Alvin Greene? Who? How did he win the United States Senate nomination in South Carolina?*

Joseph Bafumi† Michael C. Herron‡ Seth J. Hill§ Jeffrey B. Lewis¶

February 15, 2012

Abstract

Alvin Greene surprised the political world when he handily defeated Vic Rawl for the United States Senate nomination in the 2010 Democratic Primary in South Carolina. Greene had not run a campaign during the primary and appears to have been almost completely unknown prior to his surprise victory. Greene's win over Rawl, who had served eight years in the South Carolina House, was previously a circuit judge, and had in fact run a legitimate primary campaign, raised a variety of questions about how Greene could have managed to generate so much voter support. In light of lingering concerns that Greene's victory was due to malfeasance of some sort, we analyze both ballot-level and precinct-level voting data with an eye toward determining whether the 2010 Democratic Senate Primary in South Carolina appears problematic. We find that voting patterns in Greene's victory over Rawl do not exhibit unusual peculiarities and in fact are consistent with the types of regularities observed in American elections. Rawl is white and Greene is black, and this difference played a major role in Greene's victory. While this victory may have been a surprise, voters in the Greene vs. Rawl primary appear in retrospect to have behaved similarly to voters in other elections in the United States, thus lending legitimacy to Greene's win.

*The authors thank Rohail Premjee for research assistance.

†Assistant Professor of Government, Dartmouth College. 6108 Silsby Hall, Hanover, NH 03755 (jbafumi@gmail.com).

‡Visiting Professor of Applied Methods, Hertie School of Governance, and Professor of Government, Dartmouth College. Friedrichstraße 180, 10117 Berlin, Germany (michael.c.herron@dartmouth.edu).

§Postdoctoral Associate, Center for the Study of American Politics and Institution for Social and Policy Studies, Yale University. 77 Prospect Street, P.O. Box 208209, New Haven, CT 06520-8209 (seth.hill@yale.edu).

¶Associate Professor of Political Science, University of California, Los Angeles. 4289 Bunche Hall, Box 951472, Los Angeles, CA 90095-1472 (jblewis@ucla.edu).

Alvin Greene surprised the political world in June, 2010, when he handily defeated Vic Rawl for the Democratic United States Senate nomination in South Carolina. Leading up to the election, Greene had not run what anyone would call a serious campaign or even what might pass as a non-serious campaign. Greene nonetheless won nearly 60% of the primary vote over an opponent, Vic Rawl, who was the clear party favorite, at the time of the primary was sitting on the Charlestown County Council, had served eight years in the South Carolina House, was previously a circuit judge, and had run a legitimate primary campaign having raised approximately $186,000.[1] Greene's victory was a remarkable upset, and not atypical in the time frame immediately following Rawl's defeat was a headline from *The New York Times* that wondered, "Who's Alvin Greene? State Asks After Vote."[2] Upon losing the race Rawl himself averred, "I would've liked very much to be a candidate against Jim DeMint... I never saw [Alvin Greene]. I've still never met him."[3]

The story behind Greene's victory is enigmatic. As a young man—32 years old when nominated—Greene had led a relatively unheralded life prior to his primary win. During the Senate Primary campaign, Greene lived with his father, had no job, and owned neither a computer nor a cell phone. As if these factors were not enough of a political handicap, Greene had a checkered military history, having been involuntarily discharged from both the Army and the Air Force, and at the time of the primary was facing an obscenity charge involving online pornography.[4]

Greene's economic circumstances and unusual profile for a political candidate led to suspicion over how Greene had managed to pay the $10,400 filing fee required by South Carolina to qualify for ballot placement opposite Rawl.[5] James Clyburn, a prominent member of the United States House of Representatives from South Carolina, was quoted in the aftermath of Greene's victory as saying, "I don't know if he was a Republican plant... What is an unemployed guy doing paying

[1]For details on Rawl see "Unknown Dem. Shocks Party by Taking S.C. Primary," *CBS News*, June 9, 2010, available at `http://www.cbsnews.com/stories/2010/06/09/politics/main6564215.shtml` (last accessed August 31, 2011) and "Rawl begins Senate campaign," *The Post and Courier*, March 9, 2010, available at `http://www.postandcourier.com/news/2010/mar/09/09rawl/` (last accessed August 31, 2011).

[2]See "Who's Alvin Greene? State Asks After Vote," *The New York Times*, June 11, 2010, available at `http://www.nytimes.com/2010/06/12/us/politics/12greene.html` (last accessed August 31, 2011).

[3]See "Alvin Greene UPSET: Mystery Man Stuns In South Carolina Senate Primary," *The Huffington Post*, June 8, 2010, available at `http://www.huffingtonpost.com/2010/06/08/alvin-greene-upsets-vic-r_n_605365.html` (last accessed September 8, 2010).

[4]For details on Greene's life, see "At Home with South Carolina's Alvin Greene," *Time*, June 17, 2010, available at `http://www.time.com/time/politics/article/0,8599,1997061,00.html` (last accessed August 31, 2011).

[5]For details on the South Carolina filing fee, see "Dem Senate Campaign Committee Stays Out of Alvin Greene Controversy," *CBS News POLITICALHOTSHEET*, June 10, 2010, available at `http://www.cbsnews.com/8301-503544_162-20007403-503544.html` (last accessed August 31, 2011).

$10,000 to run for the United States Senate? That just doesn't add up."[6] Nonetheless, a subsequent investigation by the South Carolina Law Enforcement Division revealed that the money for Greene's filing fee came from savings from Greene's military service salary.[7]

Notwithstanding the fact that Greene was found to have qualified legitimately for the Democratic primary ballot, his electoral success spawned various allegations of misconduct. These allegations ranged from claims about improper vote counting to assertions that Greene's victory was due in part to the way he spelled his last name.

Rawl himself put forth claims of problematic vote tabulating machinery,[8] and in conjunction with his allegations formally appealed his election loss to the South Carolina Democratic Party. Rawl's lawyer during the appeal asserted that, "[H]is best guess about what happened is that every fourth vote for Rawl somehow was switched to a vote for Greene on the state's voting machines."[9] Rawl's appeal was denied by the party's Executive Committee, and subsequent to that Greene ran against, and lost convincingly to, incumbent Senator Jim DeMint in November, 2010.[10]

South Carolina state Senator Robert Ford asserted that black voters—according to the 2010 U.S. Census, over 26% of single-race, voting age South Carolina residents are black—would have known that Greene was black because his last name ends with an "e" as opposed to being simply, "Green."[11] Ford's claim about the matter of "Greene" versus "Green" happens to be wrong (we provide details later), but certainly Ford was right to highlight the importance of race in South Carolina, a state with an extensive racial history and whose Democratic party is majority African-American. Moreover, as we show, an individual named Greene is more likely to be black than an individual named Rawl.[12] Spelling notwithstanding, local South Carolina papers covered the

[6]See "Clyburn: Greene may be GOP plant," *Politico*, June 10, 2010, available at `http://www.politico.com/news/stories/0610/38358.html` (last accessed August 31, 2011).

[7]See "Senate candidate Greene paid his own S.C. filing fee," *The Seattle Times*, July 10, 2010, available at `http://seattletimes.nwsource.com/html/nationworld/2012327943_fee11.html` (last accessed September 5, 2011).

[8]See "Rawl files primary protest," *The Post and Courier*, June 15, 2010, available at `http://www.postandcourier.com/news/2010/jun/15/rawl-files-protest-in-primary-upset` (last accessed August 31, 2011).

[9]See "Can we count on machines?," *The Post and Courier*, October 18, 2010, available at `http://www.postandcourier.com/news/2010/oct/18/can-we-count-on-machines/` (last accessed September 8, 2011).

[10]See "Rawl's bid to overturn Greene win is over," June 18, 2010, *The Post and Courier*, available at `http://www.postandcourier.com/news/2010/jun/18/itsover/` (last accessed September 8, 2011).

[11]See "Update: Clyburn calls for probe in Senate race," *The Post and Courier*, June 10, 2010, available at `http://www.postandcourier.com/news/2010/jun/10/strange-twist-in-senate-race/` (last accessed August 31, 2011). According to the 2010 census redistricting data (PL 94-171) for South Carolina, there were 939,461 single-race, black individuals of voting age in the state and 3,506,755 total single-race individuals of voting age. The data are available from `http://www2.census.gov/census_2010/01-Redistricting_File--PL_94-171/South_Carolina/sc2010.pl.zip` (last accessed September 1, 2011).

[12]Various press accounts in the aftermath of the primary noted that some voters may have confused Greene with

Democratic and Republican United States Senate primaries with background on the candidates and even the occasional photograph.[13]

In light of the controversy surrounding Greene's election victory and the underlying possibility of malfeasance, we consider here two alternative explanations for Greene's electoral success. The first explanation is that, despite its surprising nature, Greene's win was in retrospect plausible and rationalizable without recourse to assertions of irregularities. In contrast, the second explanation for Greene's victory is that there was malfeasance associated with some aspect of balloting in the 2010 South Carolina Democratic Senate Primary. Such a breakdown would clearly be more troubling than the existence of a surprising election outcome in which a party favorite lost to an unheralded opponent.

These two competing explanations point toward a well-posed research question—did voting in the South Carolina Senate Democratic Primary follow patterns that are typical for American elections?—and in seeking to answer this question we find that voting patterns in Greene's victory over Rawl do not exhibit peculiarities and indeed are consistent with the types of regularities observed in American elections. Race, as we show, played a major role in Greene's victory. Voters who supported Greene also tended to vote for other black candidates competing in the 2010 Democratic Primary. We also show that voters who supported Greene—who was listed first among Senate Primary candidates—were disproportionately likely to vote for other first-listed candidates, and this phenomenon is consistent with research on ballot order effects.

Overall, our data provide evidence that Greene's primary win was legitimate, and, despite our consideration of a variety of alternatives, we find a lack of evidence that this win was marred by irregularities. While we cannot prove a complete absence of the latter, our results show that little more than standard regularities in American elections are needed to rationalize Greene's victory.

