Residual Votes in the

2008 Minnesota Senate Race

Jonathan W. Chipman[*]   Michael C. Herron[†]   Jeffrey B. Lewis[‡]

Initial Version: November 13, 2008

This Version: November 15, 2008

---

[*]Director, Applied Spatial Analysis Laboratory, Departments of Geography and Earth Sciences, Dartmouth College (Jonathan.W.Chipman@dartmouth.edu).

[†]Visiting Associate Professor of Government, Harvard University, and Associate Professor of Government, Dartmouth College (Michael.Herron@dartmouth.edu).

[‡]Associate Professor of Political Science, University of California, Los Angeles (jblewis@ucla.edu).

Abstract

The 2008 United States Senate race in Minnesota is one of the closest electoral contests in recent history: as of this writing, out of over 2.9 million ballots cast only 206 votes separate incumbent Republican Senator Norm Coleman and his Democratic challenger, Al Franken. The Minnesota Senate race is slated to be recounted starting on November 19, 2008, and a key issue in the recount will be the approximately 34 thousand residual votes associated with it. A Senate residual vote is, roughly speaking, the product of a ballot that lacks a recorded Senate vote, and in the Minnesota Senate race there is no doubt that the number of residual votes dwarfs the margin that separates Coleman from Franken. We show using a combination of precinct voting returns from the 2006 and 2008 General Elections that patterns in Senate race residual votes are consistent with, one, the presence of a large number of Democratic-leaning voters, in particular African-American voters, who appear to have deliberately skipped voting in the Coleman-Franken Senate contest and, two, the presence of a smaller number of Democratic-leaning voters who almost certainly intended to cast a vote in the Senate race but for some reason did not do so. Ultimately, the anticipated recount may clarify the relative proportions of intentional versus unintentional residual votes. At present, though, the data available suggest that the recount will uncover many of the former and that, of the latter, a majority will likely prove to be supportive of Franken.

The 2008 United States Senate race in Minnesota appears to be one of Ethe closest federal electoral battles in recent history: as of this writing, out of over 2.9 million ballots votes cast in the state only 206 votes separate incumbent Republican Senator Norm Coleman (1211565 votes) and his challenger, Democrat Al Franken (1211359).[1] Minnesota state law stipulates that an election for a federal or state office whose margin is less than one-half of one percent of the votes cast in the election qualifies for an automatic recount,[2] and as such a Senate race recount is slated to begin on November 19, 2008. Literally all ballots cast in Minnesota during the 2008 General Election will be examined as part of this review.[3]

Given the extreme closeness of the Minnesota Senate race, the residual votes cast in it—they number approximately 34 thousand—may have been pivotal to the outcome of the race.[4] Note that a *Senate race residual vote* refers to the lack of a recorded Senate vote on a given ballot. Residual votes are caused by voters who *undervote*, i.e., skip a race in question or try to cast a vote yet fail to do so,

---

[1]Technically speaking, Franken is running under the aegis of the Democratic-Farmer-Labor (DFL) party in Minnesota. The DFL, according to its web site (http://www.dfl.org/index.asp?Type=B_BASIC&SEC={F9806706-2824-4539-B3C0-CD887BEBAACF}) (last accessed November 8, 2008) should be considered the "Minnesota arm of the national Democratic Party." Accordingly, Franken will be referred to here as a Democrat. The vote margin of 206 (which as of November 11, 2008 remains uncertified) is based on reporting in "Franken now just 206 votes behind," posted on the web site of the *Minnesota Star Tribune*. See http://www.startribune.com/politics/national/senate/34200229.html?elr=KArks8c7PaP3E77K_3c::D3aDhUec7PaP3E77K_0c::D3aDhUiD3aPc:_Yyc:aULPQL7PQLanchO7DiU (last accessed on November 11, 2008).

[2]Minnesota recount procedures and guidelines can be found at http://www.sos.state.mn.us/docs/recount_guide_2008.pdf?elr=KArks8c7PaP3E77K_3c::D3aDhUxWoW_oD:EaDUiD3aPc:_Yyc:aULPQL7PQLanchO7DiU (last accessed November 8, 2008).

[3]See http://www.startribune.com/politics/national/senate/34122884.html?elr=KArks8c7PaP3E77K_3c::D3aDhUxWoW_oD:EaDUiD3aPc:_Yyc:aULPQL7PQLanchO7DiU (last accessed November 8, 2008) for the timing of the recount.

[4]Another intriguing aspect is the role of a third party candidate for the Minnesota Senate seat, Dean Barkley, who received 437389 votes, many times the margin that separates Coleman and Franken. Either Barkley siphoned votes from Coleman and turned a solid victory into a narrow one or he spoiled the Senate race for Al Franken.

or *overvote*, i.e., vote for more than the legal number of candidates in the said race (e.g., Brady et al., 2001; Tomz and van Houweling, 2003). Thus, what is being referred to in popular commentary as, say,"Minnesota Senate undervotes" should more accurately be described as "Minnesota Senate residual votes" until it is established that the residual votes are in fact undervotes.[5]

The Coleman-Franken Senate race residual vote rate is approximately 1.2%, which contrasts with a United States presidential residual vote rate in Minnesota of approximately 0.34%.[6] That this latter rate is not zero should not be considered striking. Knack and Kropf (2003*a*) show that the undervote rate in presidential elections often hovers around 2%. Moreover, a Senate race residual rate of approximately 1.2% is not inherently strking either. Indeed, it is nowhere near the exceedingly controversial residual vote rate of over 8% observed in the 2006 United States House race in Florida's 13th Congressional District, a contest that produced post-election litigation and investigations into voting machinery (e.g., Frisina et al., 2008). The only reason, it appears, that 2008 Minnesota Senate residual rate is considered worthy of attention at this point is because the associated Coleman-Franken margin is minute and in particular is two orders of magnitude smaller than the number of Senate race residual votes.

What might explain the pattern of residual votes in the 2008 Minnesota Senate

---

[5]See, for example, the November 8, 2008 *Associated Press* story titled "Most Minn. Senate 'undervotes' are from Obama turf" on the distribution of residual votes in the Minnesota Senate contest. This story be found at `http://ap.google.com/article/ALeqM5gMpTmr96V5hKIfyHT4Av4jsVQgrQD94AE8P80` (last accessed November 10, 2008).

[6]As of November 10, 2:31:10 PM, the Minnesota Secretary of State was estimating the total number of ballots cast in the state as 2920180. There were 2885052 recorded Senate votes and hence 35128 Senate residual votes. Similarly, there were 2910332 recorded presidential votes and hence 9848 presidential residual votes. See `http://electionresults.sos.state.mn.us/20081104/PrecRpt.asp?M=TPR` (last accessed November 10, 2008).

contest? In this paper we consider two explanations. The first explanation is that the residual votes cast in the Senate race reflect the actions of voters who deliberately decided to skip voting in the Coleman-Franken contest. One might call this an *intentional undervote* explanation, and deliberate abstention is certainly one reason that undervotes (and hence residual votes) exist. It is well known that residual vote rates increase as voters move down their ballots, and most rationales for this phenomenon focus on ballot fatigue, voter indifference over seemingly obscure races, and lack of information about candidates running for office. That is to say, there are many reasons that voters may decide not to vote in a particular race, and thus one possibility is that the approximately 34,000 residual votes in the Minnesota Senate race reflect a myriad of voters' choices not to participate.

