Overvoting and Representation:

An examination of overvoted presidential ballots

in Broward and Miami-Dade counties[1]

Michael C. Herron[2]   Jasjeet S. Sekhon[3]

September 28, 2001

[2]Assistant Professor, Department of Political Science, Northwestern University. Herron worked on this paper while a Postdoctoral Research Fellow, Center for Basic Research in the Social Sciences, Harvard University. Phone: 847-676-2236. Address: 601 University Place, Evanston, IL 60208–1006 (m-herron@northwestern.edu).

[3]Assistant Professor, Department of Government, Harvard University. Phone: 617-496-2426. Address: 34 Kirkland Street, Cambridge, MA 02138 (jsekhon@fas.harvard.edu)

Abstract

The closeness of the 2000 presidential election, especially in the state of Florida, has drawn attention to the importance of voting anomalies caused by ballot design, voting technology, and voter errors. In this paper we focus on a particular type of voter error: casting multiple votes for president on a single ballot. Ballots cast in this way are said to contain presidential overvotes, and we examine overvoting patterns in Broward and Miami-Dade Counties, two large and prominent counties in Florida. Using a dataset which contains electronic images of all ballots cast in these counties for the 2000 election, we identify several definitive patterns among overvoted ballots. First, we show that ballots with overvotes on non-presidential races were more likely to contain presidential overvotes compared to ballots with no overvotes elsewhere. Second, we show that ballots with presidential overvotes appear to have been cast by Democratically-inclined individuals and that Al Gore, the Democratic presidential candidate in 2000, appears on a disproportionate number of these ballots. Third and finally, we show that Broward and Miami-Dade precincts with large numbers of blacks, Hispanics, and registered Democrats tended to have high presidential overvoting rates. Overall, the evidence we present implies that a disproportionate fraction of the presidential overvotes cast in Broward and Miami-Dade Counties in the 2000 election were produced by Democrats and this diminished the vote total of Al Gore.

1 Introduction

The closeness of the 2000 presidential election, especially in the state of Florida, has drawn attention to the importance of voting anomalies caused by ballot design, voting technology, and voter errors. In this paper we focus on a particular type of voter error: casting multiple votes for president on a single ballot. Ballots with more than one vote for president are said to contain presidential overvotes, and these ballots are invalid insofar as the presidential race is concerned. Ballots with presidential overvotes are functionally equivalent to those with presidential undervotes, i.e., equivalent to ballots that contain no presidential vote at all.[1] Very little is known about the prevalence and causes of voting anomalies in general and about overvoting in particular. For example, the national overvoting rate for the 2000 presidential election is not known nor is the degree to which this rate has varied over time, by region, or by socioeconomic group.[2] Overvoting statistics for Florida, however, are known. They show, in fact, that George W. Bush won Florida, a pivotal state in the 2000 Electoral College, by significantly fewer votes than there were presidential overvotes cast in the state.[3]

In particular, there were 111,261 presidential overvotes among the 6,138,765 total ballots cast in Florida, and this constitutes a state-wide overvote rate of 1.8%. The final, certified George W. Bush–Al Gore margin in Florida was a scant 537 votes; thus, the number of presidential overvotes in Florida was approximately 207 times as large as the margin of victory that ultimately determined who became president of the United States in January, 2001. The Florida overvote rate in the 2000 election may have been unusually high, however. In the state of Texas, for instance, only approximately 0.23% of ballots cast in the 2000

---

[1]While we distinguish between presidential undervotes and overvotes, some counties and reporting units in the United States use the term "undervote" to refer to an invalid presidential vote regardless of why the vote is invalid.

[2]Nationally, approximately 2% of ballots cast in recent presidential elections do not contain valid presidential votes (The Caltech/MIT Voting Technology Project 2001). This means that the typical overvote rate is less than this approximate figure.

[3]The winner of an American presidential election is determined by electors who represent states. Hence, what determines whether overvoting was a key issue in 2000 is whether Bush-Gore margins at the state level were smaller than state-level overvote counts. The number of overvotes at the national level is not substantively that important. Details on the American presidential election system can be found in Longley and Peirce (1999).

election contained presidential overvotes. This percentage rises to 0.4%, or 14,049 total ballots, if one disregards those ballots cast using a voting system which does not permit multiple votes for a single office.

The Florida and Texas state-wide presidential overvote rates—one which falls below 1%, one which lies above—appear quite significant given the closeness of historical presidential elections. There were, notably, five states in the 2000 presidential race with Bush-Gore vote share margins of less than 1%: Florida, Iowa, New Mexico, Oregon, and Wisconsin. In 1996 the state of Kentucky produced a Bill Clinton-Robert Dole margin of less than 1%, and in 1992 two states, Georgia and North Carolina, had Bill Clinton-George Bush vote share margins of less than 1%. Had some portion of overvoted presidential ballots been cast in a valid way in the past three presidential elections, the electoral college balance might have been affected and, at least in 2000, the winner of the overall presidential race could have been different.

This cursory review of the three most recent American presidential elections implies that overvoting can affect citizen representation in its broadest sense. Elected officials—both executive and legislative—have a major role in federal policy making. If overvoting affects the identity of these officials, then it follows that this type of voter behavior can influence the policies chosen by the federal government. To foreshadow results described later, a group of voters which overvotes at a relatively high rate risks disenfranchisement and a loss of the extent to which the group's interests are represented in government.

A key consideration in an analysis of the extent to which overvoting affects representation is the question, "Who overvotes?" We explore this question by focusing on two large and prominent counties in Florida, Broward and Miami-Dade, in the context of the most recent general election. According to 2000 census figures, Miami-Dade County accounts for approximately 14% of Florida's population and is the largest county in the state. Broward County is the second most populous Florida county with around 10% of the state.[4] Thus,

---

[4]The third most populous county is Palm Beach County, but presidential voting in this county in 2000 was complicated by the county's unusual ballot format (Wand, Shotts, Sekhon, Mebane Jr., Herron and Brady 2001). Overvoting in Palm Beach County is discussed in "Over-votes Cost Gore the Election in Florida," *Palm Beach Post*, March 11, 2001.

even though we focus on only two counties, our analysis covers almost a quarter of Florida's population. Importantly, invalid presidential voting rates are comparable across Broward and Miami-Dade Counties because these two localities used identical voting systems, punch card ballot machines with central tabulation, in the 2000 general election.[5] Such machines are considered relatively error-prone and do not alert the voter if she has cast an invalid vote (The Caltech/MIT Voting Technology Project 2001).[6]

Moreover, both Broward and Miami-Dade Counties contain significant numbers of racial minorities. This is a noteworthy feature of the two counties because, in the aftermath of the 2000 election, it has been argued that presidential overvoting and invalid voting more generally have disproportionately affected minority voters (who tend to be Democrats). Indeed, this is the conclusion of the U.S. Civil Rights Commission's report on voting irregularities in Florida.[7] If true, this conclusion implies that minority racial groups may not be represented in the federal government to the extent that they should be. Our results, which will be seen shortly, generally support this conclusion.

For Broward and Miami-Dade Counties we analyze two related datasets. The first consists of electronic images of all ballots cast in 2000 in both counties. These images tell us how a single Broward or Miami-Dade resident cast votes in the many races (presidential, legislative, county-level, and so forth) that took place in 2000. We use the electronic ballot images to explore whether ballots with presidential overvotes also contain overvotes on other contested races. We also explore whether there appear to be partisan patterns in ballots with presidential overvotes, i.e., whether ballots with presidential overvotes appear to contain mainly Democratic or Republican votes on non-presidential races.[8]

---

[5]Centralized tabulation refers to the fact that Broward and Miami-Dade ballots were counted in a central location rather than in each precinct.

[6]Each county in Florida, like each county in the United States, is free to choose its own voting system for general elections. According to Florida's Department of State, there were seven different such systems in use in 2000. For instance, some counties in Florida used an optical scanning system for reading ballots and other counties used punch card readers. A single county, Union, relied on manually counted paper ballots. According to Florida state legislation passed in 2001, punch card voting is to be eliminated in the state.

[7]The June, 2001, draft of this report is available at http://www.usccr.gov/vote2000/stdraft1/main.htm.

[8]The electronic ballot images we use do not explicitly tell us whether a ballot with a presidential overvote was unambiguous in its choice of presidential candidate. We cannot determine the number of ballots, say, produced by voters who circled Al Gore's name and crossed out—accidentally voting for—various other presidential candidates. At the time of this paper's writing, a consortium of newspapers led by the National Opinion Research Center is studying the mass of Florida overvotes. If this consortium publishes its data then

Our second dataset for both Broward and Miami-Dade Counties combines precinct-level presidential voting rates with voter registration demographics. We use this dataset to determine whether the overvote rate in a Broward or Miami-Dade precinct was related to the extent that the precinct was politically liberal, populated by minority racial groups, and so forth.

The paper's combination of two distinct levels of analysis—at the ballot level and at the precinct level—strengthens its results considerably. The advantage of ballot data, and this is discussed further below, is that these data allow us to count the number of Broward and Miami-Dade ballots that had, for example, presidential overvotes and U.S. Senate race overvotes. Without ballot data, counting in this fashion would simply not be possible. Rather, this type of counting would require ecological inference, a statistical technique that needs to be treated very cautiously (Achen and Shively 1995).