Although one might think that South Carolina's reputation as a scandal-laden state lends a farcical aspect to Greene's primary victory over the better-known Vic Rawl, the question of whether a surprise election victory is legitimate can strike at the heart of democracy (e.g., Herron, Mebane,

Al Green, a black gospel musician. See "Alvin Greene: America's most unlikely politician," *The Guardian*, July 6, 2010, available at `http://www.guardian.co.uk/world/2010/jul/06/alvin-greene-south-carolina-senate` (last accessed February 14, 2012).

[13] A picture of Greene appears in "Constitution, jobs among Senate race concerns," *The Post and Courier*, May 15, 2010, available at `http://www.postandcourier.com/news/2010/may/15/constitution-jobs-among-senate-race-concerns` (last accessed August 31, 2011).

Jr. and Wand 2008). Democratic political systems depend on the public having faith in election outcomes, and lingering uncertainty over whether a victory was rightfully earned is presumably damaging to democratic institutions.

Competing Theories of Greene's Electoral Victory

The aftermath of Greene's upset victory in June, 2010, occasioned a variety of claims about the Democratic Senate Primary. As noted earlier, some claims cast aspersions on Greene's win and imply or assert directly that the victory was illegitimate. In contrast, others explain the victory and purport that it was ultimately legitimate. We now review these claims and generate testable implications which enable us to offer a judgment on Greene's win over Rawl.[14]

Legitimate Reasons for Greene's Win

What does it mean to posit that Greene's victory was legitimate? We argue that, if Greene's victory is consistent with what researchers have found to be typical electoral processes in the United States, the election result is legitimate. Specifically, if a given factor that helped Greene win tends also to help candidates like Greene win in other election environments, then the factor is supportive of a legitimate election result in the Greene vs. Rawl contest. Thus, to assess whether Greene's victory was legitimate, we need to consider a set of factors associated with typical election processes and then ask what these factors say about Greene.

Perhaps the most consequential variable in modern electoral contests is partisanship (e.g., Bartels 2000; Green, Palmquist and Schickler 2002), i.e., the willingness of a voter to support a candidate, *ceteris paribus*, when there is congruence between the candidate and said voter's party. The import of partisanship notwithstanding, this concept has little to say about the United State Senate Democratic Primary in South Carolina because both its two contestants were Democrats.

A second characteristic that has been shown to influence vote choice is ideology (Downs 1957; Enelow and Hinich 1984; Jessee 2009). Ideology could have driven vote choices in the South Carolina Senate Primary if the policies of Rawl and Greene had differed so that more liberal Democratic voters preferred one of these two candidates and more moderate Democratic voters,

[14]There is not much of a literature on Senate primaries *per se*. An exception is Abramowitz and Segal (1992), who show that Senate hopefuls involved in significant controversies are at a disadvantage in primary battles and that candidates who hold high elected office or are well-funded have an advantage. These factors, of course, worked against Greene insofar as he spent almost nothing during the course of the Democratic Primary, had never held public office, and was associated with an online pornography scandal. Perhaps Greene's saving grace was that, with so few people taking him seriously, the media appeared to have covered neither the scandal nor Greene's political inexperience.

the other. Given that Greene never articulated a clear ideology, it seems unlikely that an ideological distinction between Rawl and Greene explains Greene's win.

The absence of partisanship and ideology in the Senate Primary places this low-salience contest in a class of low-information elections such as judicial elections or ballot propositions. In the absence of using party as heuristic, research has shown that voters turn to other cues to help make appropriate choices, from group-based endorsements (Berelson, Lazarsfeld and McPhee 1954; Lupia 1994) to gender (McDermott 1997), occupation (McDermott 2005), and race (Sigelman et al. 1995). Race seems likely to be the most relevant characteristic in the Greene vs. Rawl contest because black voters make up the majority of the South Carolina Democratic primary electorate. African-Americans appear to prefer black candidates to white candidates (Bobo and Gilliam 1990; Dawson 1995; Banducci, Donovan and Karp 2004; Tesler and Sears 2010). Beyond candidate choice, the influence of race is evident in many aspects of contemporary American elections, e.g., undervote rates (Herron and Sekhon 2005) and turnout rates (Gay 2001).

If race were a key factor in the Greene vs. Rawl contest, then support for Greene should be correlated with support for other black candidates who were running for office. Moreover, places in South Carolina with more non-white residents should be associated with higher support levels for Greene, *ceteris paribus*.[15]

Greene's placement on the Democratic primary ballot—his name appeared before Rawl's on all ballots—may have aided Greene as well. Indeed, South Carolina Democratic Party Chairwoman Carol Fowler explained Greene's surprising win this way.[16] Being listed first on a ballot can be worth one to three percentage points (Koppell and Steen 2004; Ho and Imai 2006), and, while name order on a ballot may not be the most appealing reason for a candidate to win an election, we would not label a candidate name order effect in the Senate Primary illegitimate. Such an effect, were it to exist, would be consistent with the rules in place in South Carolina at the time of the 2010 Democratic Primary. And, were the effect to exist, voters who cast votes for Greene should have

[15]Some Democratic Primary voters may infer ideology from candidate race on the assumption that black South Carolinians are on average more liberal than white South Carolinians. Accordingly we cannot distinguish voters who prefer black candidates solely because they are black from voters who prefer black candidates because they are more likely to prefer liberal policies when elected.

[16]See "Accused Lewd Guy Alvin Greene Is South Carolina's Democratic Candidate for Senate," *Vanity Fair Daily*, June 9, 2010, available at `http://www.vanityfair.com/online/daily/2010/06/accused-lewd-guy-alvin-greene-is-south-carolinas-democratic-candidate-for-senate` (last accessed September 1, 2011) and "Who Is Alvin Greene?," *Mother Jones*, June 8, 2010, available at `http://motherjones.com/mojo/2010/06/alvin-greene-south-carolina` (last accessed September 1, 2011).

disproportionately voted for other first-listed candidates.

Illegitimate Reasons for Greene's Win

While one could argue that literally any influence on Greene's win that is not known to be legitimate is by construction illegitimate, this would trivialize the notion of illegitimacy. Therefore, in thinking about what an illegitimate Greene victory might look like, we draw on some of the post-election claims by those concerned about the 2010 Democratic Primary in South Carolina. These claims are centered on two key factors: voting technology problems and strategic Republicans.

With respect to the former, one could argue, as Rawl's lawyer did, that Greene won the Democratic Primary because some number of votes that were intended for Rawl were actually counted for (or switched to) Greene. This claim is in principle difficult to test because the only record of a voter's intended vote is the actual vote recorded by his or her ballot, a fact perhaps not lost upon Rawl's attorneys. To the extent we can derive an empirical proposition that follows from the claim, we test by implication through our exploration of vote patterns across individual ballots: vote-switching errors on in the Greene vs. Rawl contest should manifest themselves through odd patterns that are inconsistent with patterns of voting behavior found in other contests.

Voting technology problems have historically often made themselves evident in undervotes (e.g., Frisina et al. 2008). Because the space of possible problematic undervote patterns is extensive, we test for unusual undervotes in two ways. First, we examine whether precincts in South Carolina had Senate Primary undervote rates far out of line compared to undervote rates in the other two statewide contests in the Democratic Primary. And second, we consider whether race-base undervote patterns in the Senate Primary appear different than the corresponding patterns observed in other statewide contests in the Democratic Primary.

A second potential source of illegitimacy in Greene's win is the possibility that Republican voters took advantage of South Carolina's open primary system to vote in the Democratic Primary and purposefully elect a weak challenger to then-incumbent Senator Jim DeMint. Such an effort would have required coordination across the state, and we know of no documented evidence that efforts in this direction existed. Nonetheless, that strategic Republicans could have acted in this way is possible. Although the idea of Republican imposters might not be interpreted as illegitimate because primary voters in South Carolina are within the letter of the law to choose either Democratic

or Republican ballots, cross-party primary voting to nominate purposefully a weak candidate is presumably not the intent of South Carolina's primary system. As such we cautiously include the Republican impostor idea as a possible source of an illegitimate Greene win.

There are three observable implication of the proposition that Greene won his primary because Republicans voted for him. First, if the proposition holds then we should see more Republicans than Democrats in the 2010 Primary compared to past primaries. Second, we should also observe voters who participated in a Republican Primary in 2008 participating in the 2010 Democratic Primary. And third, we should observe a large number of individual ballots that have votes for Greene and for either no other candidates or obviously weak candidates.[17]

A third potential source of illegitimacy is a Republican operative's having paid Greene's candidate filing fee. This seems an odd choice of potential fraud for Republicans given the likelihood of *any* Democrat defeating the popular incumbent Jim DeMint in a general election. Nonetheless, while our election data do not speak to the filing fee question, having a non-family member pay the Senate filing fee for Greene would have violated South Carolina law[18] and might also have violated federal campaign finance contribution laws. As we noted earlier, the South Carolina Law Enforcement Division did investigate the filing fee matter, finding on July 9, 2010, that "[M]onies spent for Greene's filing fee were the candidate's personal funds and, therefore, no laws were violated in association with Greene's payment."[19] We are not in position to improve upon this investigation.

In summary, the competing explanations—legitimate and illegitimate, setting aside the filing fee matter—for Greene's victory have observable implications. While no test, statistical or not, can prove beyond doubt that a given election result is legitimate, we can nonetheless consider whether legitimate sources for a Greene election victory have empirical support. If none of our legitimate factors appears to have been present in Greene's election, or if some of the illegitimate factors have empirical support, then concern over the legitimacy of Greene's win is merited. If, on the other hand, we find that patterns among Greene's votes look similar to patterns found in other American

[17]This argument requires that Republican crossover voters, if they existed, did not have access to a full list of "upset" candidates to vote for in the Democratic Primary; otherwise, more than just the Greene victory would have engendered controversy after the election.

[18]S.C. Code Ann. §7-25-200 (1976). Interestingly, this kind of interference had been suggested before, motivating the implementation of the filing fee law. See `http://www.nytimes.com/1992/04/19/us/the-1992-campaign-racial-politics-in-south-carolina-a-tale-of-campaign-trickery.html` (last accessed February 13, 2012).

[19]See the news release of the report at `http://www.sled.state.sc.us/CISystem/Images/NewsPress/SNP0571.pdf` (last accessed February 7, 2012).

elections, the overall legitimacy of Greene's victory becomes a more compelling narrative.