A second and competing explanation is that the Senate race undervotes reflect the actions of voters who tried to vote in this race but for some reason were unable to register their candidate preferences in a way that Minnesota's vote tabulating technologies were able to discern. This could have happened if, say, voters did not adequately follow the directions given to them regarding how to vote using Minnesota's optical scan voting machines. That voters make errors when filling out ballots is clear (Kimball and Kropf, 2005), and indeed interface issues between voters and the machines used to record votes are not new.[7]

The subject of voter errors and the possibility that a voter marked a ballot in a way that did not register his or her preferences invariably calls into question the

_____

[7]The most detailed examination of human interface issues and voting machines is Herrnson et al. (2007). This study of electronic voting is not directly germane to Minnesota, a state that does not use what are called direct record electronic voting machines. Nonetheless, the laboratory experiments in Herrnson et al. establish that voters are capable of a wide variety of mistakes while voting, and this point is relevant to the second explanation, above, for the residual votes cast in the Minnesota Senate race.

role that vote tabulating technologies have in interpreting voter choices (this is one lesson, among many, of the Florida fiasco that occurred during the 2000 General Election), and with this in mind it is important to note that all 87 Minnesota counties use paper optical scan ballots.[8] Neither electronic voting machines, punchcards, nor lever machines were used in the 2008 Minnesota General Election.

Even though voting technology, ballot types, and ballot formats were standardized in 2008 across Minnesota, there are three distinct ways that ballots are counted in the state: by hand; by machine in a centralized location in a county ("central count"); or by machine locally in a precinct ("precinct count" or what the Minnesota Secretary of State calls "Precinct Tabulator").[9]

This sort of variance is notable because Minnesota employs a limited form of "Second Chance Voting" in its precincts. Namely, the optical scan voting machines in so-called precinct count precincts are programmed to reject or kick back to voters overvoted and entirely blank ballots.[10] This feature of a voting machine will help a voter avoid an accidental overvote (but the rejected ballot can be overridden by a voter who wishes to cast a ballot containing either no votes at all, i.e., a blank ballot, or some votes that are known to be invalid). Notably, Minnesota's optical scan machines are not programmed to kick back undervotes. Regardless, precincts

---

[8]A Minnesota voter can use an electronic device called an AutoMARK to fill out his or her paper ballot if the voter desires accessibility features that the AutoMARK provides. Note, however, that the AutoMARK is not an electronic voting machine in that it does not count votes. Rather, it assists voters in filling out paper ballots that are subsequently processed by an optical scan machine or are hand counted. For instance, the AutoMARK machines used in Minnesota allow voters to use headphones when navigating their ballots, they can magnify ballot text, and so forth. AutoMARK machines do not permit overvotes. When a voter is done completing a ballot with an AutoMARK, the AutoMARK generates a paper ballot that a voter then must bring to a counting station.

[9]The standardization of ballot types across a state is highly unusual in the contemporary United States. In the vast majority of states there is variance in vote tabulating technologies across counties (or across towns, in the case of New England states). See Niemi and Herrnson (2003) for examples of this sort of variance.

[10]Personal communication with the Minnesota Secretary of State's office on November 10, 2008.

that are hand count or central count do not allow for second chance voting.

The inconsistent availability of second chance voting is important because this feature of voting technology can help prevent unintentional overvotes. If, e.g., a voter voted for Norm Coleman and also wrote in Norm Coleman for Senate (perhaps because this hypothetical voter treated "Write-In" as an imperative) then such an overvoted Senate choice would trigger a rejected ballot in a precinct count precinct. No such triggering would occur in a central or hand count precinct, however.

Beyond intentional undervoting (due to deliberate abstention) or unintentional undervoting (due to a failure to indicate candidate choices in machine-readable ways) there are in theory other explanations for the Minnesota Senate race undervote. One such explanation involves tabulating machine malfunctions wherein a set of voters correctly followed the instructions given to them when casting their votes but the vote counting machines in which they placed their ballots failed to register their preferences.[11] Another potential explanation is deliberate election fraud. As of November 11, 2008, neither of these latter two explanations is particularly prominent, and hence the undervote problem as posed here will be conceptualized as one whose resolution requires determining whether the Coleman-Franken Senate race undervotes were the result of deliberate abstention, were unintentionally cast, or both.[12]

In the remainder of this paper we offer a simple data analysis that sheds light

---

[11]One could call this an unintentional overvote, of course, but to do so would confuse an under-vote caused by machine malfunction and an undervote caused by a voter who tried to follow voting directions but for whatever reason did not succeed in doing so.

[12]Also as of November 11, 2008, the Minnesota Secretary of State has not released any details on overvote counts by county or precinct. It remains unclear if there are any records for overvote counts in historical Minnesota elections.

on the two residual vote explanations sketched above. What follows is preliminary; it relies on uncertified vote totals; and the data on which it is based are changing regularly.

1 Spatial Patterns in Senate Race Residual Votes

Minnesota is divided into 4130 voting precincts, and the political disposition of the state as of the 2008 presidential race is characterized in Figure 1.[13] This figure shows the spatial distribution of Minnesota precincts won by Illinois Senator and now president-elect Barack Obama (blue) and Arizona Senator John McCain (red).[14] In general, Obama tended to win precincts in the densely populated core of the Twin Cities Metropolitan Area (located in the east-central part of Minnesota) as well as in the sparsely populated far north and the northeast, where mining and Great Lakes shipping perhaps provided union jobs that have contributed to a pro-Democratic culture. In contrast, McCain tended to dominate in suburban and predominantly rural-agricultural precincts that appear throughout much of the rest of Minnesota.

The Senate residual vote rate across Minnesota precincts is shown in Figure 2. Note that the light gray boundary lines in the figure denote counties and that the numerous small squares that effectively partition Minnesota generally correspond to townships. Most Minnesota townships are approximately 36 square miles in area,

---

[13]Shapefiles for Minnesota precincts were downloaded from `http://www.gis.leg.mn/html/download.htm` (last accessed November 7, 2008). All precinct-level data used in the figure and in the rest of this paper were downloaded from the web site of the Minnesota Secretary of State's web page. See `http://electionresults.sos.state.mn.us/20081104/`.