Nonetheless, our precinct-level results are based on ecological inferences. We use data at the precinct level because, on account of ballot secrecy, relevant demographic data is not available at the ballot level. However—and this is a key point—our aggregate results very nicely complement our ballot-level conclusions. Thus, even though any reliance on aggregate data is troublesome, the complementarity between ballot-level and precinct-level analyses greatly increases the confidence we have in our results.

The use of data from only Broward and Miami-Dade Counties limits the scope of our study of presidential overvoting. Both counties are pro-Democratic and, although populous, are certainly not representative of the entire state of Florida let alone the country as a whole. Unfortunately, the types of analyses we conduct cannot be extended to all counties in Florida. This is because, to our knowledge, archives which contain electronic ballot images are not available for the vast majority of Florida's counties, to say nothing of the many counties across the United States. Therefore, given the difficulty of obtaining and analyzing ballot-level overvoting data and given how little we know about overvoting rates and their covariates, we believe it is fruitful to examine closely Broward and Miami-Dade

_____

it might be possible to assess both the number and balance of overvotes that had unambiguous candidate choices.

4

Counties.

The remainder of this paper is organized as follows. In Section 2 we briefly discuss the relationship between overvoting and representation. Then, in Section 3 we pose three hypotheses as to how presidential overvoting occurs. In light of these hypotheses, Section 4 presents results from a ballot-level analysis of Broward and Miami-Dade Counties. In Section 5 we use these results in a precinct-level analysis, which includes registration demographics, of presidential overvoting in the two counties. Finally, Section 6 concludes.

2   Overvoting and Representation

As noted in the introduction, very little is known about what the presence of presidential overvotes means or the types of people who cast these spoiled votes. Thus, little is known about whether overvoting has implications for citizen representation in the federal government. Some presidential overvotes, presumably, are intentional and reflect voter alienation from politics or intense dislike of all presidential candidates. And, it seems almost certain that some overvotes are accidental. A voter, for instance, might incorrectly believe that she is supposed to vote for both a presidential candidate and a vice presidential candidate; this could lead to a double-marked presidential overvote. Or, a voter who confused "Lieberman" (the name of Al Gore's vice presidential running mate) with "Libertarian" (the political party of 2000 presidential candidate Harry Browne) might vote for both Gore and Browne.[9] Or, a voter, erroneously treating "write-in" as an imperative, might vote for a single candidate twice, once by correctly indicating a vote for this candidate in the appropriate place on a ballot and again by voting for this person in a write-in section.[10]

Suppose that all presidential overvotes cast in a given general election were accidental and that the people who cast them either misunderstood an aspect of the voting process or simply suffered an errant slip of a stylus. It is possible in this scenario that presidential overvotes are uniformly random occurrences in that all voters are equally likely to overvote

---

[9]See "Lake erred by tossing write-ins," *Orlando Sentinel*, January 28, 2001.

[10]See Ballot Image 34 from the Leon County Overvoted Presidential Race Ballot Archive, published by the Supervisor of Elections, Leon County, Florida.

accidentally. Or, it may be the case that certain types of people—supporters of Democratic candidates, say—are more prone to making errors that produce presidential overvotes. If all voters are equally likely to overvote accidentally, then overvoting may not be a cause for serious alarm insofar as representation is concerned. This is because uniformly random overvoting will not systematically bias election outcomes in a partisan direction.

If all voters are equally likely to overvote, then overvoting can only cause systematic biases in election outcomes when the overall election consists of a sequence of subelections, at least one of which has a vote share floor below which candidates are eliminated from future rounds. For instance, suppose that vote shares in the first round of a four-way race between candidates A, B, C, and D would be 20%, 21%, 29%, and 30% if all voters cast valid ballots. If an election rule mandated that candidates with less than 20% of the popular vote were eliminated from the second round and if random overvote rates were sufficiently high, then both candidates A and B could be so eliminated. This would leave an election of C versus D, whereas the absence of a vote share floor would produce a second round of B versus C versus D. One can easily imagine a scenario—the 1992 presidential election comes to mind—in which the presence of a third-party presidential candidate (B, in this example) alters overall election dynamics and has implications for who wins the election.

This simple example is useful insofar as it shows why uniformly random overvoting could in theory be an important political phenomenon with systematic influence on citizen representation. However, presidential elections in the United States do not have multiple rounds with associated vote share floors. Therefore, if overvotes are uniformly random occurrences, the presence of them will not systematically bias presidential election outcomes. Hence, overvoting can be a relatively innocuous phenomena which does not distort representation.[11]

This sanguine result on the relative non-importance of overvoting will not hold if certain types of voters are more likely to overvote than others. Suppose, for instance, that support-

---

[11]It is, however, true that presidential races are preceded by a sequence of party primaries. Although primaries do not have explicit vote floors as discussed in the example, candidates who have poor showings in these elections tend to drop out of the overall presidential race. In theory, then, sufficiently high overvote rates could suggest to a candidate that she should drop out of the presidential race when, in fact, her level of popular support does not require this.

ers of Democratic candidates are more likely to overvote accidentally in a presidential race than are comparable Republicans (or, in a similar fashion, suppose that Democratically-inclined citizens are more likely to feel politically alienated than Republicans). If this supposition holds, then vote shares of Democratic presidential candidates will suffer. Moreover, high rates of overvoting could transform a presidential election in which the Democratic candidate is a very close winner, assuming that all voters correctly mark their ballots, could into one in which the Democrat loses.

Thus, the key issue which determines whether presidential overvoting has implications for representation is the extent to which there exists a partisan bias in overvoting rates. If voter preferences are linked to overvoting propensities, then certain types of voters will have disproportionately higher overvoting rates and hence receive less government representation. This observation is the motivation for the empirical analysis of overvoting that follows.

3 Three Hypotheses as to the Identity of the Typical Overvoter

In this section we offer three hypotheses which speak to the characteristics and motivations of the typical presidential overvoter. Although the literature on voting behavior in American presidential elections lacks a theory which touches on overvoting *per se*, the hypotheses are intuitively compelling, they are based on the previous section's discussion of representation, and the paper's empirical analysis is organized around them.

Hypothesis 1 (Uniformly Random Overvoting) *Presidential overvoting is uniformly random: the probability of an individual's casting a presidential overvote depends neither on whether she overvotes or undervotes on races other than the presidential contest nor on the individual's political preferences.*

This hypothesis is consistent with the idea that all voters are equally likes to cast presidential overvotes. It is also consistent with the notion that the likelihood of a given voter being politically disaffected and deliberately spoiling her ballot—"exiting," as Hirschman (1970) might say—is independent of her political preferences.

7

Hypothesis 1 implies that Democrats do not cast presidential overvotes at a higher frequency than Republicans and that overvoters in a congressional race are no more likely to be presidential overvoters than are individuals who vote validly in the former. It follows from Hypothesis 1 that presidential overvoting induces no systematic partisan bias on election outcomes—even though in any given election uniformly random overvoting could, with some presumably low probability, alter the election outcome.

Hypothesis 1 should be considered a relatively sanguine one insofar as it implies that voter error rates or alienation levels are not correlated with underlying voter preferences or partisan affiliations. Nonetheless, if Hypothesis 1 holds then high overvoting rates might signal serious voter education deficiencies or high rates of alienation from the American political system.

Hypothesis 2 (Correlated Overvoting)   *Presidential overvoting is correlated across races: the probability that an individual casts a presidential overvote is related to whether she overvoted on races other than the presidential contest.*

If Hypothesis 2 holds, then, roughly speaking, the residents of a given county or precinct can be divided into two groups, those who essentially never overvote on presidential or other races and those who more frequently overvote. Under this hypothesis, within a single general election the same group of voters is largely responsible for presidential overvoting, congressional race overvoting, and so forth.

Hypothesis 2 is consistent with the claim that certain types of voters—perhaps relatively disaffected individuals, uneducated individuals, or first time voters—are more likely than others to cast presidential overvotes as well as overvotes on other races. Furthermore, there are two noteworthy features of the hypothesis. First, if Hypothesis 2 is true then Hypothesis 1 is not: if certain types of individuals have an elevated risk of presidential overvoting, then it follows that overvoting cannot be uniformly random. Second, although Hypothesis 2 speaks to the notion of correlation in overvoting across different races within a single general election, it says nothing about the types of voters who might be prone to this behavior. In other words, Hypothesis 1 (uniformly random overvoting) can be false

8

while Hypothesis 2 (correlated overvoting) is also false.

Finally, our third hypothesis addresses the possibility that there is partisan bias in presidential overvoting.

Hypothesis 3 (Partisan Overvoting) *Presidential overvoting has a partisan bias: the probability that an individual casts a presidential overvote is a function of her political preferences.*

There are many reasons to think that Hypothesis 3 holds. We would expect presidential overvoting to be more prevalent among individuals who are less familiar with the nuances of ballot structure and voting rules in general. Two types of individuals are likely to suffer from this lack of familiarity, those who are of low socioeconomic status and those who typically abstain from voting in national elections.