Overview of the 2010 South Carolina Democratic Primary

When South Carolinians arrived at the polls on June 8, 2010, they first had to choose between voting in the Democratic or in the Republican Primary. Total turnout in the Primary was 623,418, of whom 197,593 (almost 32%) selected Democratic ballots.

Table 1 lists the statewide and United States House contests from the Democratic Primary, races of candidates, and overall vote shares. There were neither Hispanic nor Asian candidates among the contests in Table 1, all of which are listed in the order that they appeared on the Democratic Primary ballot.

Of the gubernatorial candidates Robert Ford was listed first, and he was the only black candidate running for governor. After the gubernatorial contest, voters were asked to select one of two candidates to be nominated to run for South Carolina State Superintendent of Education. Tom Thompson, listed second among the two education candidates, is black.

Table 1: Top Contests in the 2010 Democratic Primary

Contest	Candidate	Race	Vote Share
Governor	Robert Ford	black	0.1802
	Jim Rex	white	0.2303
	Vincent Shaheen	white	0.5896
Superintendent of Education	**Frank Holleman**	white	0.5616
	Tom Thompson	black	0.4384
United States Senate	**Alvin Greene**	black	0.5896
	Vic Rawl	white	0.4104
United States House 1	Robert Burton	white	0.4430
	Ben Frasier	black	0.5570
United States House 3	Brian Doyle	black	0.3483
	Jane Dyer	white	0.6517
United States House 6	Gregory Brown	black	0.0993
	James Clyburn	black	0.9007

Note: races are listed in order that they appeared on ballots, and candidates are listed in the order that they appeared in each contest. Winning candidate names are in bold.

The third contest in the Democratic Primary was the United States Senate contest in which Greene was listed first, followed by Rawl. Neither candidate spent much money on this race:

Greene claims to have spent less than $2,000 although, as noted earlier, no one has been able to confirm any serious campaign activity. For the primary, the Rawl campaign claims to have made 220,000 robocalls as well as sending out emails and making public appearances throughout South Carolina.[20] This limited outreach appears to have made almost no impact on the South Carolina electorate. A poll by Public Policy Polling in late May, 2010, showed that only 5% of South Carolinians had a favorable impression of Rawl, 14% had an unfavorable impression, and 82% had no opinion.[21] Greene was so far off the political radar that, to the best of our knowledge, no one asked a single polling question about him leading up to the primary.

After the United States Senate contest, some Democratic voters in South Carolina had the chance to vote in a United States House Primary. The House contests on which we focus (rationale appears below) took place in Districts 1, 3, and 6, and we list these contests in Table 1.

Data and Methods

We analyze the contests in Table 1 with ballot-level images, precinct-level election returns, and the South Carolina statewide voter file. We now describe these data sources with particular attention to ballot images, a data type that remains relatively uncommon in American election research.

Ballot Images

Our first data source consists of ballot images, electronic copies of actual ballots cast, here ballots cast by Democratic Primary in-person voters, all of whom used iVotronic electronic voting machines.[22] Each image lists the set of choices made by a single voter, and ballot images are collected in voting machine log files. While individual ballot images cannot be matched back to the individuals who cast them, the images nonetheless provide valuable information about how voters made choices, information that cannot be gleaned from aggregated election statistics.

For example, while we know that a white candidate (e.g., Vincent Sheheen) was nominated in the Democratic primary to run for South Carolina governor and that a black candidate (Alvin Greene) was nominated for United States Senate, we cannot infer from election returns how many

[20]This conclusion is based on an email sent from Walter Ludwig, Vic Rawl's campaign manager, to Thomas Schaller of the University of Maryland, Baltimore County. The email is described at `http://www.fivethirtyeight.com/2010/06/something-fishy-in-south-carolina.html` (last accessed September 8, 2011).

[21]Poll results are available at `http://www.publicpolicypolling.com/pdf/2010/PPP_Release_SC_527.pdf` (last accessed September 8, 2011).

[22]Mail-in absentee ballots are optical scan. For details on South Carolina voting see `http://scvotes.org/south_carolina_voting_information_page` (last accessed September 8, 2011).

individual voters voted for both Sheheen and Greene.[23] Similarly, we cannot determine from election returns how many voters in the 2010 South Carolina Democratic Primary voted for any two black candidates, how many for three such candidates, how many for one white candidate and one black candidate, and so forth. The availability of ballot images makes these sort of determinations possible and, as will be clear shortly, allow us to say a great deal about the existence of racial voting. Other analyses of ballot data can be found in Lewis (2001), Wand et al. (2001), Herron and Sekhon (2003), Gerber and Lewis (2004), Herron and Lewis (2007), and Frisina et al. (2008).

There are 46 counties in South Carolina, and to gather as many ballot images from the 2010 Democratic Primary as possible we contacted every county by calling various county election offices and then following up our telephone calls and messages with emails. When necessary, this procedure was repeated at least three times per county. We mailed Freedom of Information Act requests to county election offices that demanded this, although our requests were not always successful (e.g., Beaufort County).

We successfully gathered Democratic Primary ballot images from 14 South Carolina counties.[24] Ballot counts along with overall ballots cast per county appear in Table 2, which also contains county population sizes, percent black, and median household income. Of immediate note is demographic variation across our counties. Discrepancies between ballots cast and ballots collected are discussed in the Appendix.

Given that our ballot data are from a subset of South Carolina precincts, it is important to consider whether they are adequately representative. In the Appendix we compare precincts in South Carolina for which we have ballot images to precincts for which we do not, and we conclude

[23]There are cases in which one could infer this quantity from election returns, but these cases rely on scenarios that almost never occur in practice. For example, had literally every voter who turned out in the Democratic Primary voted for Sheheen, then we would be able to infer from election returns the fraction of Sheheen voters who also voted for Greene.

[24]For most counties the ballot-level data provided to us combined both Democratic and Republican Primary voters. We needed to extract Democratic voters from the set of all primary voters, and we identified Democratic (Republican) voters by checking to see whether they had voted for any Democratic (Republican) candidates. There is a small set of 19 voters from our 14 counties who voted only in non-partisan races. Because we cannot determine whether these voters participated in the Democratic or Republican Primary, we set them aside from this point onward. There were also 79 voters from Charleston County and four voters from Dorchester County for whom we could not ascertain Congressional District. This is because these voters lived in split precincts that served multiple Congressional Districts, and the voters did not cast valid votes in a Congressional District primary. Although we know that these voters cast Congressional race primary undervotes, because the precincts are split we cannot determine the particular Congressional races in which the voters undervoted. We ignore this small group of voters in analyses that require knowing which Congressional contest in which a voter voted. There is also a group of 147 voters from Dillon County for whom we cannot determine the County Council contest in which the voter participated.

Table 2: South Carolina Counties with Ballot-Level Data and Ballots Cast in the Democratic Primary of 2010

County	Congressional District(s)	Ballots Collected	Ballots Cast	Population	Percent Black	Median Household Income
Anderson	3	3,599	3,663	187,126	16.0	41,399
Berkeley	1, 6	936	4,489	177,843	25.0	49,609
Charleston	1, 6	10,946	13,500	350,209	29.8	46,145
Dillon		4,558	6,475	32,062	46.1	28,653
Dorchester	1, 6	2,425	3,594	136,555	25.8	52,443
Greenville		10,663	10,856	451,225	18.1	45,917
Horry	1	1,568	4,379	269,291	13.4	41,163
Kershaw		4,647	5,337	61,697	24.6	45,268
Laurens	3	1,643	1,695	66,537	25.4	36,910
Newberry		1,211	1,313	37,508	31.0	37,263
Pickens	3	1,682	1,761	119,224	6.6	40,357
Spartanburg		5,043	5,148	284,307	20.6	40,278
Sumter	6	5,683	6,933	107,456	46.9	37,113
York		3,852	4,036	226,073	19.0	50,644

Note: population and income demographics are based on data from the United States Census. Population and percent black are from 2010 and median household income from 2009. Some South Carolina counties intersected multiple Congressional Districts that had Democratic primaries, and other counties had no Democratic Congressional primaries at all.

that the ballot image precincts, though not exactly representative of all precincts in South Carolina, are mostly similar to the latter although they are disproportionately missing black voters. We comment on this potential bias when presenting results.

Precinct Returns

Our second data source consists of precinct election return. These data have the feature of covering South Carolina entirely, and precinct returns foster analysis of candidate support levels as they vary with aggregated demographic characteristics.

The drawback to precinct returns is that they have the potential to lead to ecological fallacies, something that ballot-level data avoid. Nonetheless, there is a long history of election analysis using aggregated data and, to the extent that an ecological analysis produces patterns either consistent or inconsistent with ecological analyses of other elections, we can further bolster our claims about the legitimacy or the illegitimacy, respectively, of Greene's win.

Statewide Voter File

Finally, our third source of data that we bring to bear on the question of Greene's Primary win is the South Carolina statewide voter file. Simply put, the South Carolina statewide voter file is a list of the registered voters in the state, and the list indicates in which election(s) each voter has cast a ballot. A voter's entry in the statewide file never indicates for whom the voter voted in some election, just whether or not he or she turned out in said election. Among other things, the South Carolina statewide voter file allows us to determine the election participation histories of voters who participated in the 2010 Primary.

We access the South Carolina voter file through Catalist, LLC. Catalist is a firm that manages voter lists from across the United States, and data from Catalist also appear in Hersh (2011).[25]

Evidence in Favor of a Legitimate Greene Victory

We now use our three data sources to test competing theories—legitimate versus illegitimate—of Greene's victory. Our initial tests are for the presence of legitimate sources of Greene's surprising

[25]See `http://catalist.us` for details. We thank Eitan Hersh of the Department of Political Science at Yale University for providing us with data from Catalist. To assess the accuracy of Catalist's South Carolina turnout data, we checked whether Catalists's counts of total votes cast in the 2010 Democratic Primary, the 2010 Republican Primary, the 2008 Democratic Presidential Primary, and the 2008 Republican Presidential Primary match official state totals. Catalist data was close in three of four cases, off by 3%, 0.1%, 0.9%, and 21%, respectively, in total votes cast for the four aforementioned elections.

win. As explained previously, these source are race-based voting and ballot-order effects.