[14]The large white rectangle with an eastern-pointing-dagger at the northern tip of Minnesota is a single precinct that lies in the Koochiching Unorganized Territory. Much of the area of this precincts consists of a virtually uninhabited expanse of wetlands.
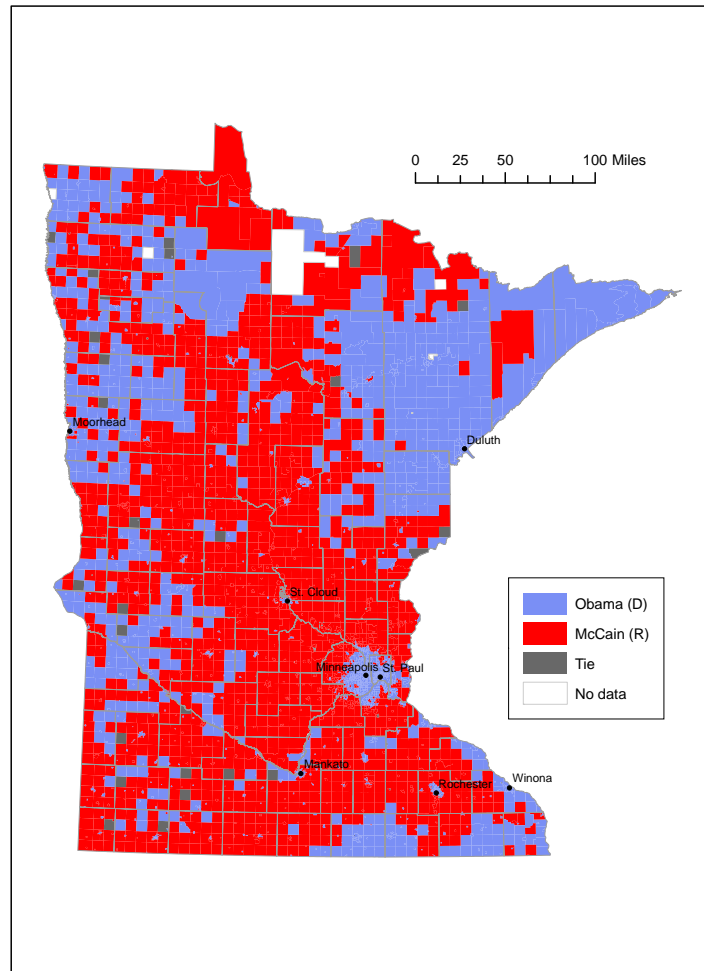
Figure 1: Outcome of the 2008 Presidential Race in Minnesota Precincts

Note: Each mapped unit is a precinct, and county boundaries are shown in gray.

and in some cases precinct boundaries follow township boundaries. In other cases a township may contain two precincts, a central town precinct and a precinct that includes all the township except for the town center. When two or more precincts have approximately equal Senate race residual vote rates, the boundaries between them are not distinguishable in Figure 2 unless these boundaries coincide with township boundaries.

It is clear from Figure 2 that there are no individual Minnesota counties that stand out as exceptionally noteworthy with respect to their Senate residual vote rates. This is an important result insofar as election administration in Minnesota is conducted at the county level. If, say, a single county employed idiosyncratic election procedures that had the effect of increasing the Senate residual vote rate, then we would almost certainly see evidence of this in Figure 2.

Further evidence of a lack of county-specific issues in Minnesota can be found in Figure 3. This figure contains a plot with 87 bars, one for each county in the state. Bar height describes Senate residual vote rates, bar color denotes whether a county had more Obama voters than McCain voters, and counties are ordered left to right by number of voters in the 2008 General Election.

It is hard to discern any sort of a meaningful pattern in Figure 3 beyond the fact that intermediate-sized counties in Minnesota tend to be disproportionately supportive of McCain (red bars are more prevalent within the middle of the figure). The heights of the bars, which are relevant for any analysis of residual votes in the Minnesota Senate race, are clearly scattered: some tall bars are red, some tall bars are blue; some tall bars are from small counties, some tall bars are from medium and large counties; and so forth.

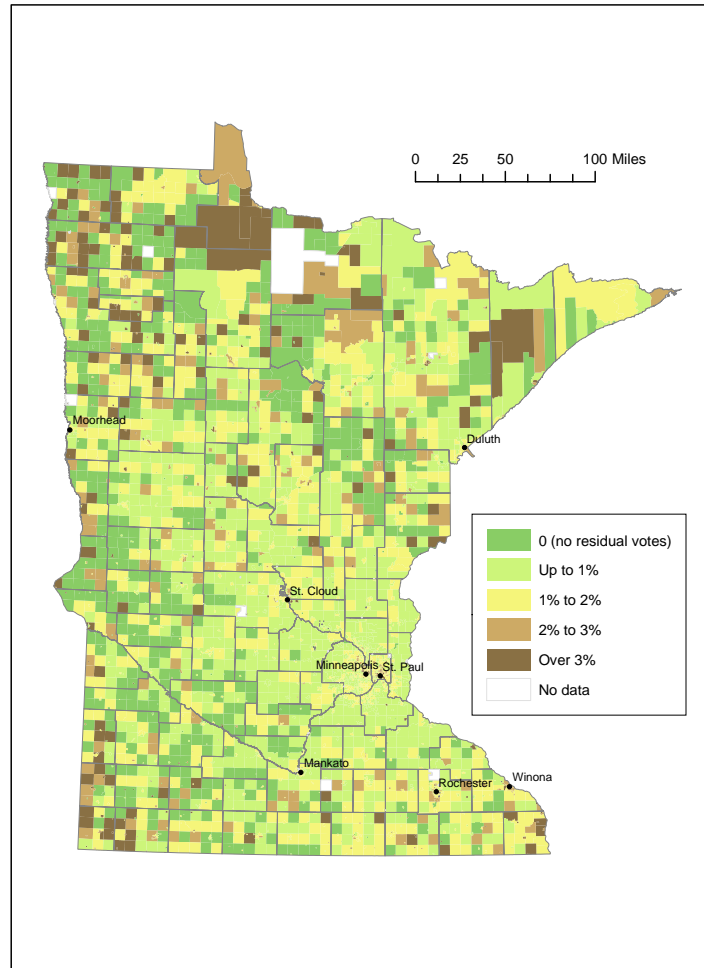It has been widely reported that most Senate race residual votes were cast