Low socioeconomic status is correlated with support for the Democratic Party (e.g. Miller and Shanks 1996). And, abstention rates in American presidential elections are higher for those who support Democratic candidates (e.g., Radcliff 1994). Hence, it is logical to hypothesize that presidential overvoting is partisan in nature and, specifically, harms Democratic presidential candidates.

Indeed, Knack and Kropf (2001) show that, in the 1996 presidential election, United States counties with relatively large black and Hispanic components also had relatively high rates of presidential undervoting and overvoting if the counties had voting systems that allowed overvoting. A similar pattern for the 2000 general election was found in Cook County, Illinois and in Fulton County, Georgia.[12] Moreover, Darcy and Schneider (1989) identify a precinct-level, racial rolloff effect in the 1986 general election in Oklahoma, Nichols and Strizek (1995) show that the percentage black within wards in Franklin County, Ohio was correlated with ballot roll-off rates in the 1992 general election, and a report authored by the U.S. House committee on Government Reform (2001) argues that low

---

[12]See "A Racial Gap in Voided Votes," *The Washington Post*, December 27, 2001 and "State Worst in Ballot Errors," *The Chicago Tribune*, April 29, 2001. Moreover, see "Many Votes Uncounted in Ohio's Poor Areas," *Columbus Dispatch*, December 17, 2000, which shows that high invalid presidential vote areas in Ohio tended to be places with relatively high poverty rates. Also see Bullock III and Dunn (1996) on the impact of district racial composition for ballot roll-off rates in Atlanta, Georgia pursuant to a 1993 municipal election.

income, high minority congressional districts across the United States had high rates of invalid presidential votes in 2000.[13] Since blacks and Hispanics tend to be Democrats, these results are tantamount to concluding that counties with Democratic proclivities tend to have higher invalid voting rates.

It is not germane to Hypothesis 3 whether this ostensible racial effect in invalid voting is due to socioeconomic status differences among Blacks, Whites, and Hispanics, racial gaps in typical education levels, or some other factor.[14] Rather, the key to the hypothesis is that some combination of these or other factors induces a partisan bias in overvoting rates such that Democrats overvote more frequently than Republicans.

If Hypothesis 3 is true, then the relatively sanguine Hypothesis 1 is not: the presence of partisan patterns in presidential overvoting implies that it cannot be the case that all voters are equally likely to cast presidential overvotes. It is possible, however, that overvoting is correlated across races on a ballot but in a way that is independent of political preferences. In other words, Hypothesis 2 can be true even when Hypothesis 3 is false.[15]

Hypothesis 3 has serious implications for the issue of voter disenfranchisement. If presidential overvoting is indeed partisan, and if Democrats overvote more frequently than Republicans, then it follows that the interests of the former will have less representation in the federal government than those of the latter.

We address this section's three hypotheses in numerical order because, to a large extent, their implications become progressively worse. If Hypothesis 1 holds, then presidential overvoting does not cause systematic, partisan biases in election outcomes. If Hypothesis 2 is true, then certain types of voters are responsible for overvotes in general; but, reassuringly, this still might not have systematic election-altering implications. In contrast, if Hypothesis 3 obtains, then presidential overvoting is a serious problem that by itself can alter election

---

[13]See Tomz and Van Houweling (2001) for another study of the 2000 election which concludes that there is a racial gap in invalid voting rates.

[14]With respect to socioeconomic status, Walker (1966) shows that county-wide rolloff rates vary by average education level in county.

[15]If for a given general election there exists a group of voters which is evenly split between Democrats and Republicans, and if this group and no other also tends to have relatively high presidential overvoting rates and on other races as well, then there will be correlated presidential overvoting as in Hypothesis 2 with no partisan bias.

outcomes in a systematic way and can influence the extent to which certain groups of voters are represented in the government.

4   Ballot Images from Broward and Miami-Dade Counties

There were ten presidential candidates on the various 2000 election ballots used across the 67 counties of Florida. Therefore, many different presidential overvoting configurations were possible for residents of Broward and Miami-Dade, the two counties at the heart of this study. Voters could overvote by voting for two presidential candidates, three candidates, four candidates, and so on all the way up to ten candidates. Indeed, Table 1 shows that overvotes of all possible number, two votes through ten votes, occurred in Broward and Miami-Dade Counties.[16]

*** Table 1 about here ***

The largest category in Table 1 is valid votes. In Broward County, approximately 97.5% of the 588,007 ballots cast in the 2000 general election contained a valid presidential vote. And, in Miami-Dade County approximately 95% of the 651,650 ballots cast there contained a valid vote. In both counties, the next most common presidential vote was an undervote and the third most common presidential pattern was a double-punched overvote. Indeed, Table 1 shows that the vast majority of ballots in Broward and Miami-Dade Counties contained either a presidential undervote, a valid presidential vote, or a double-punched presidential overvote.

Table 1 shows that 78 ballots in Broward County and 334 ballots in Miami-Dade County contained votes for all ten presidential candidates. These two numbers are very small in light of the fact that both counties had over half a million voters in 2000. Still, it is practically inconceivable to think that ballots which include votes for all presidential candidates are accidental. Instead, these ballots presumably reflect deliberate abstention in the presidential race.

Nonetheless, without very strong assumptions it is not possible to distinguish accidental

---

[16]Ballot images in encoded form were acquired from the Supervisors of Elections in Broward and Miami-Dade Counties. Image files were then decoded with software provided by Paul Nolte.

from intentional overvoting. To the point, it seems hard to imagine that all the overvotes listed in Table 1 are truly accidental, but we cannot say this for certain. One could argue that intentional overvoting has no negative representational consequences because deliberate overvoters get the presidential representation they want, i.e., none. Here, however, we take the view that all overvotes are problematic at some level, either because they are accidental or because they reflect alienation that is undesirable for the polity's sake.

Based on a Pearson χ^2 statistic of approximately 5268, which can be compared to a critical value of 29.6 for a test of level 0.001, we can easily reject the null hypothesis that Broward and Miami-Dade voters behaved in a similar way insofar as casting presidential ballots. In particular, simple algebra shows that Broward residents cast valid presidential votes in the 2000 general election at a higher rate than did Miami-Dade residents.[17] Therefore, in the ballot-level analysis that follows we treat Broward and Miami-Dade Counties as distinct and do not pool their voters.

The data used to generate Table 1—and other tables that follow—come from electronic files produced by punch card readers employed in Broward and Miami-Dade Counties. When a 2000 general election ballot was fed through such a reader, a sequence of zeroes and ones was stored in an encoded file. Each zero corresponds to a punch hole that was read as not being punched and each one corresponds to a punched hole. The electronic ballot files are used by election administration offices in case they suffer a catastrophic system failure and need to restore their election results. Moreover, the electronic ballot files are encoded by the punch card readers that produce them so that it is difficult for someone in a Supervisor of Elections office to modify official voting records.

Electronic ballot files are extremely useful because they allow direct calculations that would be impossible using only aggregate data.[18] For instance, for the vast majority of counties across the United States we can easily determine the fraction of county residents

---

[17]In the 1996 presidential election, the Broward undervote rate was 2.14% and the overvote rate 0.71%. Respective Miami-Dade rates were 1.85% and 1.17%. Hence, Miami-Dade's presidential voting error rates are not always worse, in the sense of both overvoting and undervoting, than comparable Broward County rates.

[18]Experiments designed to study the impact of balloting system and voting formats (e.g., Shocket, Heighberger and Brown 1992, Sinclair, Mark, Moore, Lavis and Soldat 2000) have access to ballot-level data. Of course, in these experiments the ballot-level data are simulated.

who cast a valid presidential ballot. Similarly, it is straightforward to calculate the fraction of voters in a county who cast a valid Senate vote (assuming that the county had such a legislative race). Knowing these two numbers would not, however, tell us the fraction of individuals who cast both valid presidential and Senate votes. With electronic ballot files or what can be thought of as ballot images, we can calculate this fraction directly without resorting to an ecological inference technique.

There are several very minor drawbacks to the use of electronic ballot files. First, these files do not report votes for write-in candidates since a voter who casts a valid write-in presidential vote makes no punches at all. However, according to the official Florida certified results, there were *no* valid presidential write-in votes in Broward or Miami-Dade Counties—rather, there were only 40 such votes in the entire state. Nonetheless, a handful of ballots may be misclassified here because of the write-in issue. A ballot that combined a valid punch for a candidate along with a write-in for some candidate would be counted in our analysis, erroneously, as containing a valid presidential vote even though it actually is an overvote.[19] Second, the electronic files tell us nothing about the position of chads that affected some Broward and Miami-Dade ballots. Therefore, double-punched overvotes which have two clean punches appear the same in our records as double-punched ballots that have one clean punch and one hanging chad that was counted as a punch. Similarly, ballots with one dimpled chad (that remained dimpled and did not fall off during a recount) are counted as undervotes. Nonetheless, short of examining by hand Broward County's and Miami-Dade County's complete set of ballots, the electronic records used here present a remarkably comprehensive picture of voting in these two locales.[20]

---

[19]Without physical access to Broward and Miami-Dade's punchcards (there are over a million of these) it is impossible to classify correctly write-in ballots. The write-in issue is almost certainly of no substantive importance to this study since at most only a handful of ballots may be misclassified.