Ballot Patterns among Statewide Contests

We start by classifying every ballot from our 14 counties as either a Greene ballot, a Rawl ballot, or a Senate Primary undervote. We then classify each ballot based on votes in the two statewide contests that preceded the Senate contest (Table 1 lists the top races on the Democratic Primary ballot). There are $4 \times 3 = 12$ possible vote combinations from these two contests because the gubernatorial contest included three candidates plus the opportunity to undervote and the education contest included two candidates plus the opportunity to undervote.

The most common of the 12 gubernatorial / education combinations is Sheheen / Holleman, the combination that includes a vote for gubernatorial winner Sheheen and a vote for Superintendent of Education winner Holleman. For this combination as well as for the 11 other combinations we calculate the fraction of ballots with votes for Greene, the fraction with votes for Rawl, and the fraction with Senate Primary undervotes. These fractions and other details are listed in Table 3.

Table 3: Greene, Rawl, and Senate Undervote Rates across Gubernatorial and Education Vote Patterns

Governor Candidate	Education Candidate	Number Black Candidates	Greene Fraction	Rawl Fraction	Undervote Fraction	Overall Count
Ford	Holleman	1	0.72	0.24	0.03	5,274
Ford	Thompson	2	0.69	0.28	0.03	4,633
Rex	Thompson	1	0.58	0.38	0.03	4,999
Sheheen	Thompson	1	0.57	0.40	0.02	12,449
Rex	Holleman	0	0.53	0.42	0.05	7,187
Sheheen	Holleman	0	0.51	0.44	0.05	17,633
Undervote	Thompson	1	0.32	0.24	0.44	103
Undervote	Holleman	0	0.22	0.11	0.68	133
Ford	Undervote	1	0.12	0.07	0.80	1,137
Sheheen	Undervote	1	0.08	0.17	0.75	2,840
Rex	Undervote	0	0.08	0.19	0.73	1,147
Undervote	Undervote	–	0.05	0.05	0.90	921

Note: combinations are listed in decreasing order by fraction of vote received by Greene.

The results in Table 3 have several implications. First, as evidenced by overall ballot counts ("Overall Count"), the most common gubernatorial and education vote combinations involved no undervotes in either of these two contests.

Second, the Ford / Holleman vote combination is the combination in which a voter selected the first candidate listed in the gubernatorial contest and then the first candidate who was running for Superintendent of Education. This combination appears on 5,274 of the 58,456 ballots described in Table 3, and one can see from the table that almost three out of four Ford / Holleman voters also cast a vote for Greene. The Ford / Holleman combination is listed first in the table, and hence this combination led to the highest conditional probability of voting for Greene among all 12 vote combinations that were possible in the first two electoral contests in the Democratic Primary.

The Greene share of the Ford / Holleman combination is consistent with a first-candidate, ballot order effect. Given available voting data, we cannot assert that there were not, say, 5,274 voters among our 14 counties in South Carolina who deliberately chose to vote for Ford, Holleman, and Greene on the basis of their policy positions or other characteristics. And, because all of the precincts in South Carolina in June, 2010, employed the same name ordering among candidates, we do not observe cases where Rawl's name is listed before Greene's. Nonetheless, the top location of the Ford / Holleman row in Table 3 is consistent with Greene being the beneficiary of the relatively small set of voters who, one might conjecture, voted for the top-listed candidate in the first three races in the Democratic Primary simply because these candidates were listed first.

Third, the most pro-Greene ballot combination after Ford / Holleman was Ford / Thompson. Both Ford and Thompson are black, and this implies that Greene tended to receive more support from voters who, before the Senate contest, voted for two black candidates. Did Ford / Thompson voters vote for Greene *because* all of these candidates were black? That question we cannot answer with our voting data, but the fact that Ford / Thompson ballots were disproportionately pro-Greene is some initial evidence that race played an important role in Greene's victory.

Fourth, combinations that involved undervotes in the gubernatorial and education contests were associated with relatively high Senate undervote rates as well. One can see this in the low positions of the Senate undervote rows in Table 3. Similar to Herron and Sekhon (2003), we find that, among top races at least, undervotes were clustered by ballot as opposed to being spread across many ballots.

In summary, Table 3, which places Senate Primary votes in the context of votes cast in the gubernatorial and Superintendent of Education races, shows that the highest conditional probability of a vote for Greene came from voters who voted for one or two black candidates or from voters

who voted for both candidates listed first on the Democratic Primary ballot. This suggests that Greene benefited from race-based and from first-candidate voting.

Even so, these facts do not account for the margin of Greene's victory, as the number of voters represented by the top two rows of Table 3 are relatively small. Greene gains 1,318 votes over Rawl from Sheheen / Holleman voters, the most common pattern, 2,086 votes from Sheheen / Thompson voters, the second most common pattern, and 793 votes from Rex / Holleman voters, the third most common pattern. Even if Greene and Rawl split the remaining 21,187 first-two-office-vote-pattern voters evenly, i.e. Greene gains no advantage from any first-candidate-only ballots or any black-black candidate ballots, Greene would still have won by 4,197 votes in our ballot images, a margin of 7% of ballots cast. Thus Greene would have won the Democratic Senate Primary even in the absence of first-candidate bias or voters who voted exclusively for black candidates in the top two races, at least given the patterns we observe in our ballot images.

Down-ballot Contests and Ballot Order Effects

We now consider all Democratic contests in the primary as opposed to only statewide contests, and we focus initially on ballot order. In particular, for each of our ballot images we calculate the fraction of votes for first-listed candidates, ignoring the Greene vs. Rawl contest. The denominator of such a fraction varies by county as, for example, Greenville County had only three Democratic contests on its Democratic Primary ballot (gubernatorial, education, and the Senate contest) while Berkeley County had four (the same three plus a Congressional race). Fraction votes cast for first-listed candidate varies from zero to one, and for each observed fraction we calculate and display in Table 4 the proportion Greene ballots, the proportion Rawl ballots, and the proportion Senate undervote.

The first row in Table 4 is the most pro-Greene of all the rows in the table, and this particular row shows that voters who voted for first-listed candidates on non-Senate contests also voted disproportionately for Greene. While this could reflect policy preferences, insofar as name order on the primary ballot was alphabetical by last name it would be rather coincidental if all of a voter's most preferred candidates had names with letters that appeared relatively early in the alphabet.

It seems more likely that the 2,264 voters who voted for all first-listed candidates did so at least in part because these candidates were listed first in their respective contests. Even so, the 2,264

Table 4: Greene and Rawl Support by Rates of First-Candidate Voting

Fraction First	Greene Fraction	Rawl Fraction	Undervote Fraction	Overall Count
1	0.75	0.23	0.02	2,264
0.75	0.69	0.28	0.03	1,023
0.8	0.60	0.35	0.06	245
0.17	0.56	0.19	0.25	723
0.67	0.55	0.41	0.04	5,462
0.6	0.55	0.36	0.10	911
0.5	0.54	0.39	0.08	16,361
0.83	0.53	0.37	0.10	140
0.25	0.52	0.33	0.15	4,371
0.33	0.50	0.39	0.11	10,310
0	0.46	0.34	0.20	14,685
0.4	0.45	0.36	0.19	1,181
0.2	0.43	0.35	0.23	780

Note: rows are listed in decreasing order by fraction of vote received by Greene. All fractions are rounded to two significant figures.

voters who behaved in this way is not very large compared to the set of ballots that we consider. Within the counties who contributed ballot images to those studied here, Greene won by 11,127 votes, many more than 2,264 who in top-ticket contests voted only for first-listed candidates.

With respect to all the rows in Table 4, the key pattern is that ballots with many first-listed candidates ("Fraction First") tend also to have a disproportionate number of Greene votes. While the table's rows are sorted based on support for Greene ("Greene Fraction"), in general high values of fraction first appear at the top of the table.

Thus, with respect to the claim that Greene won his election *solely because* the first letter in his name comes before the first letter in Rawl's name, our evidence is to the contrary. While we find good evidence that some of Greene's votes are consistent with a first-listed candidate effect, this effect does not appear to have been strong enough, at least in our ballot collection, to have secured Greene's upset.[26]

[26]One might be concerned that our estimates of a first-listed candidate effect are too small because voters with a tendency to support first-listed candidates might have ignored this tendency disproportionately in the gubernatorial primary on account of its relatively prominent candidates. To see if this concern is warranted, we re-recalculated the numbers in Table 4 ignoring the gubernatorial primary. When we do so, our results about first-listed candidates become weaker, not stronger. For example, the re-calculated fraction Greene support among voters who supported only first-listed candidates is 0.53 as opposed to 0.75, the corresponding figure in Table 4. In addition, when the gubernatorial primary is dropped the highest Greene support is found among ballots in which two-thirds of the votes

Down-ballot Contests and Race

Now, parallel to the first-candidate analysis that appears above, we extend our study of race beyond the top of the primary ballot, in this case to the set of United States House contests that appeared after the Greene vs. Rawl Senate contest. We investigate if the number of black candidate choices inour ballots predicts the Greene vs. Rawl choice.

We calculate for each ballot and for all the contests listed in Table 1 (except for the Senate Primary) the number of black candidate votes and subtract from it the number of white candidate votes. If the black minus white difference is zero for a particular ballot, then said ballot had an equal number of black candidate votes and white candidate votes. In contrast, ballots with positive differences voted for more black candidates than white candidates, and so forth.

Among our 14 counties, the observed values of black minus white range from negative three to positive five, and in Figure 1 we plot, by county and House district, the average black minus white difference for Greene voters versus average black minus white difference for Rawl voters. The plot contains a dashed 45-degree line to aid interpretation.