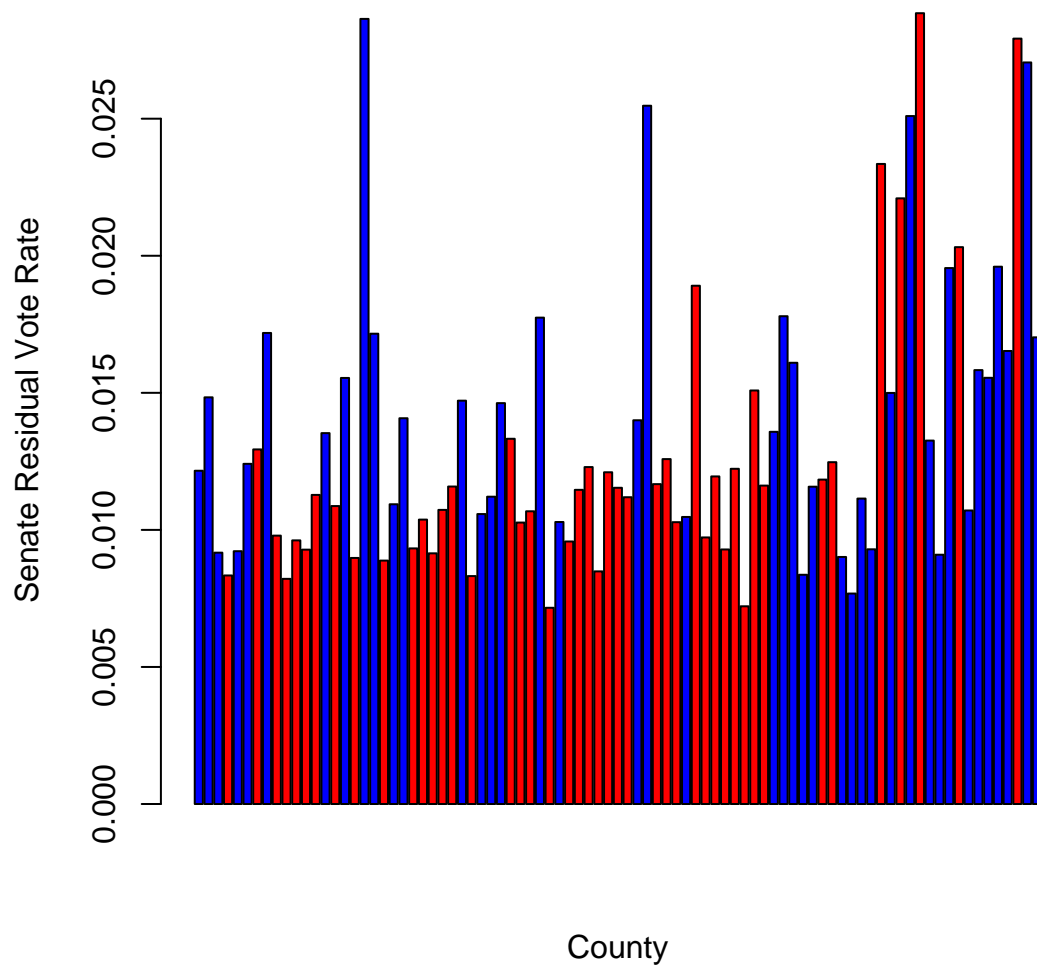Figure 2: Residual Vote Rate in the Minnesota Senate Race

Figure 3: Senate Residual Vote Rate in Minnesota Counties

Note: There are 87 counties in Minnesota, and they are ordered in the figure from greatest (left side) to smallest (right side) in terms of total number of votes cast in the 2008 General Election. A bar is colored blue if the number of Obama voters in the county exceeded the number of McCain voters and red otherwise.

in counties with majority support for Obama.[15] Nonetheless, this should not be interpreted as casting doubt on the election administration practices of counties that contain many Obama supporters. Rather, a more useful metric for assessing whether there were county-specific issues in the Coleman-Franken race is the residual vote rate within counties, and from Figures 2 and 3 we see that there do not appear to be any such issues.

Returning to precincts, there is some evidence in Figure 2 of a relative concentration of high residual vote rate precincts in the north and south parts of Minnesota. Relatedly, the further a precinct is from the urban center of Minneapolis-St. Paul, the greater the likelihood that it had a higher Senate residual vote rate.

Examining the Senate residual vote rate is useful for a rough assessment of whether any large areas in Minnesota appear to have unusually high residual vote rates. However, beyond this, the map of these rates in Figure 2 does not by itself contain any information that could allow us to determine if Senate residual votes were cast predominantly by voters who deliberately abstained from the Senate race or, in contrast, who failed to cast a recorded vote because of failure to understand how to vote in general. With this in mind, consider Figure 4, which describes the presidential residual vote rate as it varies by Minnesota precinct. The presidential race (contested in Minnesota by Obama, McCain, five third party candidates, and 16 write-in candidates), which appeared before the Senate race on all voters' ballots, can be thought of as a minimal test of voters' abilities to cast optical scan machine-recognizable votes: if a voter knew how to register his or her preferences

---

[15]See, for example, `http://www.startribune.com/politics/34116044.html?elr=KArksLckD8EQDUoaEyqyP4O:DW3ckUiD3aPc:_Yyc:aUnciaec8O7EyUsr` (last accessed November 11, 2008) and a memo by Charles Stewart, "Observations on the Close Minnesota Senate Election," downloaded from `http://web.mit.edu/cstewart/www/papers/minnesota_senate_memo_v2.pdf` (last accessed November 13, 2008).

on the presidential race, then it is natural to assume that he or she also knew how to act similarly in the Senate race.[16]

The patterns in Figure 4 are similar to those noted earlier in the discussion of the Senate race residual vote rate and the aforementioned Figure 2. In particular, no counties stick out as particularly anomalous with respect to presidential residual vote rates, and roughly speaking there is more presidential residual voting in the north and south of Minnesota than in the east-central area of the state near Minneapolis-St. Paul. A presidential residual vote rate bar plot analogous to Figure 3 is similarly devoid of obvious patterns (plot is available on request).

Nonetheless, there are differences between the presidential and Senate race residual vote rate, and the key difference is that, in most Minnesota precincts, the former is smaller. One can see this in the lighter color shadings in Figure 4 compared to colorings observed earlier.

The presidential vote rate allows us to bound the number of voters who knew how to cast machine-readable votes in the general election. If in a given precinct the presidential residual vote rate was 1%, then it follows that 99% of the voters in that location knew how to vote (or did not know but were lucky in the way that they indicated candidate choices). If the Senate race residual vote rate was 2%, then at least 1% of Senate ballots (maybe more or even all but almost certainly at least this many) are likely intentional abstentions. This suggests that the key metric for assessing the extent of intentional versus unintentional residual voting is the difference in a precinct between Senate and presidential residual vote rates.

---

[16]This is not to say that voters' abilities to register their preferences necessarily lack stochastic components. Suppose, for example, that a voter placed "X" marks in specific circles on his or her optical scan ballot as opposed to filling in these circles as directed. Some of the voter's marks could be sufficiently prominent so as to indicate a valid vote.
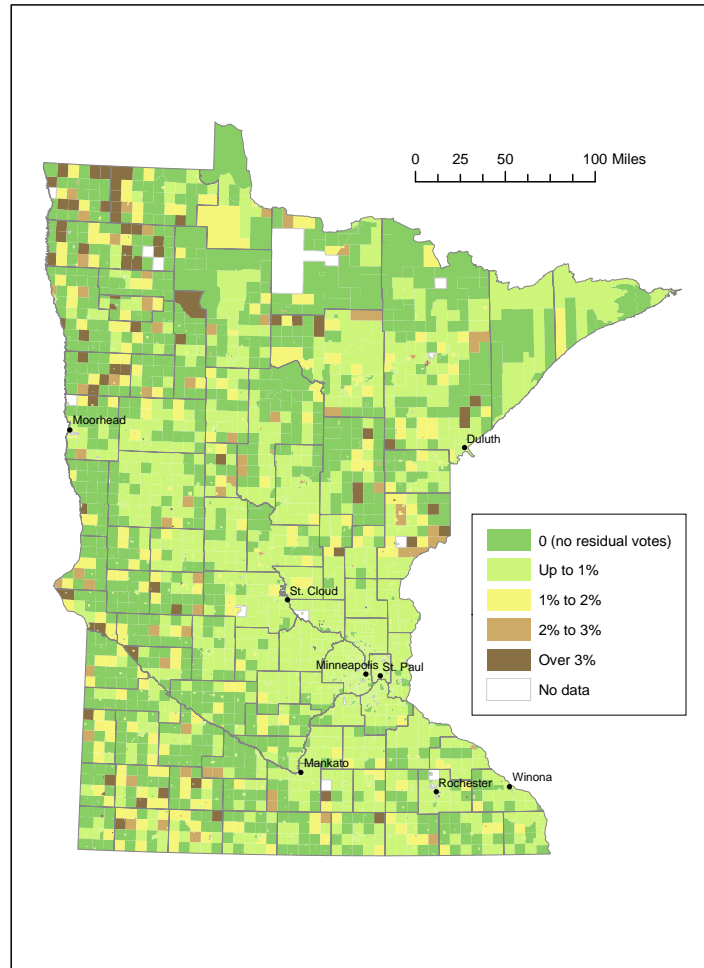
Figure 4: Residual Vote Rate for Presidential Race by Precinct.

