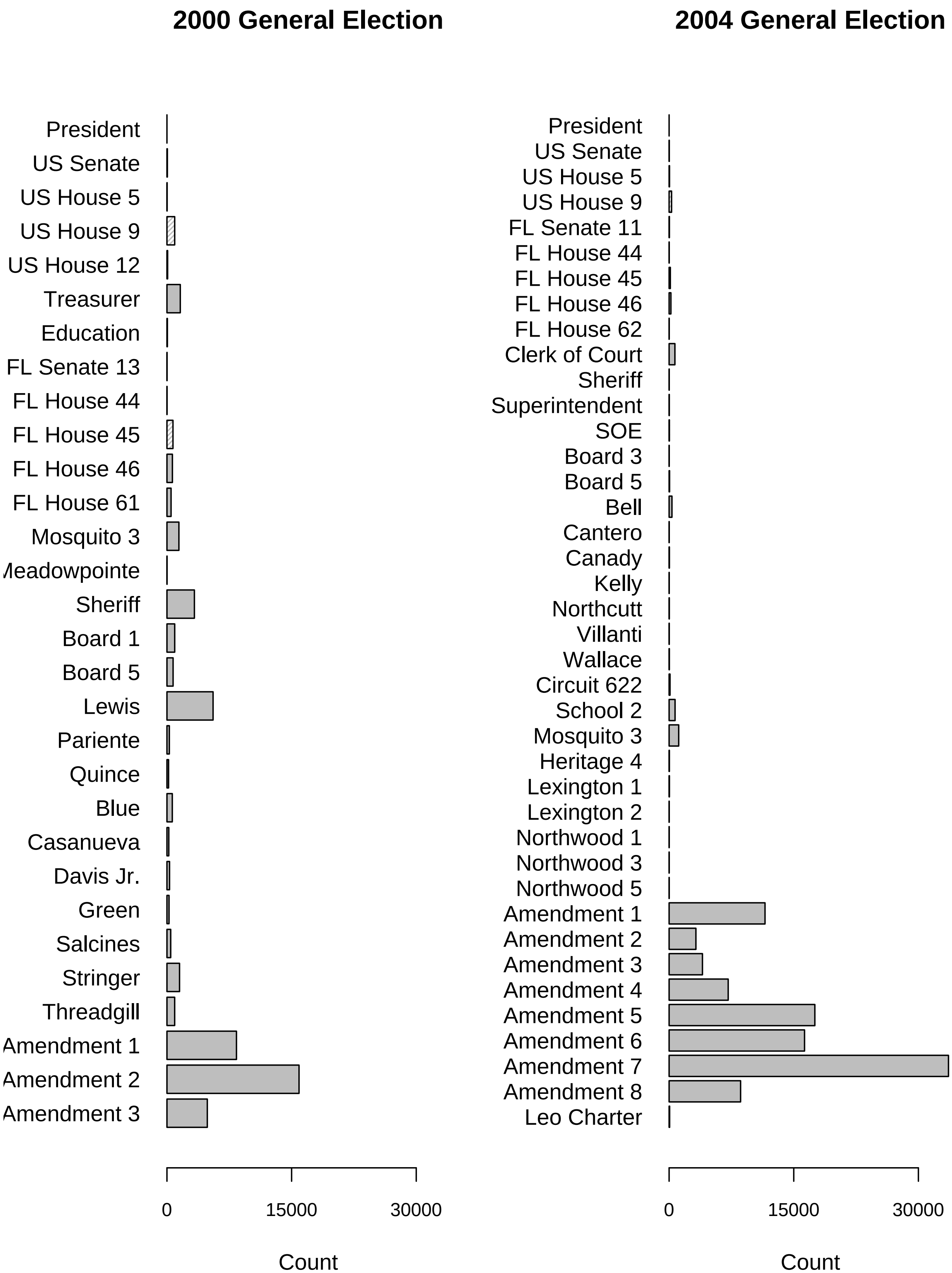
Figure 5: Ballot Location at which Rolloff Begins for Election Day Voters in 2000 and 2004