[20]It is not possible to construct a version of Table 1 for the 1996 presidential race. Neither Broward nor Miami-Dade Counties have archives of electronic ballot images comparable to those we use in our analysis of the 2000 election.

4.1 Presidential Overvotes and Overvotes Down the Ticket

The first two overvoting hypotheses involve a potential correlation between presidential overvotes and overvotes further down the ticket. In particular, Hypothesis 1 (uniformly random overvoting) posits that this correlation is zero whereas Hypothesis 2 (correlated overvoting) says otherwise.

Beyond the presidential race there were three Florida-wide races of widespread importance that were contested in the 2000 general election. First, there was a contest for a U.S. Senate seat which featured Bill McCollum (R), Bill Nelson (D), and several minor candidates; second, there was an election for Florida State Treasurer and Insurance Commissioner which included only two candidates, Tom Gallagher (R) and John Cosgrove (D); and third, the 2000 general election had a contest for Florida Commissioner of Education with three candidates, Charlie Crist (R), George Sheldon (D), and Vassilia Gazetas (no party affiliation). Henceforth, when we speak of voting "down the ticket," we mean voting on the three elections described in this paragraph.[21]

With respect to the three races down the ticket, a voter could overvote zero, one, two, or three times. Thus, in light of Hypotheses 1 and 2, Table 2 places all voters from Broward and Miami-Dade Counties into one of three groups—presidential undervote, valid presidential vote, and presidential overvote—and also classifies voters based on overvoting patterns down the ticket. The table shows, for instance, that in Broward County approximately 1.12% of the voters who cast zero overvotes down the ticket cast a presidential undervote.

*** Table 2 about here ***

There are several important features of Table 2. First, the top Broward and Miami-Dade rows in the table—the rows associated with voters who cast zero overvotes down the ticket—presumably include individuals who know how to vote, broadly speaking. Some of these people nonetheless cast presidential overvotes. Thus, it is conceivable that a significant fraction of these overvotes was intentional.

---

[21]Beyond these three races, the 2000 election ballots for both Broward and Miami-Dade County contained a number of county-wide races as well as judiciary retention votes and Florida constitutional amendments. We do not consider this group of races because some do not span both Broward and Miami-Dade Counties and others do not have explicit partisan linkages.

Second, Pearson χ^2 statistics for the separate Broward and Miami-Dade sections are extremely significant. The Broward χ^2 statistic, which compares the top four rows in Table 2, is approximately 42727 with a corresponding 0.001 critical value of approximately 22.5. The significant Pearson statistics imply that, in both Broward and Miami-Dade Counties, individuals who never cast overvotes down the ticket had significantly different presidential voting patterns than did those voters who cast one such overvote, and so forth.[22]

Third, and key to the substance of Table 2, in both Broward and Miami-Dade Counties there is a clear trend as one moves from individuals who did not cast overvotes down the ticket to those who did. The table shows that overvoting across different races in the 2000 general election is correlated as posited in Hypothesis 2. For instance, the probability of having cast a presidential overvote conditional on no overvotes down the ticket is slightly over 0.01 in Broward and around 0.02 in Miami-Dade. In contrast, these probabilities increase to 0.32 and 0.26, respectively, for those individuals who cast one overvote down the ticket. The implication of this upward shift in probabilities is that voters who made voting errors down the ticket were also prone to this same mistake on the presidential race. Furthermore, the large probability jumps (0.01 to 0.32 in Broward and 0.02 to 0.26 in Miami-Dade) suggest that there is a significant discontinuity between individuals who cast some overvotes and those who cast none. This evidence strongly supports Hypothesis 2 (correlated overvoting). It also suggests that the causes of presidential overvoting may be similar to those of overvoting down the ticket.

4.2 Presidential Undervotes and Undervotes Down the Ticket

For purposes of comparison, Table 3 presents a similar collection of results on presidential undervoting. In particular, the table shows that the probability of a presidential undervote

---

[22]The Pearson test statistics for Table 2 are based on raw counts, as opposed to percentages, and some of the corresponding counts are very small. For instance, only six voters in Broward County overvoted on all three races down the ticket and also cast a presidential undervote. Therefore, we also calculated Pearson test statistics for a two-row version of Table 2 in which all three positive overvote rows are pooled for each county. The resulting Broward and Miami-Dade test statistics are highly significant. This means that individuals in Broward and Miami-Dade Counties who cast zero overvotes down the ticket had significantly different presidential voting patterns than those who cast at least one such overvote.

is greatest conditional on three undervotes down the ticket. For instance, in Broward County a voter with three undervotes down the ticket also cast a presidential undervote with probability 0.153. On the other hand, someone with no undervotes down the ticket cast a presidential undervote with probability 0.556 (0.579 in Miami-Dade). One might infer from this that the base rate of presidential abstention conditional on a voter's turning out is approximately one-half of one percent.

<div align="center">*** Table 3 about here ***</div>

Tables 2 (overvoting) and 3 (undervoting) show that valid presidential voting rates for individuals who cast undervotes down the ticket are in general greater than comparable valid presidential voting rates for overvoters. For instance, and in the most egregious undervoting case, the probability of a valid presidential vote conditional on three undervotes down the ticket is approximately 0.52 in Broward County and approximately 0.65 in Miami-Dade. These two probabilities are much greater than the valid vote probabilities (approximately 0.15 and 0.21 in Broward and Miami-Dade, respectively) for chronic overvoters.[23]

In addition, Table 3 shows that a voter's probability of casting a presidential overvote markedly increases with the number of undervotes the voter has down the ticket. Broward County voters who cast only undervotes down the ticket have a 0.33 probability of casting a presidential overvote. The probability in Miami-Dade was 0.21. This suggests that those individuals who came to the polls in 2000 solely to cast a presidential vote were in general extremely unpracticed or inexperienced in voting. To the extent that voter inexperience is not uniformly distributed among all voters—to the point, if Democratically-inclined individuals tend to vote less frequently than Republicans—then this could induce a partisan bias in overvoting as in Hypothesis 3.

---

[23]Pearson χ^2 statistics which compare Broward and Miami-Dade rows of Table 3 are highly significant, and this implies that the presidential voting behavior of individuals who cast valid votes down the ticket was fundamentally different than the behavior associated with voters who cast undervotes down the ticket. The relevant χ^2 values are approximately 28949 and 49804 for Broward and Miami-Dade, respectively. Grouping all undervoters together as in fn. 22 produces highly significant results as well.

4.3 Partisan Bias in Presidential Overvotes

Thus far we have shown that overvotes are correlated across presidential voting and voting down the ticket and we identified a similar pattern regarding presidential undervoting. These results support Hypothesis 2 (correlated overvoting) over Hypothesis 1 (uniformly random overvoting) and suggest that a group of similar voters may have been responsible for a large fraction of the presidential and non-presidential overvotes cast in Broward and Miami-Dade Counties in the 2000 general election. However, we have not yet considered whether overvoting in 2000 had a partisan bias as posited in Hypothesis 3. We now address this possibility from two perspectives. First, we contrast individuals who voted straight Democratic down the ticket with those who voted straight Republican. Second, we analyze the frequency with which individual candidate names appeared on overvoted presidential ballots.

4.3.1 Comparison of Straight Voting Partisans Down the Ticket

Table 4 describes how individuals who cast straight Democratic and Republican votes down the ticket voted in the presidential race. Consider first the various percentages listed in the table. What is evident in both the Broward and Miami-Dade percentages is that individuals who voted straight Republican down the ticket were more likely to cast a valid presidential vote than were those people who cast straight Democratic votes down the ticket (because, for instance, 99.2>98.3 in Broward and 98.8>97.0 in Miami-Dade). Moreover, Pearson χ^2 statistics which compare the Democratic and Republican rows of Table 4 show that, within both Broward and Miami-Dade Counties, individuals who voted straight Democrat down the ticket behaved significantly differently on the presidential portion of their ballots than those who voted straight Republican.

*** Table 4 about here ***

In particular, individuals who voted Democratic down the ticket have overvoting rates—1.21% in Broward and 2.39% in Miami-Dade—which are approximately three times as large as the overvoting rates of individuals who voted Republican down the ticket—0.365% in

Broward and 0.863% in Miami-Dade. Thus, individuals who look like Democrats were around *three times more likely* to overvote in the presidential race as were ostensible Republicans. This is a striking finding and it constitutes compelling evidence in support of the partisan overvoting hypothesis—Hypothesis 3.[24]

Moreover, the absolute difference in what appear to be Democratic and Republican presidential overvoting rates is substantial. In Broward County, the Republican valid presidential voting rate advantage is approximately $1.21-0.365=0.845\%$ while in Miami-Dade the comparable advantage is approximately $2.39-0.863=1.53\%$. A rough estimate of the implication of these two differences is as follows.