This is the subject of the regression analyses that follow.

2 Explaining the Senate Race Residual Vote Rate

This section presents several sets of regression results. First considered are regressions that seek to explain patterns in the difference between the Senate and presidential residual vote rates among Minnesota precincts; these regressions speak to the phenomenon of intentional undervoting, as will be explained. The second set of regressions analyzes the presidential residual vote rate, and these shed light on unintentional residual votes.

2.1 Differences between Senate and Presidential Residual Vote Rates

The regressions in this section use a combination of precinct voting returns from 2008 and precinct returns based on the 2006 midterm elections in Minnesota.[17] To determine which precincts from 2006 still existed in 2008 (precincts tend to change boundaries based on population shifts) we overlaid electronic maps of the two sets of precincts. Based on such an overlay, we treat a 2008 precinct as equivalent to a 2006 precinct if at least 95% of the area of the 2006 precinct is contained in the 2008 precinct and vice versa. Any regression that uses a combination of 2006 and 2008 data is based on the set of 2008 precincts that satisfy that criteria.[18] We note that the voting technology used across Minnesota counties in 2006 was the same

---

[17]Regarding 2008 precincts in particular, all the regression analyses that follow use only those Minnesota precincts for which the Minnesota Secretary of State has reported (as of November 12, 2008) a non-zero number of voters. In addition, the regressions do not use the very few precincts that, also as of November 12, have negative presidential or Senate race residual vote rates associated with them. These presumably reflect reporting errors that will be corrected.

[18]One 2008 precinct, "Stillwater W-2 P-5" in Washington County, does not appear in the 2008 precinct shapefile and is dropped from the analysis.

as the technology used in the state in 2008.[19]

In addition, some of the regressions that appear shortly use census data from the 2000 census. These data are available at either the block level (so-called Stanford Form 1 variables) or the block group level (so-called Standard Form 3 variables).[20] We allocate census blocks (there are 200,222 in Minnesota) and block groups (4082) to 2008 precincts by overlaying maps (here, precinct maps and block or block group maps) and then allocating census counts by area.[21]

Table 1 includes estimates for a linear model in which precinct-level differences between Senate residual vote rates and presidential vote rates are regressed on various precinct features. The table, like the other regression tables that follow, displays coefficient estimates, estimated standard errors, t-statistics, and associated p-values.

The implications of Table 1 are as follows. First, the more Democratic a precinct as measured by the vote share it contributed to the Democratic candidate in the 2006 Minnesota gubernatorial race, the greater than Senate-presidential residual vote rate difference.[22] Is this consistent with the idea that Democratic

---

[19]Personal communication with the Minnesota Secretary of State via phone conversation on November 13, 2008.

[20]Minnesota's Public Law 94-171 data are not useful for our purposes because many of the precincts that existed prior to the 2000 census have changed boundaries in the intervening years.

[21]Ideally, of course, we would not be forced to rely on 2000 census data for research on the 2008 General Election. While the United States Census Bureau makes projections of various demographic variables by county during years that fall between decennial censuses, it does not update block or block group level data in this way. Lastly, the American Community Survey is not conducted at a level that allows its use in a precinct-level analysis.

[22]The 2006 Minnesota governor's race was won narrowly by Republican Tim Pawlenty (46.69%) over Democrat Mike Hatch (45.73%). This race is presumably more informative about a precinct's partisan disposition because the corresponding 2006 Senate race, between Democrat Amy Klobuchar (58.06%) and Republican Mark Kennedy (37.94%), was not close. See `http://electionresults.sos.state.mn.us/20061107/ElecRslts.asp?M=S&Races=0331` (last accessed November 13, 2008) for 2006 governor race results, and see `http://electionresults.sos.state.mn.us/20061107/ElecRslts.asp?M=S&Races=0102` (last accessed November 13, 2008) for 2006 Senate race results. If we include in

---

Table 1: Regression of Senate minus Presidential Residual Vote Rate Difference

	Estimate	Std. Error	t value	Pr>|t|
Intercept	-0.0089	0.0027	-3.25	0.0012
2006 Democratic Governor	0.0079	0.0012	6.70	0.0000
2006 Senate Residual Vote Rate	0.1392	0.0209	6.66	0.0000
2006 Governor Residual Vote Rate	0.0309	0.0221	1.40	0.1630
2008 E.D. Registration	0.0228	0.0016	14.29	0.0000
2006-2008 Fraction Change in Voters	0.0061	0.0006	10.16	0.0000
Precinct Count	0.0019	0.0015	1.25	0.2130
Central Count	-0.0010	0.0023	-0.42	0.6717

Note: The dependent variable is the Senate residual vote rate minus the presidential vote rate. Each observation is a Minnesota precinct, and observations are weighted by the number of voters in 2008. The variable "2006 Democratic Governor" refers to the fraction of total gubernatorial votes received by the Democratic gubernatorial candidate in the 2006 midterm election; "2006 Senate Residual Vote Rate" and "2006 Governor Residual Vote Rate" are self-explanatory; "2008 E.D. Registration" is the fraction of 2008 votes cast by election day registrants; "2006-2008 Fraction Change in Voters" is the rate of increase or decrease in the number of voters between the 2006 and 2008 elections; "Precinct Count" and "Central Count" are indicators for vote tabulating methods (hand count is the omitted category); and, 86 county indicator variables are in the model yet do not appear in the table.

voters tend to make more voting errors and, *ceteris paribus*, have disproportionately high residual vote rates? No. Rather, it is consistent with the idea that, the more Democratic a precinct, the greater the number of voters in it who voted for president and skipped the Coleman-Franken Senate contest.[23]

The use of a differenced dependent variable in Table 1 bears emphasis here. If a given type of voter was on average unable to cast machine-readable votes across all races on a ballot, then this voter type would *not* be highlighted in Table 1. Why? The effect of this voter type would on average be canceled out when the presidential residual vote rate was subtracted from the Senate residual vote rate.