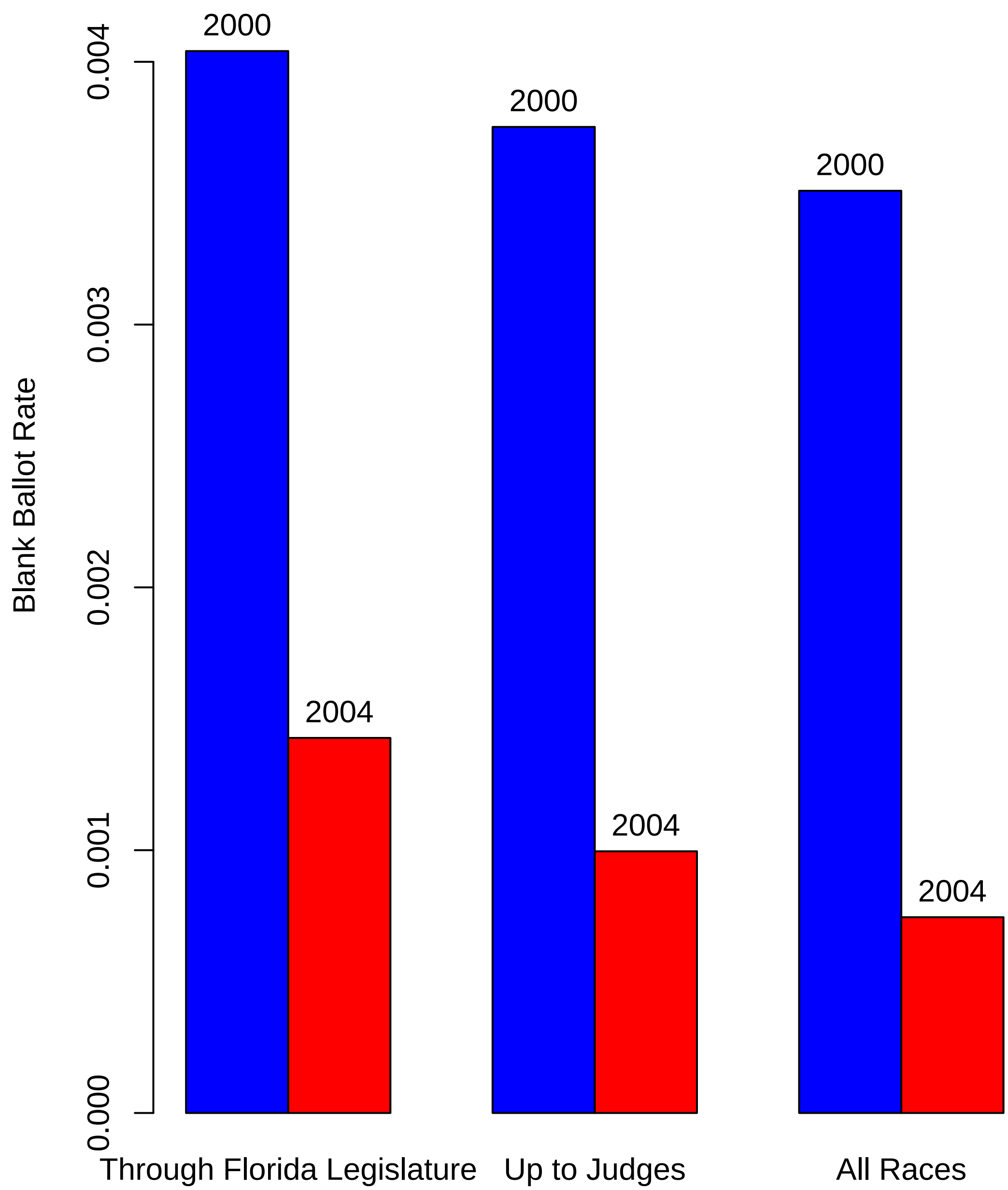
Note: Based on 131,655 election day voters in 2000 and 127,526 election day voters in 2004. Shaded races were not contested by multiple, major party candidates.

Figure 6: Ballot Locations after which Top Choice is Made Uniformly by Election Day Voters in 2000 and 2004