Suppose that all Broward and Miami-Dade residents who voted straight Democrat (Republican) down the ticket in the 2000 election actually intended to vote for the Democratic (Republican) presidential candidate as well. If this is true, then in Broward County Al Gore lost $257619 \times 0.0121 \approx 3122$ votes and George Bush lost $132047 \times 0.00365 \approx 482$ votes. The consequent net loss to Gore in Broward due to overvoting is thus $3122 - 482 = 2640$ votes, which is about four times as many votes by which he lost the state of Florida and hence the 2000 presidential election. In Miami-Dade County, calculations based on similar assumptions about straight ticket voting imply that Gore lost approximately 3207 net votes. Therefore, combining these totals implies that Gore lost approximately 5847 net votes because of presidential overvoting in Broward and Miami-Dade Counties alone.

However, Broward and Miami-Dade Counties are pro-Democratic counties, and this certainly accentuates the corresponding net vote loss suffered by Gore due solely to overvoting. Furthermore, some individuals who voted straight party down the ticket presumably would have voted against this party at the presidential level. Nonetheless, the numbers above do not even take presidential undervoting into consideration, and Table 4 shows that doing so would increase the estimated net loss suffered by Gore.[25] Furthermore, because Broward

---

[24]In addition, we conducted a multinomial logistic regression analysis, results of which are available by request, where the three-category dependent variable is presidential undervote, valid vote, or overvote. For covariates we used votes down the ticket, votes on three Florida constitutional amendments that appeared on the 2000 Broward and Miami-Dade ballots, and undervote/overvote indicators for all of these contests. Results of the multinomial logistic analysis are qualitatively identical to those reported here: individuals who cast valid Democratic votes down the ticket were more likely to cast presidential overvotes than voters who cast valid Republican votes down the ticket.

[25]Calculations are as follows: The undervoting-related Broward County losses to Gore and Bush are

and Miami-Dade tend to support Democratic presidential candidates over Republicans, there is almost certainly a pro-Democratic overvoting gap among voters who cast exactly two valid Democratic votes down the ticket (recall that Table 4 only analyzes straight voting partisans down the ticket). Thus, there are reasons to think that the net loss figures derived above are conservative.

The twelve correlations in Table 4 tell a similar story. For instance, the upper-left of the twelve numbers indicates that the correlation between a voter's casting a straight Democratic vote down the ticket and undervoting in the presidential rate is -0.0524. In absolute terms the twelve correlations are rather small in magnitude because most 2000 election voters in Broward and Miami-Dade Counties cast valid presidential votes. Furthermore, most straight ticket Democrat voters and Republican voters also cast valid presidential votes; thus, the middle column of correlations contains relatively small numbers.

Nonetheless, there are subtle differences among the correlations that are important. In both counties, the correlation between straight Democrat voting down the ticket and valid presidential voting is lower in absolute terms than that between straight Republican voting down the ticket and valid presidential voting. In other words, the relationship between Democratic voting and valid presidential voting is weaker than that between Republican voting and valid presidential voting.

With respect to overvoting and undervoting correlations, there are in Table 4 four Democratic/Republican pairs of interest. Three of them (the exception, discussed below, is the Democratic/Republican pair of presidential undervoting correlations in Broward County) are logical in light of aforementioned results. For example, in Broward County the correlation (-0.0104) between presidential overvoting and straight Democratic voting down the ticket is greater in absolute terms than the correlation (-0.0459) between presidential overvoting and straight Republican voting down the ticket. In other words, the relationship between Democratic voting and presidential overvoting is stronger than that between Republican voting and presidential overvoting. This correlation-based finding and others like

---

$257619 \times 0.00508 \approx 1309$ and $132047 \times 0.00461 \approx 609$, respectively, which lead to a net loss to Gore of approximately 700 votes. In Miami-Dade County, the Gore and Bush losses are $205725 \times 0.00613 \approx 1261$ and $197278 \times 0.00342 \approx 675$, respectively, which lead to a net loss to Gore of approximately 586 votes.

19

it complement the percentages in the left most columns of Table 4.

There is one exceptional pair of correlations—that involving undervoting in Broward County. The pair is exceptional because, according to it, there is a stronger relationship between Republican voting and presidential undervoting (correlation of -0.343) than between Democratic voting and presidential undervoting (correlation of -0.524). This is odd in light of earlier results on the relationship between ostensibly Democratic voters and presidential undervoting.

We suspect that the Democratic and Republican Broward County undervoting correlations are exceptional because there is a weaker relationship between undervoting and partisanship than there is between overvoting and partisanship. Later in the paper we present a precinct-level analysis of undervoting in Broward and Miami-Dade Counties, and the analysis highlights this point—i.e., that overvoting has more of a partisan bias than does undervoting.

4.3.2 Frequency Analysis of Overvoted Presidential Ballots

If the implications of the prior analysis, which suggests that Democrats cast presidential overvotes more frequently than Republicans, are correct, then we should expect to see votes for the 2000 Democratic presidential candidate Al Gore appearing on a disproportionate number of overvoted ballots. Hence, Table 5 describes the frequency with which candidate names appeared on ballots with presidential overvotes.

*** Table 5 about here ***

With respect to all overvoted presidential ballots from Broward County, Gore's name appears on almost 90%. In contrast, of the 524,834 valid presidential votes cast in Broward, Gore's was chosen on 359,255, or approximately 66%, of them. Notably, beyond Gore no other candidate's name appears on even 50% of the overvoted presidential ballots. Indeed, the candidate whose name appearance frequency is closest to Gore's in Broward County is Harry Browne. This could reflect the confusion, alluded to earlier, that the name of Gore's running mate ("Lieberman") is similar to the political party of Harry Brown ("Libertarian"). After Gore and Browne, Table 5 shows that the rest of the presidential candidates

in Broward County appeared on approximately 17% to 31% of overvoted ballots.

In Miami-Dade County, the same basic pattern is evident—Gore's name is on approximately 84% of the overvoted presidential ballots, and the second more frequent name is appear is Monica Moorehead's (about 58%). In contrast to Broward, though, minor presidential candidates appeared with much greater frequency on overvoted ballots in Miami-Dade County. This reflects the fact that in Miami-Dade a greater proportion of ballots had three, four, or five presidential votes (see Table 1).

The rightmost two columns of Table 5 describe candidate name frequencies for double-punched presidential overvotes only. In some sense these ballots represent the closest to normal of the presidential overvotes cast in the 2000 general election, and it is important to ensure that results based on the leftmost two columns of the table are not solely a function of highly anomalous ballots with a very large number of punches.

Patterns in the double-punched ballots described in the table are very similar to those which apply to all overvoted presidential ballots. Namely, Gore's name appears most frequently (89.5% in Broward and 78.6% in Miami-Dade)—far out of proportion to the partisan bias of the counties. In both counties, Browne (42.9% in Broward and 25.6% in Miami-Dade) and Bush (35.3% in Broward and 45.1% in Miami-Dade) appear next in the list of most frequently punched candidates. Following the top three there is a large drop off in name appearance frequency. In Miami-Dade County, the minor candidates that appeared so frequently on presidential overvotes generally disappear once overvotes of more than two votes are ignored.

5   Precinct-level Analysis of Presidential Overvoting and Undervoting in Broward and Miami-Dade Counties

In this section we build on our ballot-level findings by incorporating demographic variables in a precinct-level analysis of presidential voting patterns in Broward and Miami-Dade Counties. Our objective is determining whether overvoted and undervoted presidential ballots were cast in relatively pro-Democratic precincts. If so, this would corroborate the

apparent pro-Democratic bias of the ballots described in the previous section.

For each election day precinct i in Broward or Miami-Dade County, let V_i^{valid} be the fraction of ballots with valid presidential votes, V_i^{under} the fraction with presidential undervotes, and V_i^{over} the fraction with presidential overvotes. By construction, $V_i^{\text{under}} + V_i^{\text{valid}} + V_i^{\text{over}} = 1$. For reasons described below, we also define $\tilde{V}_i^{\text{under}} = V_i^{\text{under}} + 0.01$, $\tilde{V}_i^{\text{valid}} = V_i^{\text{valid}} + 0.01$, and $\tilde{V}_i^{\text{over}} = V_i^{\text{over}} + 0.01$, so that these latter three values are all strictly positive.[26]

We wish to study the precinct-level characteristics, like racial composition, that are associated with high levels of presidential undervoting and overvoting. It would be preferable, of course, to study voter-level characteristics by ballot, i.e., it would be useful to know if a given ballot had been cast by a registered Democrat, an individual who is a member of a racial minority, and so forth. However, since ballots are secret this type of information can never be known. Thus, let X_i denote a vector of precinct characteristics where the components of this vector reflect the demographic and political composition of precinct i's registered voters. The elements of X_i and their mean values, broken down by county, are listed in Table 6. For instance, the average fraction of registered voters who were Democrats is approximately 0.53 in Broward County and 0.46 in Miami-Dade County.[27] As highlighted in the table, a key distinction between Broward and Miami-Dade Counties is that the latter is much more Hispanic than the former. As will be seen shortly, this has implications for how we compute the impact of racial composition changes for valid presidential voting rates.