Second, from Table 1 we observe that precincts with large Senate residual vote rates in 2006 had large Senate-president residual vote rate differences. This is consistent with the idea that precincts with large numbers of people who cast residual votes historically voted for president and did not vote in the Senate race.

Third, we see that what might be called a voter registration or turnout surge—measured by either the fraction of election day registrants in 2008 or an increase in the total number of voters in 2008 compared to the number in 2006—is associated with relatively more Senate race residual votes and relatively fewer presidential residual votes. What might explain this? One possibility is that turnout-inducing

---

the Table 1 regression and in other regression models the 2006 Senate Democratic vote share, our results do not change appreciably.

[23]It is important to emphasize here where is meant by consistent and what in fact is not known due to data limitations. Namely, the conclusion that ardent Democrats skipped the Senate race after voting for president *is* consistent with the results in Table 1. However, because we only have access to aggregate voting data at the precinct level, we cannot determine, say, if the same people who voted in the 2008 presidential race also voted in the 2008 Minnesota Senate race. Throughout this paper, therefore, we will consider whether statistical results are consistent with the two explanations for residual votes noted in the introduction, but the nature of the data that we have, and in particular the lack of ballot image data which could resolve the question of whether a few voters in Minnesota cast multiple residual votes or many voters cast a few, means that results' being consistent with an explanation does not guarantee with certainty that the explanation is correct.

excitement (where it occurred) based on the Obama-McCain presidential race did not translate into excitement at the Senate level. What might explain this, i.e., why the Coleman-Franken Senate race may not have generated the same level of excitement as the presidential contest, is beyond the scope of this paper.

There is no evidence in Table 1 of any vote counting technology effects. Likewise, county indicator variables, whose 86 associated estimates do not appear in the table, are relatively unremarkable (results available on request). While many of the county indicator variables are statistically significant at conventional confidence levels, none has a magnitude that is noteworthy. This is consistent with earlier comments made about Minnesota counties in the context of Figures 2, 3, and 4.

Table 2 contains a regression model similar to the one just discussed but the present model is augmented with precinct demographics based in part on census data. Briefly, this table has a number of notable implications.

First, there are no explicitly partisan effects insofar as the estimate for "2006 Democratic Governor" is not statistically significant at conventional confidence levels. This it not surprising: precinct-level demographics are presumably proxying for partisanship, and the next point is evidence of this.

Second, the more African-American a precinct, the greater the Senate-presidential residual vote rate difference. Did African-American voters turn out and vote in the presidential race and then skip the Coleman-Franken Senate race? The aggregate results in Table 2 cannot answer this question definitely: perhaps white voters who live in precincts with African-American voters behaved this way. Nonetheless, the African-American regression estimate is consistent with findings in Herron and Sekhon (2005), who show that white voters have disproportion-

20

Table 2: Regression of Senate minus Presidential Residual Vote Rate Difference with Census Data

	Estimate	Std. Error	t value	Pr>|t|
Intercept	-0.0063	0.0026	-2.47	0.0136
2006 Democratic Governor	-0.0011	0.0014	-0.74	0.4566
2006 Senate Residual Vote Rate	0.0920	0.0200	4.61	0.0000
2006 Governor Residual Vote Rate	0.0194	0.0206	0.94	0.3480
2008 E.D. Registration	0.0128	0.0017	7.43	0.0000
2006-2008 Fraction Change in Voters	0.0027	0.0006	4.41	0.0000
Indian Reservation	0.0002	0.0010	0.20	0.8416
African-American	0.0162	0.0026	6.28	0.0000
Indian/Native American	0.0003	0.0029	0.12	0.9071
Asian	0.0391	0.0037	10.46	0.0000
Hawaiian	-0.1063	0.0798	-1.33	0.1830
Other Race Group	0.0078	0.0581	0.13	0.8938
Multiple Race	0.0307	0.0272	1.13	0.2596
Hispanic	0.0003	0.0034	0.08	0.9397
Age 18-20	0.0404	0.0020	20.29	0.0000
Age 21-29	0.0036	0.0025	1.45	0.1458
Age 65 and Up	0.0140	0.0017	8.13	0.0000
Low Education	-0.0012	0.0019	-0.67	0.5031
Medium Education	0.0001	0.0004	0.18	0.8577
High Education	0.0131	0.0137	0.96	0.3377
Precinct Count	0.0011	0.0014	0.77	0.4433
Central Count	-0.0018	0.0021	-0.84	0.3983

Note: The dependent variable is the Senate residual vote rate minus the presidential vote rate, and observations are weighted by the number of voters in 2008. Variable definitions are the same as in Table 1 with the following additions: "Indian Reservation" is an indicator for whether a precinct lies in a reservation;[24] "African-American" is the fraction of the precinct that is African-American based on 2000 census data; other race groups and age groups are self-explanatory; "Low education" is the fraction of precinct residents who have not spent time in college; "Medium Education" is the fraction of precinct's residents whose education spans some college time through a bachelors degree; "High Education" is the fraction of a precinct's residents who have doctorates; and, 86 county indicator variables are in the model yet do not appear in the table.

ately high undervote rates in races with a dominant African-American candidate and African-American voters have disproportionately high undervote rates in races with a dominant white candidate.[25]

Third, even controlling for precinct demographics, there is a clear association between what might be called a turnout surge and excessive Senate residual votes. As before, this is consistent with the theory that excessive turnout for the 2008 general election in Minnesota was driven by the presidential contest and not the Coleman-Franken Senate contest.

Fourth, there is evidence in Table 2 of age effects insofar as precincts with large numbers of young and older residents had disproportionately high Senate-presidential residual vote rate differences. At the moment these effects, assuming that they reflect the behaviors of young and old voters, are hard to explain. They can be rationalized (one might speculate that older Democratic voters were not pleased with the candidacy of comedian Al Franken) but not in a particularly convincing fashion. We conjectured that the young voter effect in Table 2 might reflect the locations of colleges and universities across Minnesota; these institutions contain students who perhaps care about a presidential race yet are indifferent about Minnesota politics. Nonetheless, modeling the locations of educational institutions did not modulate the age effects in Table 2.

And fifth, there is no evidence in Table 2 for education effects. The importance of this non-finding will become clear shortly.