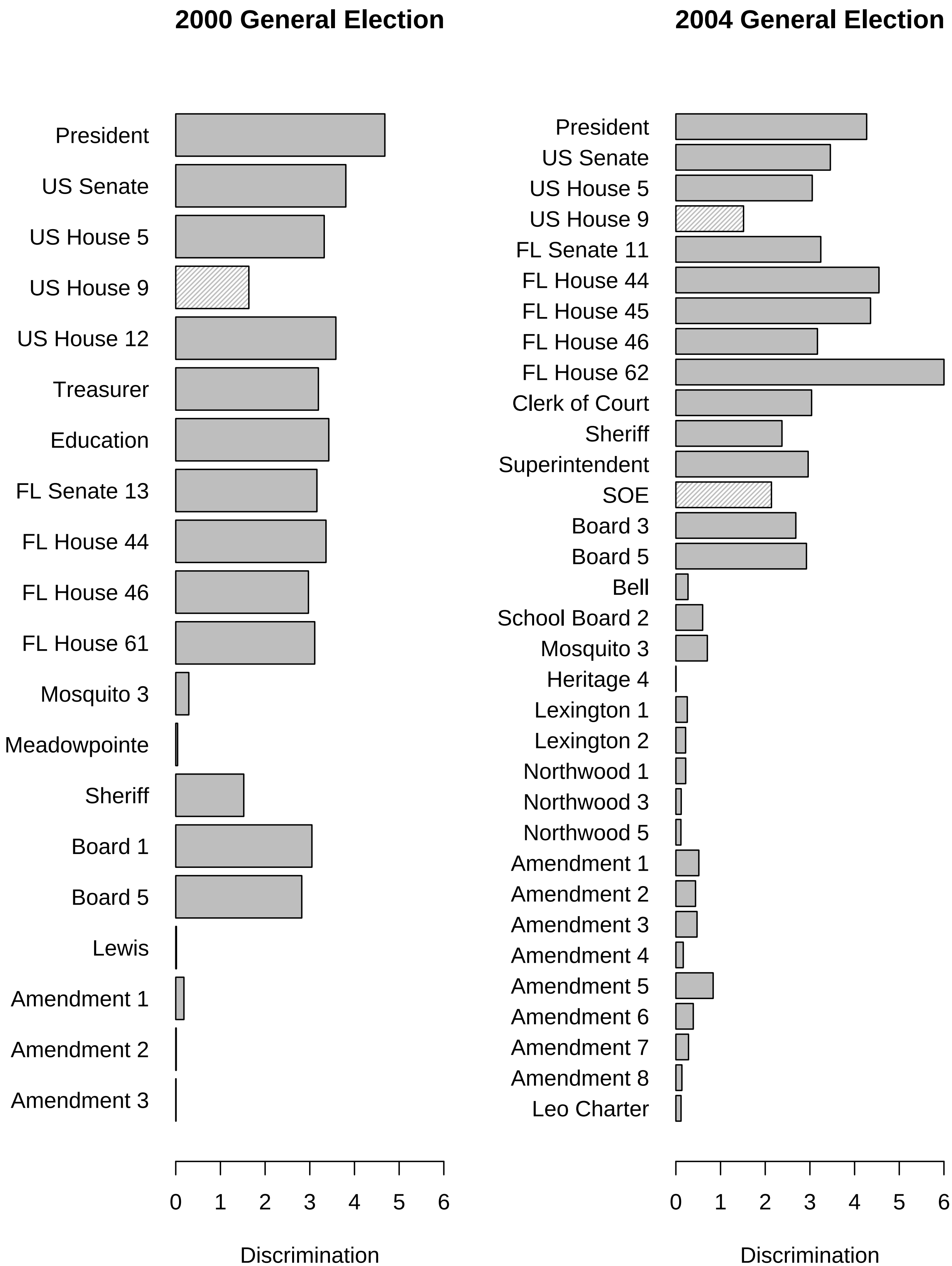
Note: Based on 131,655 election day voters in 2000 and 127,526 election day voters in 2004. Shaded races were not contested by multiple, major party candidates.

Figure 7: Blank Ballot Rates among Election Day Voters in 2000 and 2004



Note: Based on 131,655 election day voters in 2000 and 127,526 election day voters in 2004.

Figure 8: Partisanship Discrimination in 2000 and 2004 Races



Note: Each bar represents the discrimination score for a given race. Shaded races were not contested by multiple, major party candidates.