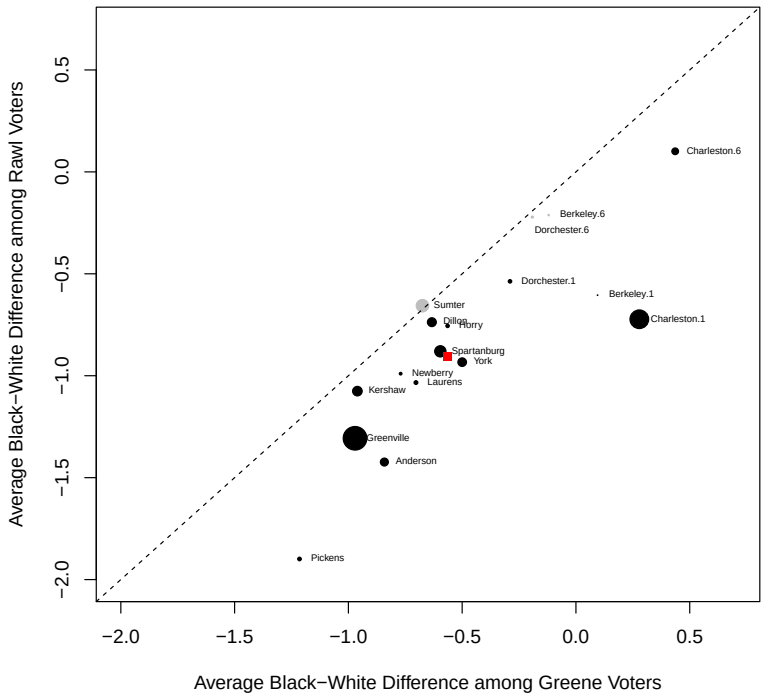
Each point in Figure 1 represents a county. For example, the point "Charleston.1," which denotes the part of Charleston County that intersects United States House of Representative District 1, has for Greene voters an average black minus white difference of approximately 0.3 and for Rawl voters, approximately -0.7. Because 0.3 is greater than -0.7, in Charleston.1 ballots with Greene votes had disproportionately more black candidate votes than ballots with Rawl votes. The colors of Figure 1's black and grey dots, respectively, denotes whether a difference-in-difference (Greene black minus white difference minus Rawl black minus white difference) is significant at the 0.05 level, and the red square is the average for all ballots described in the figure.[27]

All the points in Figure 1 except that associated with Sumter County fall to the right of the 45-degree line in the figure. This means that, for all but one of the counties in our sample, Greene voters voted on average for more black candidates than did Rawl voters. Moreover, all of the statistically significant Greene minus Rawl differences are to the right of the 45-degree line; thus, when the Greene black minus white candidate support rate is statistically different from the corresponding

cast were for first-listed candidates. As such, our inclusion of the gubernatorial contest in Table 4 does not appear to produce a downward bias in our estimates of ballot order effects.

[27]The p-values for each difference-in-difference are based on standard difference-in-means tests.

Figure 1: Black minus White Candidate Support among Greene and Rawl Voters by County



Note: describes the distribution of average black candidate minus white candidate support rates by county as well as the average (solid red square) for all Greene and Rawl voters among the 14 counties. County dots are proportional to number of Democratic Primary ballots, and the numbers after some county names indicate county voters within a particular Congressional District. Dots that are black connote counties (or part of counties) in which the difference between the Greene black candidate minus white candidate support rate and the Rawl rate is significantly different from zero at the 0.05 level. Grey dots are associated with statistically insignificant differences.

Rawl rate, it is always greater. From this we know that, holding county constant, Greene voters chose more black candidates than Rawl voters.

We hold county constant in Figure 1 to ensure that our earlier results on race were not the product of a single county. Given the plethora of points in the figure to the right of the 45-degree line, we know that race affected the Greene vs. Rawl Primary in general and not just in a handful of counties.

Of course, given that Greene was such a relative unknown prior to the primary, one might legitimately wonder whether South Carolina Democrats were aware of Greene's race. Obviously we cannot know which voters were conscious of the fact that Greene and Rawl were not of the same

race, but earlier we noted that South Carolina state Senator Robert Ford had asserted that the name "Greene" is more black compared to the name "Green." According to data from the 2000 Census, the odds of being black given a last name of "Greene" compared to being black given a last name of "Green" are 0.603, i.e., less than one, meaning that, contrary to Ford, "Green" is more black than "Greene" nationwide.[28] Of course this comparison is not particularly relevant because the important question is, is "Greene" more likely to be black than "Rawl?" The answer is yes: the odds are 1.739: an individual with a surname "Greene" is 1.7 times more likely to be black than an individual with the surname "Rawl." As such, if voters made inferences about Greene's and Rawl's races based solely on their surnames, they would on average draw correct inferences.

Although Greene was unknown, Rawl was somewhat less so. Might this asymmetry be reflected in race-based voting? Indeed the answer here is yes, and one can observe an interaction between race and Rawl support in Table 5.

For each county or subset of a county that intersects a Congressional District, this table displays average black-minus white differences for Greene voters, Rawl voters, and Senate undervoters. The black minus white differences in Figure 1 are based on the Greene and Rawl numbers in Table 5, and the table contains average black minus white differences for Senate undervoters as well.

The rightmost column of Table 5 is the difference between the undervote black minus white difference and the Rawl black minus white difference, and note that these differences are always positive. This means that Senate Primary undervoters were more likely to select black candidates over white candidates compared to Rawl voters.

Why might the result that undervoters were relatively pro-black compared to Rawl voters be notable, and what might it say about Rawl and race? To answer this question, suppose that of voters who participated in the 2010 South Carolina Democratic Primary there were three general types: pro-white, pro-black, and racially neutral. Of course there are ranges of pro-white voters, i.e., some who will only vote for white candidates and some who weakly prefer white candidates, but for the purpose of this discussion it is best to think simply of pro-white, pro-black, and racially neutral voters.

When a pro-white voter considers the candidates for the Senate contest, she has a reasonable

[28]The relevant census name database covers surnames that appeared at least 100 times in the 2000 Census. It is available at `http://www.census.gov/genealogy/www/data/2000surnames/names.zip` (last accessed August 26, 2011).

Table 5: Black and White Candidate Support Rates for all Voter Types

County	Overall Count	Black - White Greene	Black - White Rawl	Black - White Undervote	Under - Rawl Difference
Greenville	9585	-0.97	-1.31	-1.21	0.09
Charleston.1	7577	0.28	-0.72	-0.00	0.72
Sumter	5053	-0.67	-0.66	-0.53	0.12
Spartanburg	4583	-0.60	-0.88	-0.67	0.21
Kershaw	3883	-0.96	-1.08	-0.91	0.16
Dillon	3610	-0.63	-0.74	-0.28	0.46
York	3484	-0.50	-0.93	-0.75	0.19
Anderson	3236	-0.84	-1.42	-1.21	0.22
Charleston.6	2720	0.44	0.10	0.75	0.65
Laurens	1538	-0.70	-1.03	-0.96	0.07
Pickens	1481	-1.21	-1.90	-1.70	0.20
Dorchester.1	1418	-0.29	-0.54	-0.52	0.02
Horry	1404	-0.56	-0.76	-0.74	0.01
Newberry	1094	-0.77	-0.99	-0.74	0.25
Dorchester.6	803	-0.23	-0.29	-0.13	0.16
Berkeley.6	557	-0.12	-0.21	0.14	0.35
Berkeley.1	260	0.09	-0.60	-0.10	0.50

Note: rows are listed in decreasing order by number of voters. County refers to either a complete county or the part of a county that intersects a given United States House of Representative District.

idea that Rawl is white due to campaign exposure and Rawl's history in South Carolina. So, by this logic, such an individual is unlikely to cast an undervote in the Senate contest since she can be assured of voting for a white candidate if she votes for Rawl.

In contrast, a pro-black voter in the primary will also know that Rawl is likely white but will not be as confident about Greene's race because of Greene's relative obscurity. Such an individual cannot guarantee that, by voting for Greene, she is voting for a black candidate. If undervotes reflect uncertain voters or voters who choose not to participate in a contest because they are not sure what is going on, then we would expect undervoters in the Senate race to be relatively pro-black.

This is what is evident in Table 5. Undervoters, it is clear, are not on average pro-white. This is consistent with the idea that Democratic Primary voters with a strong preference for voting white voted for Rawl and that voters with a strong preference for voting black tended to participate less frequently in the Senate Primary because they did not know if there was a black candidate in the

Greene vs. Rawl contest.[29]

Precinct-Level Analysis of Race-based Voting

We turn now to a precinct-level analysis of voting in the 2010 South Carolina Democratic Primary. While our precincts cover virtually the entire state of South Carolina, unlike our ballot images that are drawn from 14 counties, an analysis based on precincts has the drawback of relying on ecological correlations. Our ballot image-precinct tradeoff is one in which data coverage is balanced against certainty of inference.

In Table 6, we present hierarchical regression results for each of the three statewide contests in the Democratic Primary plus two models that predict the additional votes per precinct obtained by Greene over Ford (black gubernatorial candidate) and Thompson (black education candidate). Our models are hierarchical as precincts are nested within counties, they include both precinct-level and county-level predictors, and they allow for randomly-varying county intercepts. All primary voters had the chance to participate in the three statewide contests analyzed in the left-most contests in Table 6, and we analyze all such contests as opposed to the Greene vs. Rawl contest alone because, if this latter contest were problematic in some sense, regression results for it would look different than regression results from the gubernatorial and education primaries.[30]

What is notable about the first three columns in Table 6 is the similarities between the Greene contest and the other, uncontroversial statewide races from the South Carolina Democratic Primary. For example, for Ford, Thompson, and Greene, the greater a precinct's fraction of non-white voters, the greater the candidate's vote share.[31] This race effect is strongest for Greene followed by Ford, but the three non-white estimates in Table 6 are positive and significant for all black candidates.[32]

[29]The Appendix shows that the coverage of our ballot images implies that they fail to include as many black voters as they should. If the pro-black voters supporting Greene were themselves disproportionately black, which is good conjecture, then the sample bias in our ballot image data makes our results conservative: by not including enough black ballots, we are disproportionately missing the most consistently pro-black voters, and hence our findings about racially-polarized ballots are understated.

[30]Source for county level data (percent with at least bachelor's degrees) is the 2006-2010 American Community Survey, and source for precinct level data (including vote, percent voters non-white and percent turnout) is the South Carolina State Election Commission. See `www.scvotes.org` (last accessed February 1, 2012).

[31]Although South Carolina only tracks white versus non-white voters across individual precincts, registration data from the state shows that almost all non-whites are black. For example, as of January, 2012, there were 2,730,651 registered voters in South Carolina of whom 765,227 were black and 1,901,025, white. The sum of white and black voters (2,666,252) is approximately 97.6% of all registered voters. See `http://www.scvotes.org/files/VR\%20Totals\%20(Race-County-Gender-Age)\%202012-01-15.xls` (last accessed February 11, 2012).