*** Table 6 about here ***

To study the relationship between the components of X_i and the presidential voting behavior of precinct i's residents we use the following two-equation regression model:

$$\ln\left(\frac{\tilde{V}_i^{\text{under}}}{\tilde{V}_i^{\text{valid}}}\right) = \beta_0 + X_i'\beta + \epsilon_i^1 \tag{1}$$

$$\ln\left(\frac{\tilde{V}_i^{\text{over}}}{\tilde{V}_i^{\text{valid}}}\right) = \gamma_0 + X_i'\gamma + \epsilon_i^2, \tag{2}$$

[26]We ignore absentee returns in our precinct-level analysis of Broward and Miami-Dade Counties. This is because a single absentee precinct generally includes a number of election day precincts. Hence, analyzing absentee precincts requires aggregating beyond the level of election day precincts.

[27]Registration demographics for Broward and Miami-Dade Counties were provided by the Supervisor of Elections in each county.

where β_0 and γ_0 are intercept terms, β and γ are the key parameter vectors to be estimated, and ϵ_i^1 and ϵ_i^2 are disturbance terms assumed to be mean zero and uncorrelated with the elements of X_i.

Although there are three presidential voting variables per precinct, these variables sum to one (or to 1.03, based on $\tilde{V}_i^{\text{under}}$, $\tilde{V}_i^{\text{valid}}$, and $\tilde{V}_i^{\text{over}}$) and hence can be modeled with only two equations. Moreover, we use a natural logarithmic transformation to smooth out the ratios in equations (1) and (2). Note that the two ratios on the left hand sides of the equations are always defined because the denominators are greater than zero by construction (avoiding division by zero) and the numerators are also positive (avoiding logarithmic errors).

Equations (1) and (2) can be estimated separately by ordinary least squares—results are in Table 7—because each equation contains the same set of explanatory variables. Before discussing substantive results, though, it is important to note that the estimated Miami-Dade interaction terms in the table are jointly significant. This is consistent with the rejection of Broward and Miami-Dade pooling described in the earlier ballot-level analysis. Furthermore, even with Miami-Dade racial interactions, we can reject ($p < 0.001$) that the base rates of presidential overvoting in Broward and Miami-Dade Counties are similar. In other words, demographic variables notwithstanding, there is something about residents of Miami-Dade County—perhaps political inexperience or less educational attainment—that leads them to cast a large number of invalid presidential votes.[28]

With the estimates in Table 7 it is possible to conduct experiments on the consequences for presidential voting of changes in precinct-level demographics. The three upcoming figures result from applications of the following procedure. We choose one variable in X_i to analyze and hold constant all other variables. Then, we change the chosen component of X_i and, based on the estimates in Table 7, this induces new values of the logarithms of the ratios in equations (1) and (2). Because $\tilde{V}_i^{\text{under}} + \tilde{V}_i^{\text{valid}} + \tilde{V}_i^{\text{over}} = 1.03$, we can solve these ratios for new values (and standard errors via the delta method) of V_i^{under}, V_i^{valid}, and V_i^{over} that are based on the modified X_i.

---

[28]We have also estimated equations (1) and (2) separately for each county. Results are qualitatively identical to those in Table 7.

Thus, Figure 1 displays the consequences for presidential voting patterns in Broward and Miami-Dade Counties of letting the percent black in a precinct vary from zero to 50%. Based on the numbers in Table 6, for both plots the "other" race group is fixed at 5%; and, fraction Hispanic is fixed at 6% in Broward County and 40% in Miami-Dade. Hence, as the fraction black per precinct increases, the fraction white commensurately decreases. Figure 1 plots both the estimated undervote and overvote rate as fraction black per precinct varies and also displays 95% confidence bands around estimated presidential voting rates.[29]

Figure 1 shows that, controlling for political affiliations and the other precinct-level demographics in Table 6, Broward and Miami-Dade presidential overvoting in the 2000 general election was greatest in those precincts which contained a large fraction of black registered voters. Although this is also true of presidential undervoting, the undervoting result is not statistically significant (note that the undervoting confidence bands for precincts with few black registrants overlap the corresponding bands for precincts that are heavily black). Simply put, precincts with large black populations had, *ceteris paribus*, relatively low valid presidential voting rates. This is very much consistent with evidence about minority voting described in the context of Hypothesis 3.[30]

*** Figure 2 about here ***

A similar analysis of the impact of Hispanic registration on presidential overvoting is captured in Figure 2. The patterns in this latter figure are very similar to those seen earlier. Namely, as the fraction Hispanic in a precinct increases, so does the rate of presidential overvoting. This result is clearly statistically significant in Miami-Dade County but not so in Broward County. Broward County contains relatively few Hispanics—see Table 6—and therefore projections based on large Hispanics fractions in the county are very noisy. This is the source of the large confidence intervals in the left panel of Figure 2.

---

[29]Confidence bands displayed in Figure 1, as well as in other figures, are based on heteroskedastic-consistent standard errors (White 1982).

[30]The relative insignificance of the undervote results could be attributed to so-called dimpled chads. A ballot with a single dimpled chad among the ten president chads is one that contains a punch for president that dented but did not remove the chad which, to a mechanical punch-card reader, indicates a vote. Ballots with one dimpled chad vote for president are counted here as undervotes as opposed to valid votes, and this could attenuate the relationship between precinct-level undervote rates and demographics.

Hispanic residents of Miami-Dade are heavily Cuban-American (50.4%) compared to Broward County Hispanics (18.7% are Cuban-American).[31] And, Cuban-Americans are significantly more pro-Republican than other Hispanics (Moreno and Warren 1992, Moreno and Warren 1992–1993). According to exit poll data collected by Voter News Service, in 2000 George W. Bush received 78 percent of the Hispanic Cuban-American vote in Florida while he received only 49 percent of Hispanic vote in Florida. Nonetheless, Figure 2 shows that Hispanics appear to overvote excessively regardless of whether they favor Democratic or Republican policies. Thus, what the figure suggests is that the source of the Hispanic overvote effect is socioeconomic status.

Cuban-American Hispanics notwithstanding, members of minority groups tend to be Democratically inclined. And, our earlier analysis of ballot images from Broward and Miami-Dade Counties shows that overvoted presidential ballots looked rather Democratic down the ticket. Thus, the precinct-level results summarized in Figures 1 and 2 are consistent with our ballot-level results and suggest very strongly that what appear to be Democratic ballots with presidential overvotes are indeed Democratic ballots.

Further evidence of this can be seen in Figure 3. This figure depicts the consequences for presidential undervoting and overvoting of changes in the fraction of a precinct that registered Democratic for the 2000 general election. The figure is based on varying a single component of the precinct-level demographic vector X_i while holding all over variables, racial and otherwise, at their means.

*** Figure 3 about here ***

The influence of a precinct's having a large contingent of registered Democrats is clear: the more Democrats in a precinct, the greater the precinct's presidential overvoting rate. Similarly, in Broward and Miami-Dade Counties presidential undervoting is greatest, *ceteris paribus*, in precincts with large fractions of registered Democrats. However, this relationship is not nearly as strong as that between Democrats and presidential overvoting.

Assuming that the vast majority of presidential overvotes are accidental as opposed to intentional, our identification of a Democratic effect in presidential overvoting even when

---

[31]Source for the Cuban Hispanic data is Census 2000 Summary File 1, 100-Percent data, PCT11.

controlling for precinct racial composition suggests that the root of this effect is socioeconomic status. The average Democratic resident of Broward or Miami-Dade County is, based on traditional sources of support for Democratic policies, almost certainly of lower socioeconomic status than the corresponding Republican. If individuals of lower socioeconomic status understand less about voting and have a more difficult time comprehending voting instructions, then we would expect to see greater invalid voting rates in precincts that are heavily populated by Democrats. Indeed, Figures 1 and 3 show that this expectation holds.[32]

It is useful to consider the implications of Figures 1, 2, and 3 for the three hypotheses that motivated this paper. Recall that Hypothesis 1 implies that all voting residents of a county have an equal likelihood of casting a presidential overvote. If this hypothesis were correct, then the overvoting curves (and undervoting curves as well) in the figures should be flat. They clearly are not, and hence we can reject Hypothesis 1 based on our precinct-level results. This corresponds well with the rejection of Hypothesis 1 based solely on ballot data.

The three figures do not speak directly to Hypothesis 2 (correlated overvoting) since the figures only examine overvoting in a single race. However, the figures are very supportive of Hypothesis 3 (partisan overvoting) in that they imply that voters with Democratic inclinations are more likely to overvote than voters with Republican preferences.[33]

6 Discussion

Our analysis of 2000 election voting patterns shows that Broward and Miami-Dade County presidential overvotes were cast disproportionately by Democratically-inclined voters. This conclusion follows from an analysis of electronic ballot images—which shows that ballots with presidential overvotes appear Democratic down the ticket—combined with an

---

[32] An interesting hypothesis related to this finding—but one that cannot be tested with the data used here—is that the invalid vote rate is higher in presidential elections that have relatively high turnout rates. This is a logical hypothesis as high turnout elections presumably bring to the polls many individuals who typically do not vote and are of low socioeconomic status.

[33] Of course, one could argue that Republicans prone to presidential overvoting choose to live in highly Democratic areas and, moreover, commit presidential voting errors which make their ballots appear Democratic down the ticket. This is a rather contrived scenario.

aggregate analysis of Broward and Miami-Dade precincts. This latter analysis shows that, *ceteris paribus*, precincts with relatively large numbers of blacks, Hispanics, and registered Democrats had relatively high presidential overvoting rates. Overall, the weight of our ballot-level and aggregate analysis implies that Al Gore lost votes in 2000 due to overvoting. We found similar, although not as compelling, results for presidential undervoting.