Figure 9: Dropoff Propensities by Precinct in 2000 and 2004

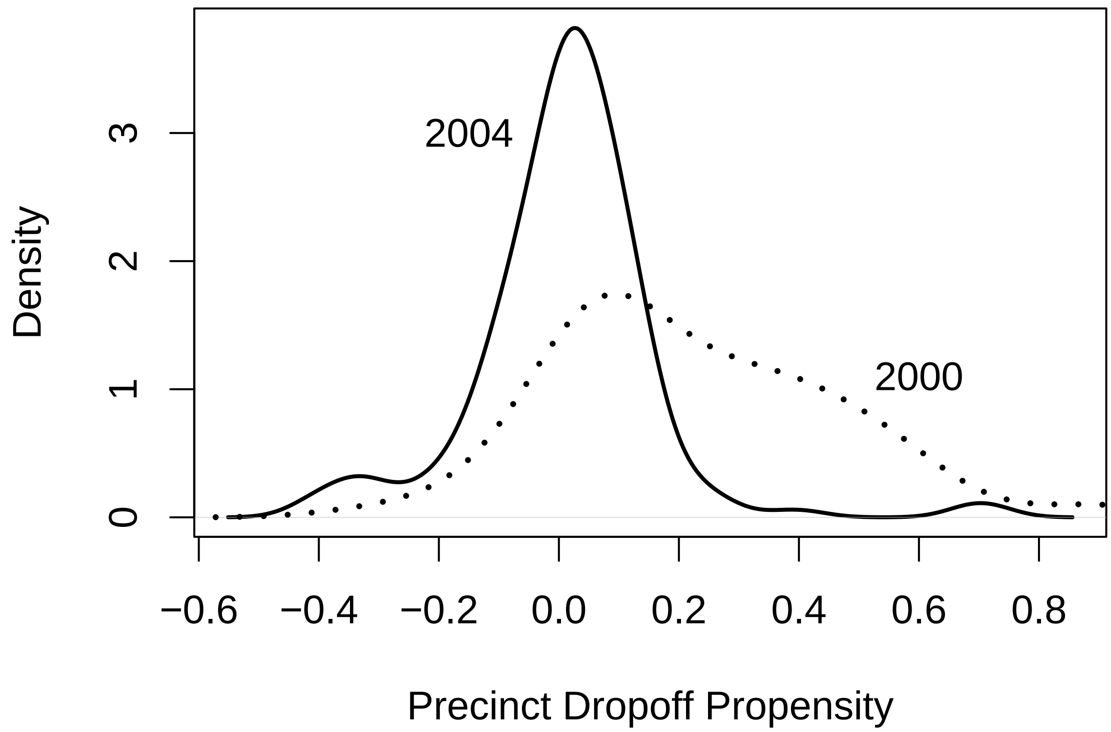
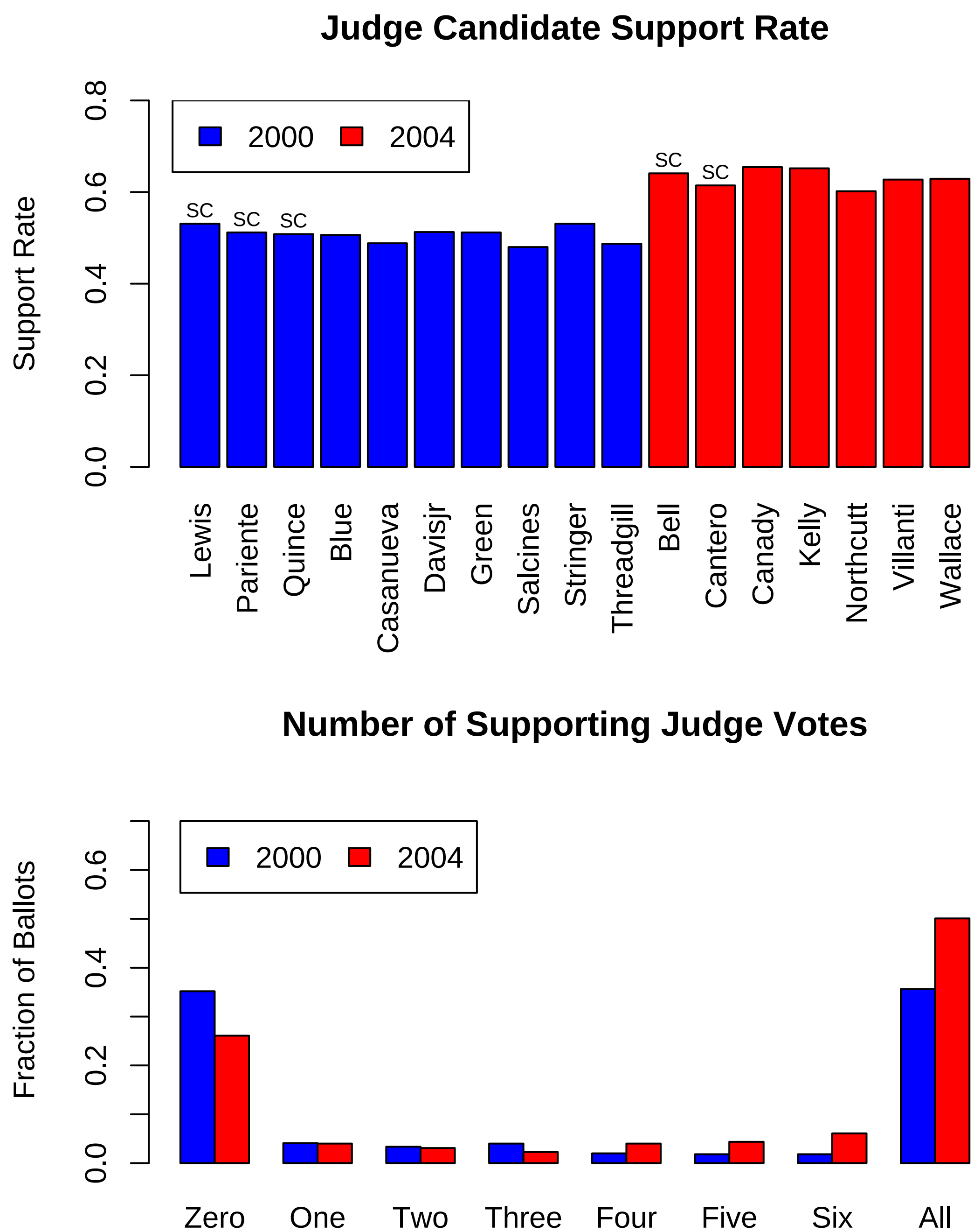