Probably the most notable finding of those noted above is that concerning the relationship between a precinct's racial composition and the difference in it be-

---

[25]This type of voting behavior cannot explain the Asian demographic effect in Table 2. This effect remains difficult to explain at present.

tween Senate and presidential residual vote rates. Of further relevance to this is Figure 5, which plots by Minnesota precinct the fraction of a precinct that is African-American versus the Senate-Presidential residual vote rate difference.
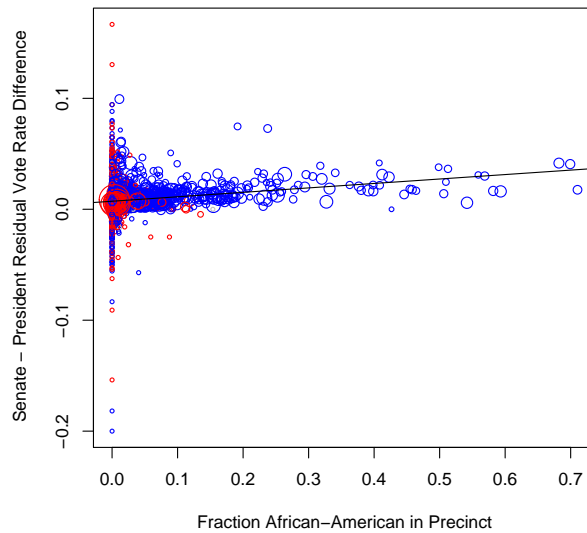

Figure 5: Racial Compositions of Precincts and the Senate-Presidential Residual Vote Rate Difference

Note: Each open circle represents one Minnesota precinct. A precinct's circle is colored blue if Obama had more votes than McCain in the precinct, red if the opposite obtained, and gray if a tie occurred. Precinct circles are roughly proportional to the number of voters who cast ballots in the 2008 General Election, and a regression line (weighted by number of voters) is superimposed on the blue, red, and gray circles.

One can see in Figure 5 that, as the fraction of a precinct becomes more African-American, the greater the gap between the Senate residual vote rate and the presidential residual vote rate. Of course we do not know which African-

23

Americans voted, we do not know if the same African-Americans (or whites) intentionally undervoted in both the presidential and Senate races, and we do not know which African-Americans in our precincts were citizens and/or registered voters. Nonetheless, the only behavioral explanation for Figure 5 that we currently know of would suggest that African-American voters were deliberately undervoting in the Coleman-Franken contest after voting in the presidential race.

The solid black regression line in Figure 5 is sloped upward (the slope is approximately 0.4). This positive slope is not driven by the small number of Minnesota precincts that were heavily African-American. If the slope is re-calculated using precincts that were at most 40% African American, the slope increases by a small amount (to 0.49).

Overall, what we see from Tables 1 and 2 is a fair amount of evidence that is consistent with explanations for intentional undervoting in the Minnesota Senate race. The findings about a turnout surge and about precincts with African-American residents seem particularly compelling.

2.2 Analyis of Presidential Residual Vote Rate

The evidence adduced above suggests that some Minnesotans deliberately skipped—intentionally undervoted in—the Coleman-Franken Senate race when voting in the 2008 General Election. Beyond this result, how might we know if some voters unintentionally undervoted in the Senate race, i.e., did not know how to indicate a candidate preference so that an optical scan machine could read it?

To address this question we analyze the presidential residual vote rate across Minnesota precincts. The reason we focus on the presidential contest in particular is because this race can be thought of as bounding the existence of voting problems.

Table 3: Regression of Presidential Residual Vote Rate

	Estimate	Std. Error	t value	Pr>|t|
Intercept	0.0071	0.0019	3.75	0.0002
2006 Democratic Governor	0.0029	0.0008	3.52	0.0004
2006 Senate Residual Vote Rate	0.0238	0.0145	1.64	0.1005
2006 Governor Residual Vote Rate	0.0468	0.0153	3.06	0.0023
2008 E.D. Registration	-0.0017	0.0011	-1.58	0.1138
2006-2008 Fraction Change in Voters	0.0000	0.0004	0.00	0.9964
Precinct Count	-0.0029	0.0011	-2.74	0.0062
Central Count	0.0003	0.0016	0.18	0.8548

Note: The dependent variable is the presidential residual vote rate, and observations are weighted by the number of voters in 2008. Variable definitions are the same as in Table 1. County indicator variables are in the model yet do not appear in the table.

If a voter had difficulty voting in the first race on his or her ballot, it is reasonable to assume that he or she suffered voting problems elsewhere, conditional on trying to vote elsewhere. Similarly, if a voter knew how to indicate in a machine-readable way a presidential preference, then presumably he or she knew how to indicate a Senate preference as well.

Table 3 contains results for a linear model that regresses presidential residual vote by precinct on a variety of precinct-level variables. The model includes county effects as well, but these are not remarkable and are not shown (results available on request).

Several conclusions follow from results displayed in Table 3. First, precincts that are disproportionately Democratic as measured by 2006 governor race voting behavior had disproportionately high presidential residual vote rates in 2008. We observed a Democratic precinct effect in our analysis of the Senate–presidential residual vote rate difference earlier (see Tables 1 and 2) but the interpretation of

such an effect is different when the relevant dependent variable is the 2008 presidential residual vote rate as opposed to a difference in rates. Namely, the Democratic effect in Table 3 is consistent with the notion that Democratic voters are more likely than their Republican counterparts to cast unintentional undervotes (e.g., Herron and Sekhon, 2003; Tomz and van Houweling, 2003; Ansolabehere and Stewart III, 2005).

Second, there is evidence in Table 3 that precincts with previous undervoters had high presidential residual vote rates in 2008. This could be interpreted as either evidence of voters who regularly choose to undervote in top races or who regularly try but fail to communicate their candidate choices in machine-readable manners.

Finally, we see from the second to last row of Table 3 that precincts with Minnesota's implementation of second chance voting had lower presidential undervote rates. These precincts presumably have no or almost no presidential overvotes, and the "Precinct Count" finding is consistent with previous research which shows that precinct count optical scan voting tends to produce relatively low residual vote rates (because, it is believed, this technology alerts voters to overvotes and possibly undervotes).[26]

Presumably Democrats do not spoil their ballots on account of partisanship *per se*. Thus, Table 4 includes various precinct-level demographics in a regression model of the presidential residual vote rate.

There are several notable findings from Table 4. The first is that, once racial demographics are included, precinct partisanship as measured by support for the Democratic gubernatorial candidate in 2006 is no longer statistically significant at

---

[26]An alternative explanation is that voters who regularly and intentionally overvote happen to live in places that employ central count voting. This is possible but seems implausible.

Table 4: Regression of Presidential Residual Vote Rate with Census Data

	Estimate	Std. Error	t value	Pr>|t|
Intercept	0.0055	0.0019	2.92	0.0035
2006 Democratic Governor	0.0011	0.0011	1.03	0.3053
2006 Senate Residual Vote Rate	0.0258	0.0148	1.75	0.0803
2006 Governor Residual Vote Rate	0.0330	0.0153	2.16	0.0305
2008 E.D. Registration	-0.0023	0.0013	-1.82	0.0687
2006-2008 Fraction Change in Voters	0.0006	0.0004	1.34	0.1808
Indian Reservation	-0.0014	0.0007	-1.94	0.0525
African-American	0.0007	0.0019	0.34	0.7326
Indian/Native American	0.0029	0.0021	1.36	0.1745
Asian	-0.0015	0.0028	-0.56	0.5780
Hawaiian	0.0211	0.0590	0.36	0.7210
Other Race Group	-0.0248	0.0429	-0.58	0.5628
Multiple Race	0.0024	0.0201	0.12	0.9034
Hispanic	0.0022	0.0025	0.89	0.3731
Age 18-20	-0.0000	0.0015	-0.02	0.9818
Age 21-29	0.0013	0.0018	0.72	0.4702
Age 65 and Up	0.0127	0.0013	9.95	0.0000
Low Education	0.0005	0.0014	0.38	0.7022
Medium Education	0.0003	0.0003	0.85	0.3971
High Education	-0.0370	0.0101	-3.65	0.0003
Precinct Count	-0.0032	0.0010	-3.05	0.0023
Central Count	0.0004	0.0015	0.27	0.7863

Note: The dependent variable is the presidential residual vote rate, and observations are weighted by the number of voters in 2008. Variable definitions are the same as in Table 2. County indicator variables are in the model yet do not appear in the table.

conventional confidence levels. This parallels earlier results. Second, even with demographics, precinct count precincts have, *ceteris paribus*, relatively lower residual vote.