Our identification of a partisan bias in presidential overvoting is important for two reasons. First, presidential overvotes are invalid and, therefore, voters who cast them effectively disenfranchise themselves. This can impact representation insofar as groups of voters with high overvote rates can lose the ability to elect officials who represent their interests. In particular, our statistical results imply that numerous Democrats in Broward and Miami-Dade Counties discarded opportunities to ensure that their preferences for Al Gore were counted.

Second, and more narrowly, a partisan bias in presidential overvoting is important because it can change election results. This is particularly noteworthy in the context of the extremely close 2000 presidential election. Indeed, it is natural to ask, did overvoting cost Al Gore the presidency? If one thinks that the vast majority of overvotes cast in 2000 were accidental, then it almost certainly did. Under some very simple and conservative assumptions, we calculated that overvoting in Broward and Miami-Dade Counties led to a net Gore loss of almost 6,000 votes, and this is more than ten times as many votes by which Gore lost the 2000 race. Notably, the 6,000 figure is based on overvotes from only two (Democratically biased) counties in Florida, albeit the two largest counties in the state. Because of a lack of data availability, unfortunately, we cannot extend our analysis to the entire state of Florida. Nonetheless, given the evidence that Democrats overvote more frequently than Republicans, it seems certain that, had all overvoted ballots cast in 2000 been cast in a valid way, Gore would have won the state of Florida and the presidency.

References

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

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

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

Hirschman, Albert O. 1970. *Exit, Voice, and Loyalty.* Cambridge: Harvard University Press.

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

Longley, Lawrence D. and Neal R. Peirce. 1999. *The Electoral College Primer 2000.* New Haven, CT: Yale University Press.

Miller, Warren E. and J. Merrill Shanks. 1996. *The New American Voter.* Cambridge: Harvard University Press.

Moreno, Dario and Christopher Warren. 1992. The Conservative Enclave: Cubans in Florida. In *Latinos in the 1988 Election*, ed. Rodolfo O' de la Garzas and Louis DeSipio. Boulder, CO: Westview Press.

Moreno, Dario and Christopher Warren. 1992–1993. "The Cuban Community and the 1992 Presidential Election." *Harvard Journal of Hispanic Politics* 6:27–36.

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

Radcliff, Benajmin. 1994. "Turnout and the Democratic Vote." *American Politics Quarterly* 22(3):259–76.

Shocket, Peter A., Neil R. Heighberger and Clyde Brown. 1992. "The Effect of Voting Technology on Voting Behavior in a Simulation Multi–Candidate City Council Election: A Political Experiment of Ballot Transparency." *The Western Political Quarterly* 45(2):521–37.

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

The Caltech/MIT Voting Technology Project. 2001. "Residual Votes Attributable to Technology: An Assessment of the Reliability of Existing Voting Equipment." Working draft available from http://www.vote.caltech.edu/.

Tomz, Michael and Robert Van Houweling. 2001. "Who Spoils Their Ballots? Voting Equipment and Race in the 2000 Presidential Election." Working Paper, Stanford University.

U.S. House Committee on Government Reform. 2001. "Income and Racial Disparities in the Undercount in the 2000 Presidential Election." Report prepared by the Special Investigations Division.

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):448–63.

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

White, Halbert. 1982. "Maximum Likelihood Estimation of Misspecified Models." *Econometrica* 50(1):1–26.

Table 1: Presidential Punches by Ballot

Number of Punches	Broward	Miami–Dade
0	6686	14352
1	573396	619574
2	5087	9647
3	836	2345
4	501	1503
5	522	1458
6	235	625
7	150	443
8	203	471
9	313	898
10	78	334
Total	588007	651650

Table 2: Patterns in Overvoting Across Different Races

	Overvotes down the ticket	Presidential Vote		
		Under	Valid	Over
Broward	0	1.12	97.8	1.08
	1	3.33	64.6	32.1
	2	4.65	28.7	66.7
	3	8.96	14.9	76.1
Miami-Dade	0	2.17	95.6	2.28
	1	3.98	69.0	27.0
	2	5.91	52.3	41.8
	3	4.76	21.4	73.8

Note: figures are rounded percentages

Table 3: Patterns in Undervoting Across Different Races

	Undervotes down the ticket	Presidential Vote		
		Under	Valid	Over
Broward	0	0.556	98.6	0.898
	1	2.32	93.1	4.56
	2	10.6	86.0	3.41
	3	15.3	51.8	32.9
Miami-Dade	0	0.579	97.5	1.88
	1	3.94	89.4	6.63
	2	12.4	82.7	4.88
	3	14.2	65.1	20.7

Note: figures are rounded percentages

Table 4: Partisan Patterns in Presidential Overvoting

		Presidential Vote			Correlation		
	Down the Ticket	Under	Valid	Over	Under	Valid	Over
Broward	Democrat	0.508	98.3	1.21	-0.0524	0.0434	-0.0104
	Republican	0.461	99.2	0.365	-0.0343	0.0573	-0.0459
Miami-Dade	Democrat	0.613	97.0	2.39	-0.0735	0.0604	-0.0139
	Republican	0.342	98.8	0.863	-0.0836	0.113	-0.0753

Note: left three columns are rounded percentages;
right columns are correlations

Table 5: Name Frequency on Ballots with Presidential Overvotes

Candidate	All Overvoted Ballots		Double–punched Ballots	
	Broward	Miami-Dade	Broward	Miami-Dade
George W. Bush	30.5	27.2	35.3	45.1
Al Gore	88.9	84.0	89.5	78.6
Harry Browne	47.4	54.9	42.9	25.6
Ralph Nader	24.8	53.5	9.0	9.9
James Harris	22.5	55.5	4.7	6.5
John Hagelin	21.9	48.3	4.2	4.9
Pat Buchanan	18.2	47.4	3.7	7.5
David McReynolds	17.6	45.8	1.7	3.8
Howard Phillips	18.3	49.7	2.8	5.4
Monica Moorehead	24.9	57.9	6.3	12.8
Total Ballots	7925	17724	5087	9647

Note: figures are rounded percentages; candidates are in official Florida order

34

Table 6: Average Precinct–level Demographics of Registered Voters

Variable	Broward County	Miami-Dade County
Democrat	0.530	0.464
Black	0.154	0.209
Hispanic	0.0603	0.403
Other Race	0.0468	0.0456
Male	0.452	0.442
Age 21–29	0.107	0.135
Age 30–55	0.495	0.463
Age 56–64	0.110	0.125
Age 65 and up	0.264	0.236

Note: all figures are fractions of registered voters

Table 7: Regression Estimates

Variable	Equation 1	Equation 2
Constant	-5.12***	-4.65***
	(0.684)	(0.669)
Miami-Dade Indicator	0.0168	-0.289***
	(0.0611)	(0.0598)
Democrat	0.360***	0.936***
	(0.133)	(0.130)
Black	0.190**	0.854***
	(0.0925)	(0.0905)
Black × Miami-Dade	0.497***	0.700***
	(0.0797)	(0.0780)
Hispanic	0.893***	0.822***
	(0.293)	(0.287)
Hispanic × Miami-Dade	-0.509	0.448
	(0.312)	(0.305)
Other	-0.613	2.54***
	(0.800)	(0.783)
Other Race × Miami-Dade	1.77*	2.05**
	(1.00)	(0.979)
Male	-1.15***	-0.454**
	(0.232)	(0.227)
Age 21–29	1.10	0.967
	(0.808)	(0.791)
Age 30–55	1.33*	-0.127
	(0.724)	(0.708)
Age 56–64	1.73**	0.364
	(0.777)	(0.760)
Age 65 and up	1.85***	0.819
	(0.680)	(0.666)
R^2	0.411	0.729
F	64.9***	250***

Note: estimated standard errors in parentheses;
* denotes $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$