Figure 10: Judge Retention Vote Statistics among 2000 and 2004 Election Day Voters



Note: "SC" in the upper panel denotes Florida Supreme Court justices; all other judges for whom votes are reported are members of a Florida District Court of Appeal.

Table 1: Two Views on Invalid Vote Rates among Comparable Races

Race	Undervote Rate in 2000	Undervote Vote Rate in 2004	Difference	z
President	0.0111	0.00456	-0.00650	-18.7
U.S. Senate	0.0276	0.0309	0.00328	4.95
Pasco County Sheriff	0.0317	0.0440	0.0124	16.5
First Judicial Retention Vote	0.215	0.160	-0.0556	-36.2
First Amendment	0.0926	0.0362	-0.0564	-58.3

(a) Comparison of Undervote Rates

Race	Residual Vote Rate in 2000	Residual Vote Rate in 2004	Difference	z
President	0.0265	0.00456	-0.0220	-45.0
U.S. Senate	0.0342	0.0309	-0.00337	-4.83
Pasco County Sheriff	0.0331	0.0440	0.0109	14.4
First Judicial Retention Vote	0.220	0.160	-0.0606	-39.3
First Amendment	0.0935	0.0362	-0.0573	-59.0

(b) Comparison of Residual Vote Rates

Note: Undervote and residual vote rates are based on 131,655 election day voters in 2000 and 127,526 election day voters in 2004. The difference between undervote and residual vote rates in 2000 is that residual vote rates include overvote rates. Overvotes were not permitted by Pasco County's touchscreen machines in 2004.

Table 2: Discrimination as a Function of Election Year and Office Type

Variable	Partisanship model	Dropoff model
Year 2004	0.46	0.58*
	(0.29)	(0.18)
Federal	3.10*	-0.81*
	(0.40)	(0.24)
State	3.49*	0.03
	(0.30)	(0.27)
County	1.70*	0.34
	(0.32)	(0.18)
Judge	-0.03	0.05
	(0.12)	(0.14)
Local	-0.26*	-0.43*
	(0.10)	(0.19)
Constant	-0.05	2.20*
	(0.22)	(0.21)
$\hat{\sigma}$	0.85	0.55
F	25.3	6.72
R^2	0.77	0.46
N	53	54

Note: Heteroskedastic-consistent standard errors are in parentheses, and * *denotes* $p < 0.05$.

Table 3: Regressions of Estimated Partisanships and Dropoff Propensities by Precinct.