[32]The Greene point estimate in Table 6 is significantly greater than the Ford and the Thompson estimates. For Ford, the key calculation is $\frac{0.31-0.22}{\sqrt{0.02^2+0.03^2}} \approx 2.50$ and for Thompson, $\frac{0.31-0.13}{\sqrt{0.02^2+0.01^2}} \approx 8.05$.

Table 6: Hierarchical Regression Results for Statewide Contests in the 2010 Democratic Primary

	Ford	Thompson	Greene	Greene-Ford	Greene-Thompson
			Precinct-level Variables		
Percent non-white voters	0.22	0.13	0.31	0.10	0.19
	(0.03)	(0.01)	(0.02)	(0.03)	(0.02)
Percent turnout	-0.03	-0.14	-0.26	-0.30	-0.15
	(0.02)	(0.04)	(0.06)	(0.06)	(0.04)
Intercept	-35.72	-14.33	-8.56	9.36	0.17
	(0.97)	(2.16)	(3.02)	(3.06)	(1.98)
			County-level Variables		
Percent at least Bachelors	-0.04	-0.07	-0.13	-0.58	-0.03
	(0.01)	(0.02)	(0.03)	(0.03)	(0.02)
			Variance		
Precinct	43.40	68.80	83.24	95.48	155.28
County	8.55	9.48	10.08	14.23	27.21

Note: regressions are based on 2,115 precincts and are weighted by turnout. Four South Carolina precincts were not included in the regressions because they had zero turnout, and a fifth precinct was not included because it reported more votes cast in the Senate Primary than Democratic ballots cast. Percent turnout is calculated as the number of individuals who voted in the Democratic Primary over the number of registered voters, and all variables are measured in percents minus 50. Robust standard errors appear in parentheses.

The results in Table 6 are consistent with the proposition that non-white voters in the Democratic Primary tended to vote disproportionately for non-white candidates.

According to Table 6, precincts with high levels of turnout had significantly reduced vote shares for Thompson and Greene, suggesting that more participatory electorates were less likely to vote for the black candidate in each of these races. The effect is in the same direction but not statistically significant for Ford, ostensibly the most legitimate of these candidates.[33] Similarly at the county-level, a more educated populace is significantly associated with reduced vote share for all three candidates. This effect is strongest for Greene and weakest for Ford, as we might expect given the relatively greater legitimacy of Ford's candidacy.[34] In general, these Ford, Thompson, and Greene contests show a similar dynamic. Insofar as no one disputed the outcomes of the former two contests, the similarity across all three contests in Table 6 implies that there is not an especially

[33]The Greene point estimate for turnout is significantly different than the Ford estimate ($t \approx -3.63$) but statistically indistinguishable from the Thompson estimate ($t \approx -1.66$).

[34]The Greene county-level turnout estimate is significantly different than the associated Ford estimate ($t \approx -2.85$) but not different than the Thompson estimate ($t \approx -1.66$).

compelling reason to believe Greene's win was unusual.

The final two columns in Table 6 use our set of covariates to predict the additional vote share per precinct obtained by Greene over Ford and over Thompson. The results for these two models are very similar. We find, for example, that Greene benefitted relative to Ford and Thompson by increased non-white voters per precinct, suggesting that race was a major factor in explaining Greene's margin above and beyond the vote shares received by other black candidates. And, higher turnout per precinct disadvantaged Greene relative to the other black candidates; this is consistent with the logic, explicated above, of the results in the left hand side of Table 6.

Generally speaking, the results in the right most two columns of Table 6 do not look unusually problematic, and they are in line with the idea that the factor that aided Greene above all else was race. From the regressions in Table 6 it appears that participants in the 2010 South Carolina Democratic Primary voted as if they had knowledge of Greene's racial identity.

Beyond the hierarchical regression models, which point to race as a key variable, we also estimated candidate vote shares by voter race using precinct observations and King's (1997) ecological inference method. With this method we find strong evidence of race-based voting for Greene, estimating that Greene received 59.7% of the votes of non-white voters in the primary compared to 37% of the votes of white voters.[35] The estimates suggest that Greene did almost 23 percentage points better with non-white primary voters than white primary voters. We also estimate that non-white voters preferred the black candidates in the other two statewide races, with our ecological inference estimates suggesting that Ford did 18.6 points better among non-white voters (24.1% versus 5.5%, respectively) and Thompson did 10.8 points better (42.2% versus 31.4%, respectively). As such, our ecological inference results suggest that non-white voters preferred black candidates and that Greene did especially well because of a large advantage among these voters.

(Lack of) Evidence of an Illegitimate Greene Victory

We have provided evidence in support of two legitimate sources, race and ballot order effects, for Greene's primary victory. The case for an illegitimate victory must thus be strong, and we turn now to assessing the two previously described claims, voting technology problems and strategic Republicans, that would imply that Green's victory was illegitimate.

[35] Full ecological inference results are available from the authors.

Voting Machine Problems

How can we infer if there were voting machine problems in the Senate Primary? In the absence of concrete evidence that such problems were present—and to the best of our knowledge there is no such evidence—we argue that an appropriate approach is to examine patterns in Senate Primary residual vote rates.

The residual vote rate in a given electoral contest is the fraction of individuals who turned out yet did not cast a valid vote in said contest. Residual votes consist of undervotes (abstentions and otherwise missing votes) and overvotes (votes rendered invalid by the presence of more than the permitted number of choices). The iVotronic voting machines used by all in-person South Carolina voters in June, 2010, do not allow for overvotes, and thus to a very close approximation in the 2010 Democratic Primary a precinct's residual vote rate is equal to its undervote rate.[36]

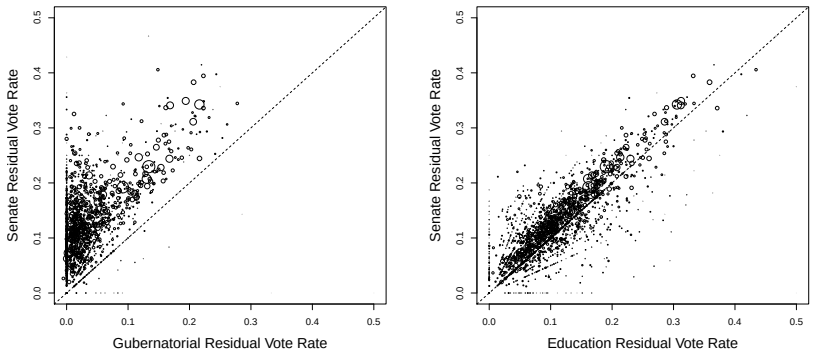
The literature on residual votes is extensive, and two regularities stand out. First, residual vote rates tend to be relatively low in contests featured at the tops of ballots. And, residual vote rates for minority voters tend to be higher than residual vote rates for white voters. See, for example, Vanderleeuw and Engstrom (1987), Ansolabehere (2002), Knack and Kropf (2003), and Tomz and van Houweling (2003).

Insofar as unusual patterns in residual vote rates can point to problems with voting machines and their interfaces, if there were such problems in the Greene vs. Rawl Senate Primary we might expect to observe residual vote rates in this primary that are at odds with residual vote rates observed elsewhere on the Democratic Primary ballot. The appropriate comparison here for the Senate Primary is other races in the overall South Carolina Democratic Primary. We do not know, nor does the literature make clear, what the "correct" residual vote rate is for something like a Democratic Senate Primary, contested by two candidates, in a southern state. We can therefore check whether residual vote rates in the 2010 South Carolina Senate Primary look anomalous compared to residual vote rates in the gubernatorial Primary, but we cannot assess whether Senate Primary residual vote rates are on average "large" or "small" in a general sense.

To this end, Figure 2 contains two plots.[37] The plots compare residual rates at the precinct

[36]While absentee voters in South Carolina could have cast overvotes, we have no data on overvote rates. Slightly under six percent of votes cast in the 2010 Primary were absentee. See `http://www.scvotes.org/2010/05/13/fact_sheets` (last accessed October 20, 2011).

[37]These two plots drop the Adamsville precinct in Marlboro County as this precinct, according to the South



(a) Governor (b) Education

Note: each panel plots residual rates by precinct. Dot sizes are proportional to Democratic ballots cast by precinct, and 45-degree lines are superimposed for clarity.

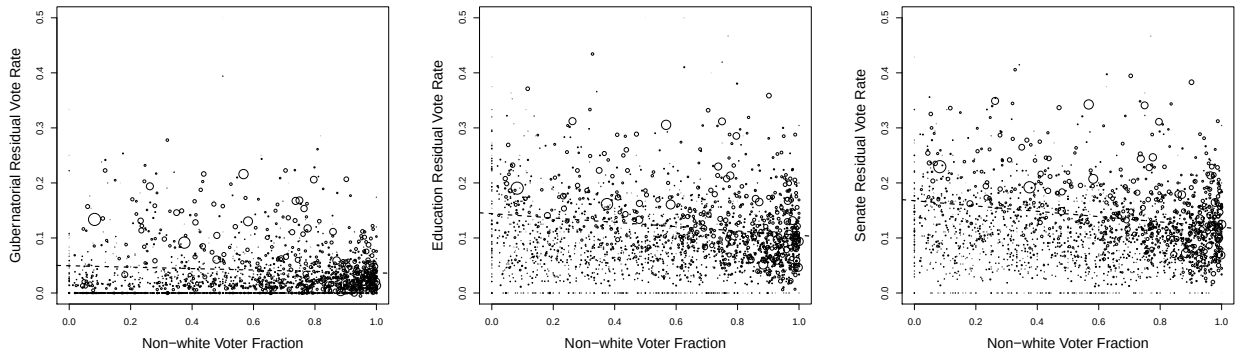
Figure 2: Residual Vote rates in Statewide Contests in the Democratic Primary

level for the Governor Primary and the Senate Primary—Figure 2a—and for the Superintendent of Education Primary and the Senate Primary—Figure 2b. Each dot in each plot denotes a precinct, and dot size is proportional to Democratic ballots cast. Several minuscule precincts in South Carolina are not visible in Figure 2 because they had very high residual vote rates that, if pictured, would distort the scales of the two residual vote rate plots.