Third, precincts with large concentrations of older residents have disproportionately high presidential residual vote rates. The most compelling explanation for this result, it seems, is that older voters had greater difficulty make machine-readable marks on their ballots. And, fourth, we see that precincts with the most highly educated residents had the fewest presidential residual votes. Like the ostensible older voter result, this education-based result, which can also be found in Knack and Kropf (2003*b*), is conceivably related to a voter's facility at indicating machine-readable candidate preferences on a paper ballot.

If the above conjectures are correct, then it follows that a non-trivial fraction of the Senate race residual votes contain unintentional undervotes with Senate-related ballot marks. Whether these marks, should they exist, ultimately indicate voter intent is a matter of Minnesota state law. Nonetheless, the statistics here suggest that, independently of whether these hypothetical ballots are legally determined to have evidence of intent, the voters who cast them intended to participate in the Coleman-Franken Senate race.

It is important to note that the coefficient estimates in Tables 3 and 4 are not extremely large. For instance, and with respect to the latter table, if a precinct were to move from zero percent age 65 and up voters to all such residents, then the presidential residual vote rate would rise by approximately 1.2% (assuming all such residents vote). If 700 ballots were cast in this precinct (this number is slightly more than the mean number of ballots cast per precinct in the 2008 General Election), then this would imply that perhaps seven ballots represent voters who

did not know how to cast votes in a machine-readable way. Of the seven associated voters, our earlier results indicate that some will have intentionally undervoted in the Senate race. Only the remainder from this group will be voters who attempted to cast Senate race votes. Whatever figure one uses for this latter quantity will be small, particularly in light of the fact that the number seven was based on a precinct completely changing from non-older votes to only older votes.

Nonetheless, of the votes that have valid Senate marks as defined by Minnesota state law, we suspect that a small plurality will be supportive of Franken. This reflects the fact that, historically speaking, Democrats accidentally spoil their ballots at higher rates than Republicans; we have seen evidence of this phenomenon as applied to Minnesota in Table 3; and highly educated voters according to exit polls broke for Coleman over Franken.[27] To the extent that there are any voter education effects at all, then, they will almost certainly be in Franken's favor.

Regarding our evidence on precincts with older residents ages 65 and up, exit polls imply that they preferred Coleman to Franken by 43% to 42%. This difference is almost certainly statistically negligible. Accordingly, then, we suspect that the invalid Senate votes cast by older voters and recovered in a recount will break approximately evenly for Coleman and Franken.

3 The Impending Recount

Clearly, the most important outcome from the impending recount of the Minnesota Senate race will be the determination of whether Norm Coleman or Al Franken will be serving the citizens of Minnesota in Congress for the next six years. Based

---

[27]See http://www.cnn.com/ELECTION/2008/results/polls/#val=MNS01p1 (last accessed November 13, 2008).

on the currently available data which has been reviewed here, it appears plausible that Coleman's 206-vote lead as of November 12, 2008 may be reversed by the recount process. But we also note that the recount itself will represent an ideal natural experiment that sheds light on the dynamics of residual voting in United States elections. What makes this experiment virtually unprecedented in United States electoral history is the combination in a single state of (nearly) uniform voting procedures, the availability of actual paper ballots as marked by the voters, and the systematic re-examination by hand of literally every ballot cast. Recounts are not new, and the features of elections that lead to accurate vote counts have been studied by others (Ansolabehere and Reeves, 2004), but what is currently transpiring in Minnesota is nonetheless a rare conjunction in United States elections.

We cautiously predict that the recount will find a small net increase in votes in favor of Franken (due to unintentional residual votes that are resolved through the recount process) and a larger number of presumably intentional residual votes that reflect the choices of voters who abstained from voting in the Senate race. Should the outcome of the recount follow this pattern, we would conclude that our models of the residual vote phenomenon are reasonable and that differences in residual vote rates on a single ballot can support inferences about voter intent. On the other hand, if the recount produces a large majority of additional votes for Coleman, or a much larger fraction of unintentional than intentional residual votes, then we will need to re-evaluate what we think we know about analyzing differences in residual vote rates. In either case, the results of the recount will certainly be of interest to political scientists for reasons beyond the simple question of who will be representing the state of Minnesota in Washington starting in January, 2009.

References

Ansolabehere, Stephen and Andrew Reeves. 2004. "Using Recounts to Measure the Accuracy of Vote Tabulations: Evidence from New Hampshire Elections 1946-2002." Working paper, Massachusetts Institute of Technology.

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

Brady, Henry E., Justin Buchler, Matt Jarvis and 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.

Frisina, Laurin, Michael C. Herron, James Honaker and Jeffrey B. Lewis. 2008. "Ballot Formats, Touchscreens, and Undervotes: A Study of the 2006 Midterm Elections in Florida." *Election Law Journal* 7(1):25–47.

Herrnson, Paul S., Richard G. Niemi, Michael J. Hanmer, Benjamin B. Bederson, Frederick C. Conrad and Michael W. Traugott. 2007. *Voting Technology: The Not-So-Simple Act of Casting a Ballot*. Washington, D.C.: Brookings Institution Press.

Herron, Michael C. and Jasjeet S. Sekhon. 2003. "Overvoting and Representation: An examination of overvoted presidential ballots in Broward and Miami-Dade counties." *Electoral Studies* 22(1):21–47.

Herron, Michael C. and 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.

Kimball, David C. and Martha Kropf. 2005. "Ballot Design and Unrecorded Votes on Paper-Based Ballots." *Public Opinion Quarterly* 69(4):508–529.

Knack, Stephen and Martha Kropf. 2003*a*. "Roll-Off at the Top of the Ballot: International Undervoting in American Presidential Elections." *Politics & Policy* 31(4):575–595.

Knack, Stephen and Martha Kropf. 2003*b*. "Voided Ballots in the 1996 Presidential Election: A County-Level Analysis." *Journal of Politics* 65(3):881–897.

Niemi, Richard G. and Paul S. Herrnson. 2003. "Beyond the Butterfly: The Complexity of U.S. Ballots." *Perspectives on Politics* 1(2):317–26.

Tomz, Michael and 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.