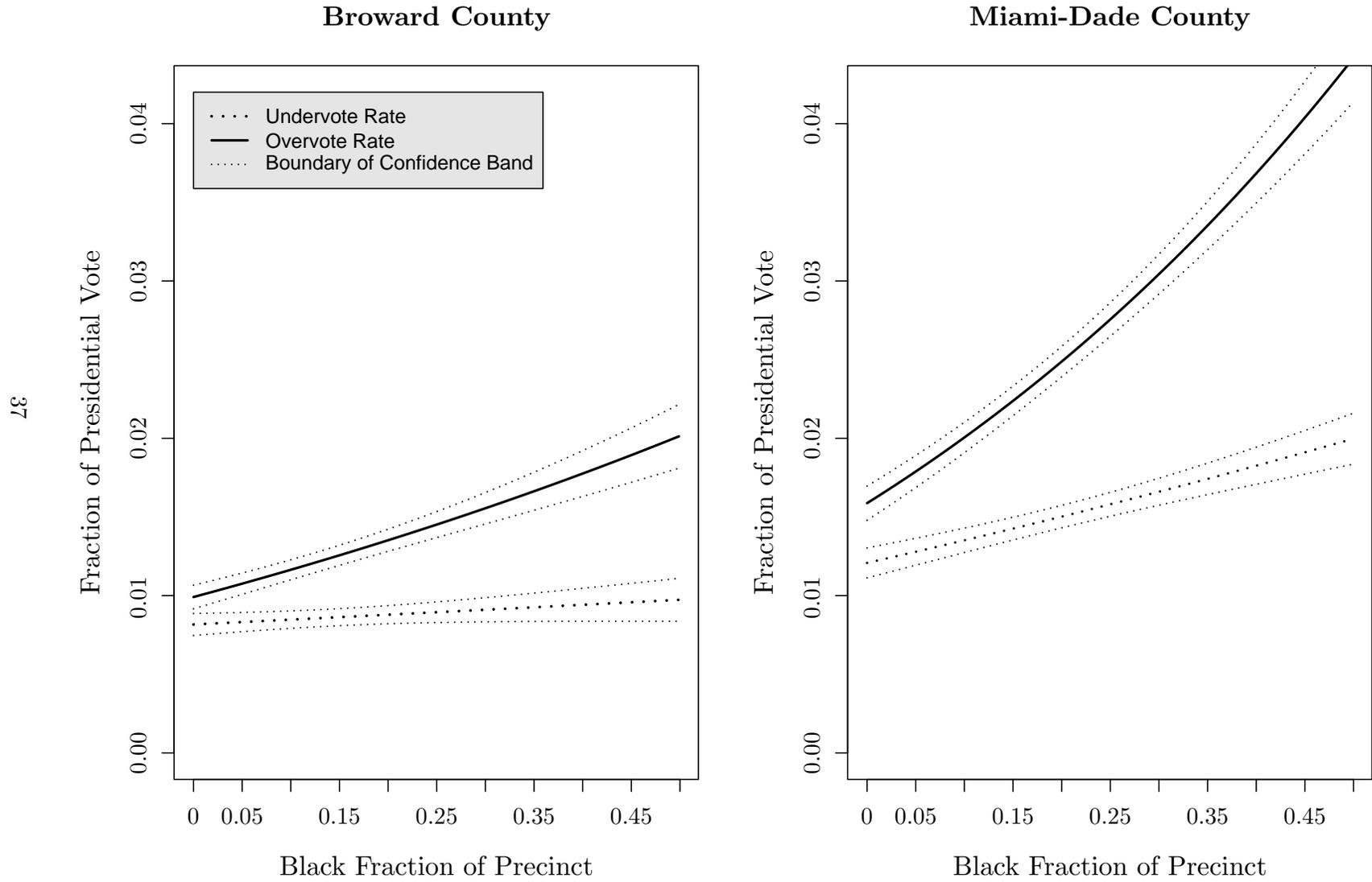Figure 1: Presidential Voting and Fraction Black in Precinct

Broward County

Miami-Dade County

Fraction of Presidential Vote

Black Fraction of Precinct

· · · · Undervote Rate
——— Overvote Rate
· · · · Boundary of Confidence Band

Figure 2: Presidential Voting and Fraction Hispanic in Precinct

Broward County
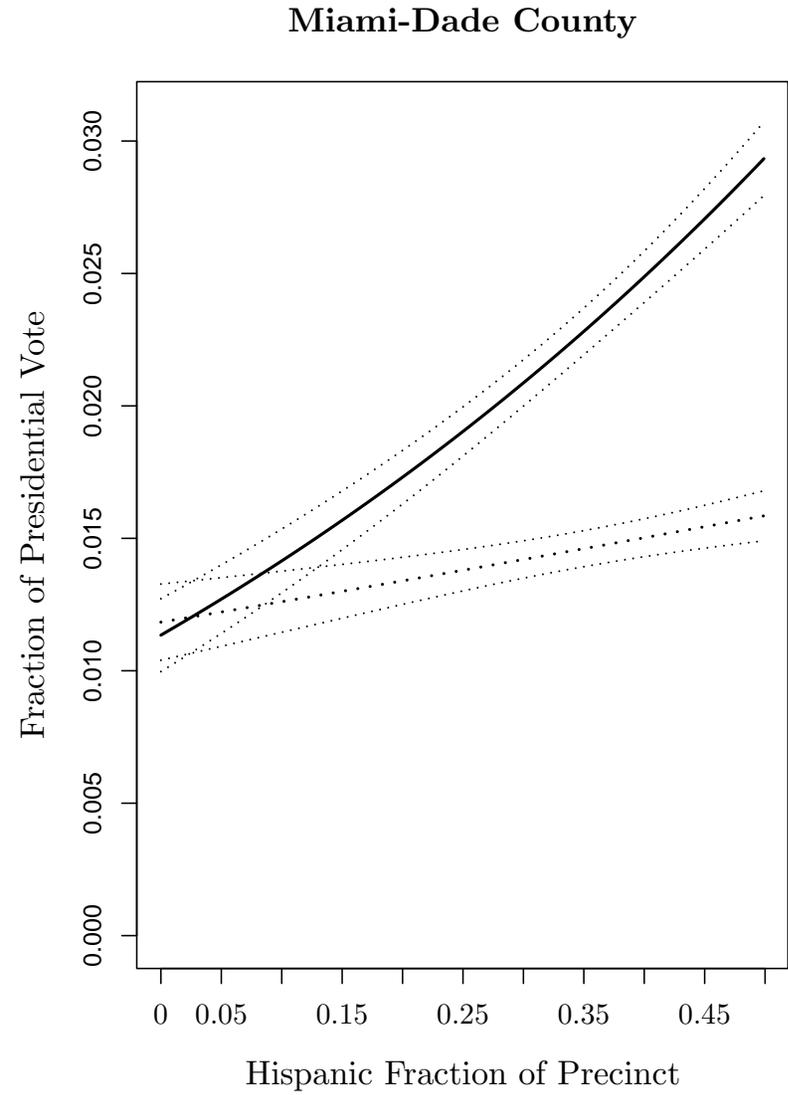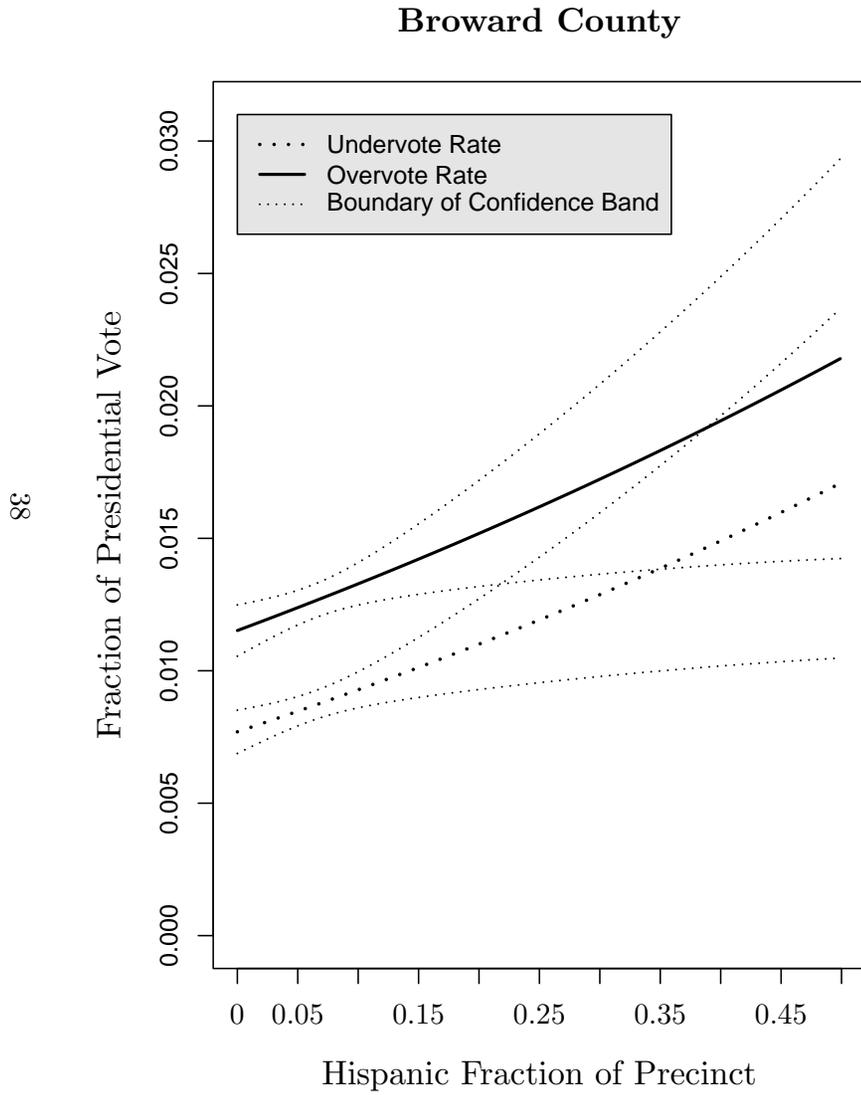
Miami-Dade County

Figure 3: Presidential Voting and Fraction Democrat in Precinct (Broward and Miami-Dade Counties)