Figure 2a shows that residual vote rates in the Senate Primary were almost always greater than residual vote rates in the gubernatorial contest. Note that the cloud of points in Figure 2a is in general above the 45-degree line. This sort of pattern is typical for American elections: the gubernatorial contest was first on the Democratic Primary ballot, and it not uncommon for initial contests to have relatively low residual vote rates.

In Figure 2b we see a similar yet more muted pattern. Residual vote rates in the Senate Primary were slightly elevated compared to residual vote rates in the Superintendent of Education contest, and more importantly the cloud of points in Figure 2b is very close to the 45-degree line. If there were something very wrong with South Carolina's voting machines in way that caused strange non-vote patterns in the Senate Primary, the cloud of points in Figure 2b would not follow such a seemingly linear pattern.

Carolina Election Commission, had more votes cast in the Senate Primary than Democratic ballots cast. This is obviously impossible, and we suspect that this Adamsville problem reflects a data entry error.



(a) Governor (b) Education (c) Senate

Note: each panel plots residual rates by precinct. Dot sizes are proportional to Democratic ballots cast by precinct, and regression lines (weighted by Democratic ballots cast) are superimposed for clarity.

Figure 3: Residual Vote rates and Voter Race in Statewide Contests in the Democratic Primary

Thus, Figure 2 is consistent with existing empirical evidence on residual votes. There is little evidence in either of the panels in the figure that suggests problematic voting machines and, by association, an illegitimate victory for Greene.

Continuing our exploration of residual vote rates, Figure 3 contains three plots, each of which describes the relationship between voter race (x-axis) and a residual vote rate (y-axis) in the gubernatorial, education, and Senate primaries. Each plot also contains a dashed regression line.

Is there any evidence of illegitimacy in Figure 3? One could note that the negative sloping regression lines seem odd insofar as, in most studied electoral environments, precincts that have more minority voters have relatively high, not low, residual vote rates, *ceteris paribus*. Thinking conspiratorially, one could argue from Figure 3c that white voters in the Senate Primary were somehow pushed to cast residual votes.

The problem with this logic, outside of the possibility that it is an ecological fallacy, is that we observe negative regression slopes in all three panels of Figure 3. Thus, if South Carolina's voting machines were somehow racially discriminatory in the Senate Primary, then they must have operated similarly in both the Governor Primary and the Superintendent of Education Primary. To the best of our knowledge there have been no assertions that the latter two contests were affected by voting machine problems. Thus, while a marker of potentially problematic voting machinery

that affected the Senate Primary would be a regression slope in Figure 3c that was different than the slopes in Figure 3b and 3c, no such marker exists.

The similarity across electoral contests displayed in Figure 3 in conjunction with the regularities observed in Figure 2 together make the allegation of voting machine problems in the Greene vs. Rawl Senate Primary hard to sustain. Among all statewide contests in the Democratic Primary, such allegations were made *only* about this primary and yet a standard indicator of voting machine problems fails to highlight it in a meaningful way. While such an argument cannot dismiss with certainty that there were voting machinery problems in the Senate Primary, it certainly lowers the likelihood that such problems existed.

Strategic Republicans

Finally, we consider the possibility that strategic Republicans participated in the South Carolina Democratic Primary in order to torpedo the candidacy of Rawl and nominate a weak challenger, Alvin Greene, who would then face incumbent Senator Jim DeMint. We consider a number of ways that strategic Republicans might have acted in support of this objective.

Ideally we would like to identify potential Republican voters in South Carolina and then check whether they participated in the 2010 Democratic Primary. South Carolina, however, is not a party registration state, and thus we cannot simply check to see how many registered South Carolina Republicans voted in the 2010 Democratic Primary. Nonetheless, for states like South Carolina without party registration, one can use primary vote history to infer the party preference of individual voters on the assumption that individuals are more likely to vote in the primary of their preferred party. Toward this end we consulted the South Carolina voter file and tabulated 2010 Primary turnout by 2008 Primary turnout.

The 2008 Primary in South Carolina was an important primary in the 2008 presidential nominating process, and it featured competitive Democratic and Republican contests; the presidential primaries were eventually won by Barack Obama and John McCain. We assume that the 2008 Primary choice of a voter, Democrat or Republican, is a fair measure of his or her partisanship. Then, if it were the case that a large number of strategic Republicans participated in the 2010 Democratic primary, we should observe a large number of registrants who had voted in the 2008 Republican Primary voting in the 2010 Democratic Primary.

While the statewide South Carolina voter file identifies 17,811 registrants who voted in the 2010 Democratic Primary after voting in the 2008 Republican Primary, we also find 40,899 registrants who voted in the 2008 Democratic Primary and yet voted in the 2010 *Republican* Primary. Specifically, about 11.4% of the 2010 Democratic Primary electorate were "Republicans" who voted in the Republican Primary in 2008, and about 11.6% of the 2010 Republican Primary electorate were "Democrats" who voted in the Democratic Primary in 2008.

These two numbers—11.4% and 11.6%—do not provide much support for a hypothesis that large numbers of Republican voters in 2010 undermined the candidacy of Vic Rawl by voting for Greene. If one were to assert that this happened, one would also have to motivate why so many apparent Democrats crossed over to vote in the Republican Primary in 2010.[38]

As another check on Republican crossover voting, we consider the fraction of Democratic ballots in the 2010 South Carolina Primary compared to the fraction of Democratic ballots in the 2006 South Carolina Primary. The former should be relatively larger if Republican crossover voting occurred to a significant extent. We use 2006 as a comparison here because there was a presidential primary in neither 2006 nor 2010. In 2010, when crossover voting could be alleged to have influenced Greene's victory, this fraction was 0.32; in other words, about 32 out of 100 primary ballots cast in 2010 were cast in the Democratic Primary. In 2006, when to the best of our knowledge there were no serious allegations of Republican crossover voting, Democratic fraction ballots was 0.36. Because the 2006 fraction is greater than the 2010 fraction, the two fractions are inconsistent with the notion that there was a surge of crossover voting in 2010.[39]

There is thus little evidence from voter registration statistics that strategic Republicans targeted Greene. Nonetheless, we now look for evidence of this phenomenon among our ballot images.

If strategic Republican voters were strongly focused on the election of Alvin Greene via the undermining of Rawl, among statewide contests in the Democratic Primary they might have chosen to participate only in the Senate Primary. Within our set of ballot images, of the 921 ballots out of 58,456 total that contained undervotes in both the gubernatorial and Superintendent of Education contests, 46 contained votes for Greene versus 49 which had votes for Rawl, essentially an even split

[38]Our results on a lack of crossover voting are consistent with those in Alvarez and Nagler (1997), who study a variety of primaries and conclude that crossover voting is not extremely common.

[39]See `http://www.enr-scvotes.org/SC/16117/28318/en/summary.html`, `http://www.scvotes.org/files/06primaryOfficial/RepublicanStatewide.pdf`, and `http://www.scvotes.org/files/06primaryOfficial/DemocratStatewide.pdf` for data (all last accessed on September 8, 2011).

(the remaining 826 contained a Senate race undervote). This constitutes little evidence of targeted Greene-supporting behavior.

On the other hand, perhaps strategic Republican voters eschewed undervoting and instead sought to support weak candidates in the Democratic primary so as to strengthen Republican general election candidates. How might we find evidence of this behavior?

Perhaps the strongest black candidate who ran in the 2010 South Carolina Democratic Primary was James Clyburn, a candidate from United States House of Representatives District 6; see Table 1. Strategic Republicans who voted against Clyburn should be expected to have voted disproportionately for Greene, and in Table 7 we tabulate votes in the Senate Primary by votes in the District 6 contest. If strategic voters sought to undermine the candidacies of strong candidates Rawl and Clyburn, we should see a correlation between Rawl-Clyburn votes and between Greene-Brown votes, Brown being Clyburn's opponent.

Table 7: Greene Support by U.S. House 6 Vote

	Senate Contest		
	Undervote	Greene	Rawl
House 6 Undervote	272	645	404
	(0.206)	(0.488)	(0.306)
House 6 Brown	75	312	214
	(0.125)	(0.519)	(0.356)
House 6 Clyburn	920	4366	2090
	(0.125)	(0.592)	(0.283)

Note: cell entries are counts of ballot images showing that two-vote combination. Numbers in parenthese are fractions aggregated across columns. Not all rows sum to one due to rounding.

Of 601 ballot images that contain votes against James Clyburn and for Gregory Brown, Table 7 shows that only slightly more than half (312) contain a vote for Greene. Moreover, District 6 undervoters were practically as supportive of Greene (support rate of 0.448) as were anti-Clyburn voters (0.519). We thus find little evidence that voters strategically chose weak candidates over strong ones in both the District 6 and Senate Primaries.

Finally, we consider the set of ballots with votes for Jim Rex (worst vote share in the gubernatorial contest) and Tom Thompson (losing education candidate). Did voters who supported these two candidates disproportionately vote for Greene in an effort to insert weak Democrats in the 2010 General Election? Of the 4,999 ballots with Rex and Thompson votes, approximately 58.3%

of them contained Greene votes. This number is a fraction smaller than Greene's overall vote share in the Democratic Primary, thuswe have providing little evidence of strategic Republican behavior.

Conclusion

We began by drawing attention to a puzzle: how did Alvin Greene, an unknown candidate with no political background, defeat Vic Rawl, a much better-known candidate with a wealth of experience, in the 2010 South Carolina United States Senate Primary? The most compelling part of our answer to this question is race. Race is not infrequently a key feature of American elections, and our evidence shows that the Greene vs. Rawl primary does not look particularly different from what one might expect from a contest that featured two candidates, one white and one black. Beyond race, we also showed that Greene benefitted slightly from being listed above Rawl on the Primary ballot although this ballot order effect was not pivotal to Greene's victory.

Our finding that race appears to have been used as a voting heuristic is in line with other research on vote choice in low-information settings where voters use whatever cues and shortcuts they can to make choices that approximate what they would do in the presence of more information (Berelson, Lazarsfeld and McPhee 1954; Lupia 1994; McDermott 1997, 2005; Sigelman et al. 1995). Regardless of whether the use of race as an heuristic is effective, the presence of this heuristic in the Greene vs. Rawl contest is consistent with the operation of a legitimate election.