	Partisanship Model		Dropoff Model	
	2000	2004	2000	2004
Intercept	-0.103	1.22*	0.410	0.490
	(0.433)	(0.420)	(0.579)	(0.385)
Fraction Democratic	-1.41*	-1.87*	0.762*	-0.425
	(0.406)	(0.326)	(0.361)	(0.239)
Median Family Income	0.142*	0.129*	-0.101*	0.0114
	(0.0258)	(0.0289)	(0.0400)	(0.0251)
Percent Grade School	0.591	0.447	-0.328	-0.244
	(0.794)	(0.724)	(1.06)	(0.704)
Percent High School	0.344	-0.988*	-0.473	-0.507
	(0.701)	(0.442)	(0.694)	(0.454)
Percent Bachelors Degree	-1.28	-0.230	1.20	-0.419
	(0.755)	(0.425)	(0.711)	(0.451)
R^2	0.570	0.603	0.177	0.0814
$F\ p-value$	$<$0.001	$<$0.001	$<$ 0.001	0.0283

*Note: Heteroskedastic-consistent standard errors are in parentheses, and * denotes $p < 0.05$. Fraction Democratic refers to the Democratic share of the Democratic and Republican vote share in a precinct. And, Median Family Income is measured in ten of thousands of 1999 dollars.*

Table 4: Decomposing Changes in Voter Participation in Two Judicial Retention Races

	Lower quartile precinct in:	
	2000	2004
Lewis (2000)	0.87	0.91
Bell (2004)	0.93	0.95

Table 5: Participation and Yes/No Choices in Judicial Retention Votes among Pasco County Election Day Voters, 2000 and 2004

Number of Votes	Number of Supporting Retention Votes											
	0	1	2	3	4	5	6	7	8	9	10	Total
0	20286											20286
1	426	2186										2612
2	311	416	2176									2903
3	405	454	1066	5151								7076
4	52	122	252	405	1094							1925
5	85	55	101	160	201	717						1319
6	50	43	56	92	119	168	391					919
7	92	43	54	74	96	115	158	524				1156
8	226	120	109	115	134	164	242	283	1348			2741
9	460	190	191	195	217	298	320	369	496	2902		5638
10	16548	1850	1879	1770	1418	1880	2151	2738	3332	4228	47286	85080
	38941	5479	5884	7962	3279	3342	3262	3914	5176	7130	47286	131655

(a) 2000 General Election

Number of Votes	Number of Supporting Retention Votes								
	0	1	2	3	4	5	6	7	Total
0	17048								17048
1	270	1820							2090
2	292	232	1510						2034
3	78	97	164	693					1032
4	166	66	150	189	1085				1656
5	75	51	62	93	144	560			985
6	195	91	99	98	166	230	1104		1983
7	16043	3457	2273	2348	3225	4265	6279	62808	100698
	34167	5814	4258	3421	4620	5055	7383	62808	127526

(b) 2004 General Election

Table 6: Changes in Florida County Voting Technology and First Judicial Retention Outcome

Technology Type	Counties	Lewis Fraction	Bell Fraction	Difference	Votes in 2000	Votes in 2004	z
Punchard $\rightarrow$ DRE	12	0.730	0.743	0.0134	2315379	3035552	35.0
Punchard $\rightarrow$ PC Optical	12	0.682	0.721	0.0394	404410	582890	42.2
PC Optical $\rightarrow$ PC Optical	17	0.714	0.714	-0.000433	1164662	1486191	-0.775
CC Optical $\rightarrow$ PC Optical	22	0.684	0.701	0.0170	617907	827006	21.9
CC Optical $\rightarrow$ DRE	2	0.709	0.732	0.0229	124620	166369	13.7

Note: Lewis was the first judicial retention candidate in the 2000 general election and Bell was the first in 2004. There are 67 counties in Florida but Martin and Union Counties do not contribute to this table. The voting technologies used in them are described in the column denoted "Technology Type" where the first technology noted was used in 2000 and the second in 2004. Moreover, "DRE" refers to touchscreen machines; "Optical" to optical scan voting technology; "PC Optical" to precinct-count optical scan; and "CC Optical" to central count optical scan.

References

Ansolabehere, Stephen. 2002. "Voting Machines, Race, and Equal Protection." *Election Law Journal* 1(1):61–70.

Ansolabehere, Stephen & Charles Stewart III. 2005. "Residual Votes Attributable to Technology." *Journal of Politics* 67(2):365–389.