We found no evidence of voting technology problems in the Senate Primary, nor did we uncover evidence that Republicans in South Carolina strategically participated in the Democratic Primary with the intent of swinging the Senate contest to Greene. Our conclusion, then, is that Alvin Greene was the legitimate winner of the 2010 Democratic Primary in South Carolina.

Our findings are independent of the matter as to whether Republicans in South Carolina encouraged or even facilitated Greene's primary run. Similarly, the findings do not shed light on why Greene ran and why he appears to have thought that spending $10,400 to run for a Senate nomination was a useful thing to do.

Greene was practically unheard of, yet voters acted as if they knew his race. While election scholars commonly assume that regular campaign activities like advertising, get-out-the-vote efforts, and media coverage are necessary for electorates to learn about candidates running for office, it may

be possible that formal advertising and thorough news coverage are not necessary conditions for the racial characteristics of candidates to filter through electorates. Might social networks, informal conversation, and intermittent local news coverage be enough to transmit this information? The extent of correlated voting for Greene and other black candidates suggests that, somehow, many Democratic voters knew to vote for Greene on the basis of race despite Greene's lack of a campaign. Future research may want to investigate how information on Greene's race may have propagated, whether in general such propagation is limited to salient candidate features such as race, or if other less-salient candidate characteristics might also filter through the electorate in informal ways.

Lastly, our analysis demonstrates the value of ballot records in an exercise designed to assess election legitimacy. Copies of ballots cast allow citizens, politicians, and academics to assess the fairness of election outcomes, and our analysis is one example of how ballot images can identify reasonable individual voting patterns that generate surprising aggregate election results. Access to anonymized ballot records after an election is an important part of election transparency, and we have used this transparency to diagnose the legitimacy of a surprising election outcome, finding that Alvin Greene's 2010 victory, even if surprising, does not appear to be cause for concern about election administration in South Carolina.

Appendix

For counties where our count of ballots is much lower than the number of ballots cast (e.g., Berkeley and Horry), we are likely missing logs from some iVotronic voting machines.[40] The explanation for this is in part a reflection of the fact that ballot logs are not automatically produced by such machines. Logs can be generated post-election by administrators, but the process required to produce logs goes beyond the process required to produce candidate vote totals.[41] In counties

[40] We confirmed this likelihood through phone calls on October 20, 2011, with election administrators from Horry and Dorchester Counties and on October 27, 2011, with administrators from Charleston County. South Carolina's voting machines were programmed to allow voting in either party primary due to the open primary status of the state. Thus, if only a few voters voted in the Democratic Primary on a voting machine whose data were not downloaded, a vote log could be off by a few votes. On the other hand, if a machine from a precinct in which many voters participated in the Democratic Primary did not have data downloaded, a vote log could be off by many votes. The Charleston County administrator, Patrick Lee, who formerly worked for the state and for ES&S, the vendor who produces the voting machines, said training for election administrators was not always as thorough as it could have been, and conjectured that much of the missingness in our data was due to human error in implementing the audit of ballot logs (telephone call of October 27).

[41] Each iVotronic voting machine has a card—essentially a USB key—that stores voting data. This card is not part of the process used to report candidate vote totals; instead, the card is used if election administrators purposefully set out to create ballot images, among other things. See, for example, `http://www.scvotinginfo.com/wp/wp-content/uploads/2011/03/iVoOperatorsManual.pdf` (last accessed October 18, 2011).

where our ballot counts are very close to countywide totals (e.g., Anderson and Greenville), the difference is likely due either to votes cast on alternative, optical scan absentee machines, or, as above, to missing voting machines used only sparingly during the 2010 Primary.

As our ballot data are from a subset of South Carolina precincts, it is important to consider whether they are adequately representative. To this end consider Table A1, which contrasts the precincts that are covered by our ballot images ("covered precincts") versus those that are not ("uncovered precincts").

Table A1: Covered versus Uncovered Precincts

	Covered Precincts	Uncovered Precincts
Count	949	1167
Fraction White Voters	0.410	0.330
Fraction Alvin Greene	0.588	0.601

Table A1 shows that, first, the set of 949 covered precincts for which we have ballot data are more white than the set of uncovered precincts.[42] This sample bias—undercoverage of non-whites—should make our results on race-based voting conservative. Second, the set of covered precincts is only slightly more pro-Greene than the uncovered precincts. That the former are more pro-Greene than the latter is consistent with our undercoverage of non-whites. In any case, the sample bias here is likely to be slight given the similar Greene shares of 0.588 and 0.601.[43]

[42] According to the 2005-2009 American Community Survey, 98.7% of South Carolina residents report being of one race; of this group 67.4% is white and 28.2%, black. Hence we treat South Carolina as a state that has white and black residents; this assumes that the South Carolina residents who are neither black nor white are evenly split between these two categories in terms of voting behaviors. See `http://factfinder.census.gov/servlet/ACSSAFFFacts?_event=&geo_id=04000US45&_geoContext=01000US\%7C04000US45&_street=&_county=&_cityTown=&_state=04000US45&_zip=&_lang=en&_sse=on&ActiveGeoDiv=geoSelect&_useEV=&pctxt=fph&pgsl=040&_submenuId=factsheet_1&ds_name=null&_ci_nbr=null&qr_name=null&reg=null\%3Anull&_keyword=&_industry=` for the South Carolina Fact Sheet that contains this racial data (last accessed October 19, 2011).

[43] Our data on voter race at the precinct level were gathered from the South Carolina Election Commission, and these data are for Democratic Primary voters. Nonetheless, race-based precinct totals (white voters plus non-white voters) are very close to, but do not exactly match, counts of Democratic ballots cast by precinct that are also provided by the South Carolina Election Commission. For the purposes of Table A1 and the regressions that use precinct-level racial statistics, we assume that fractions of white voters by precinct are correct even given the slight inconsistencies between the two aforementioned sources of data.

References

Abramowitz, Alan I. and Jeffrey A. Segal. 1992. *Senate Elections.* Ann Arbor, MI: The University of Michigan Press.

Alvarez, R. Michael and Jonathan Nagler. 1997. "Analysis of Crossover and Strategic Voting." Unpublished working paper, available at `http://www.hss.caltech.edu/SSPapers/wp1019.pdf`.

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

Banducci, Susan A., Todd Donovan and Jeffrey A. Karp. 2004. "Minority Representation, Empowerment and Participation." *Journal of Politics* 66(2):534–556.

Bartels, Larry M. 2000. "Partisanship and Voting Behavior, 1952-1996." *American Journal of Political Science* 44(1):35–50.

Berelson, Bernard R., Paul F. Lazarsfeld and William N. McPhee. 1954. *Voting: A Study of Opinion Formation in a Presidential Campaign.* Chicago: University of Chicago Press.

Bobo, Lawrence and Frank Gilliam. 1990. "Race, Sociopolitical Participation, and Black Empowerment." *American Political Science Review* 84(2):377–393.

Dawson, Michael C. 1995. *Behind the Mule: Race and Class in African-American Politics.* Princeton, NJ: Princeton University Press.

Downs, Anthony. 1957. *An Economic Theory of Democracy.* New York: HarperCollins Publishers.

Enelow, James M. and Melvin J. Hinich. 1984. *The Spatial Theory of Voting: An Introduction.* New York: Cambridge University Press.

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.

Gay, Claudine. 2001. "The Effect of Black Congressional Representation on Political Participation." *American Political Science Review* 95(3):589–602.

Gerber, Elisabeth R. and Jeffrey B. Lewis. 2004. "Beyond the Median: Voter Preferences, District Heterogeneity, and Political Representation." *Journal of Political Economy* 12(6):1364–1383.

Green, Donald P., Bradley Palmquist and Eric Schickler. 2002. *Partisan hearts and minds: political parties and the social identities of voters.* New Haven, CT: Yale University 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.

Herron, Michael C. and Jeffrey B. Lewis. 2007. "Did Ralph Nader Spoil a Gore Presidency? A Ballot-Level Study of Green and Reform Party Voters in the 2000 Presidential Election." *Quarterly Journal of Political Science* 2(3):205–226.

Herron, Michael C., Walter R. Mebane, Jr. and Jonathan N. Wand. 2008. "Voting Technology and the 2008 New Hampshire Primary." *William & Mary Bill of Rights Journal* 17(2):351–374.

Hersh, Eitan. 2011. "At the Mercy of Data: Campaigns Reliance on Available Information in Mobilizing Supporters." Unpublished Manuscript, available at `http://www.eitanhersh.com/uploads/7/9/7/5/7975685/atthemercyofdata.pdf`.

Ho, Daniel E. and Kosuke Imai. 2006. "Randomization Inference With Natural Experiments: An Analysis of Ballot Effects in the 2003 California Recall Election." *Journal of the American Statistical Association* 101(475):888–900.

Jessee, Stephen A. 2009. "Spatial Voting in the 2004 Presidential Election." *American Political Science Review* 103(1):59–81.

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

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

Koppell, Jonathan GS and Jennifer A. Steen. 2004. "The Effects of Ballot Position on Election Outcomes." *Journal of Politics* 66(1):267–281.

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

Lupia, Arthur. 1994. "Shortcuts Versus Encyclopedias: Information and Voting Behavior in California Insurance Reform Elections." *American Political Science Review* 88(1):63–76.

McDermott, Monika L. 1997. "Voting Cues in Low-Information Elections: Candidate Gender as a Social Information Variable in Contemporary United States Elections." *American Journal of Political Science* 41(1):270–283.

McDermott, Monika L. 2005. "Candidate Occupations and Voter Information Shortcuts." *Journal of Politics* 67(1):201–219.

Sigelman, Carol K., Lee Sigelman, Barbara J. Walkosz and Michael Nitz. 1995. "Black Candidates, White Voters: Understanding Racial Bias in Political Perceptions." *American Journal of Political Science* 39(1):243–265.

Tesler, Michael and David O. Sears. 2010. *Obama's Race: The 2008 Election and the Dream of a Post-racial America.* Chicago: University of Chicago Press.

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.

Vanderleeuw, James M. and Richard Engstrom. 1987. "Race, Referendums, and Roll-Off." *Journal of Politics* 49(4):1081–92.

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." *American Political Science Review* 95(4):793–810.