Berinsky, Adam J. 2005. "The Perverse Consequences of Electoral Reform in the United States." *American Politics Research* 33(4):471–491.

Brady, Henry E., Justin Buchler, Matt Jarvis & John McNulty. 2001. "Counting All the Votes: The Performance of Voting Technology in the United States." Report published by the Survey Research Center and Institute of Governmental Studies, University of California, Berkeley.

Card, David & Enrico Moretti. 2005. "Does Voting Technology Affect Election Outcomes? Touch-Screen Voting and the 2004 Presidential Election." NBER Working Paper 11309.

Gerber, Elisabeth R. & Jeffrey B. Lewis. 2004. "Beyond the Median: Voter Preferences, District Heterogeneity, and Political Representation." Forthcoming, *Journal of Political Economy*.

Hanmer, Michael J. & Michael W. Traugott. 2004. "The Impact of Voting by Mail on Voter Behavior." *April* 32(4):375–405.

Herrnson, Paul S., Benjamin B. Bederson, Richard G. Niemi, Frederick Conrad, Michael J. Hanmer & Michael Traugott. 2006. "The Not So Simple Act of Voting: An Examination of Voter Errors with Electronic Voting." Unpublished working paper, University of Maryland, College Park.

Herrnson, Paul S., Owen G. Abbe, Peter L. Francia, Benjamin B. Bederson, Bongshin Lee, Robert M. Sherman, Fred Conrad, Richard G. Niemi & Michael Traugott. 2005. "Early Appraisals of Electronic Voting." Forthcoming, *Social Science Computing Review*.

Herron, Michael C. & Jasjeet S. Sekhon. 2005. "Black Candidates and Black Voters: Assessing the Impact of Candidate Race on Uncounted Vote Rates." *Journal of Politics* 67(1):154–177.

Herron, Michael C. & Jeffrey B. Lewis. 2005. "Did Ralph Nader Spoil a Gore Presidency? A Ballot-Level Study of Green and Reform Party Voters in the 2000 Presidential Election." Unpublished working paper, Dartmouth College.

Jacob, Brian A. & Steven D. Levitt. 2003. "Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating." *The Quarterly Journal of Economics* 118(3):843–877.

Knack, Stephen & Martha Kropf. 2002. "Who uses inferior voting technology?" *PS: Political Science and Politics* 35(3):541–8.

Lewis, Jeffrey B. 2001. "Estimating Voter Preference Distributions from Individual-Level Voting Data." *Political Analysis* 9(3):275–297.

Mebane, Jr., Walter R. 2003. "The Wrong Man is President! Overvotes in the 2000 Presidential Election in Florida." Forthcoming, *Perspectives on Politics*.

Merzer, Martin. 2001. *The Miami Herald Report: Democracy Held Hostage*. New York: St. Martin's Press.

Nichols, Stephen M. 1998. "State Referendum Voting, Ballot Roll-off, and the Effect of the New Electoral Technology." *State and Local Government Review* 30(2):106–117.

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

Poole, Keith T. & Howard Rosenthal. 1997. *Congress: A Political–Economic History of Roll Call Voting*. New York: Oxford University Press.

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

Sinclair, Robert C., Melvin M. Mark, Sean E. Moore, Carrie A. Lavis & Alexander S. Soldat. 2000. "An electoral butterfly effect." *Nature* 408:665–666.

Stewart III, Charles. 2004. "The Reliability of Electronic Voting Machines in Georgia." Caltech/MIT Voting Technology Working Paper 23.

Stewart III, Charles. 2005. "Residual Vote in the 2004 Election." Caltech/MIT Voting Technology Working Paper 25.

The Washington Post. 2001. *Deadlock: The Inside Story oF America's Closest Election*. New York: PublicAffairs.

Tomz, Michael & Robert P. van Houweling. 2003. "How Does Voting Equipment Affect the Racial Gap in Voided Ballots?" *American Journal of Political Science* 47(1):46–60.

United States Commission on Civil Rights. 2001. "Election Reform: An Analysis of Proposals and the Commission's Recommendations for Improving America's Election System.".

Wand, Jonathan N., Kenneth W. Shotts, Jasjeet S. Sekhon, Walter R. Mebane, Jr., Michael C. Herron & Henry E. Brady. 2001. "The Butterfly Did It: The Aberrant Vote for Buchanan in Palm Beach County, Florida." *American Political Science Review* 95(4):793–810.