

Voting Technology and the 2008 New Hampshire Primary

Michael C. Herron

Walter R. Mebane

Jonathan N. Wand

Repository Citation

Michael C. Herron, Walter R. Mebane, and Jonathan N. Wand, *Voting Technology and the 2008 New Hampshire Primary*, 17 Wm. & Mary Bill of Rts. J. 351 (2008), <http://scholarship.law.wm.edu/wmborj/vol17/iss2/4>

VOTING TECHNOLOGY AND THE 2008 NEW HAMPSHIRE PRIMARY*

Michael C. Herron,** Walter R. Mebane, Jr.,*** & Jonathan N. Wand****

EXECUTIVE SUMMARY

We address concerns that the reported vote counts of candidates running in the 2008 New Hampshire presidential primaries were affected by the vote-tabulating technologies used across New Hampshire.

- In the Democratic primary, Hillary Clinton was more successful in New Hampshire wards that used Accuvote optical scan vote-tabulating technology than was Barack Obama, receiving 4.3% more of the vote there (40.2% for Clinton versus 35.9% for Obama). In contrast, Clinton did worse than Obama in wards that counted paper ballots by hand, trailing by 6.1% (33.7% versus 39.8%).
- In the Republican primary, Mitt Romney trailed John McCain by 3.6% in Accuvote wards and by 15% in wards that counted ballots by hand.
- In New Hampshire, the choice of vote-tabulating technology is made ward by ward, and electronic technology was used in wards that typically differ demographically and politically from wards that count ballots by hand. Wards that selected electronic tabulation are disproportionately from the southeast part of New Hampshire, and they tend to be more densely populated and more affluent. Accuvote and hand count wards have also typically produced divergent voting patterns in elections prior to the 2008 primary. This context makes it plausible that most or all of the observed differences between vote-tabulating technologies in the votes candidates received reflect such background differences and not anything inherent in the tabulation methods.

* The authors thank Jasjeet S. Sekhon for computing resources and advice and thank Anthony Stevens of the Office of New Hampshire Secretary of State for comments on an earlier draft of this paper and for providing data. The authors are listed alphabetically.

** Associate Professor of Government, Dartmouth College.

*** Professor of Political Science and Professor of Statistics, University of Michigan.

**** Assistant Professor of Political Science, Stanford University, and Robert Wood Johnson Health Policy Scholar, University of Michigan.

- Using a subset of New Hampshire wards that have similar demographic features and voting histories but differ in their vote-tabulating technologies, we find no significant relationship between a ward's use of vote-tabulating technology and the votes or vote shares received by most of the leading candidates who competed in the 2008 New Hampshire presidential primaries. Among Hillary Clinton, John Edwards, Dennis Kucinich, Barack Obama, and Bill Richardson in the Democratic primary and among Rudy Giuliani, Mike Huckabee, John McCain, Ron Paul, and Mitt Romney in the Republican primary, we observe a significant difference only in the votes counted for Edwards, and that difference is small (a deficit of between 0.6% and 3.4% in the hand-counted votes).
- With respect to Hillary Clinton's surprise victory in the Democratic primary and the differences across vote-tabulating technologies in Clinton's and others' votes, our results are consistent with these differences being due entirely to the fact that New Hampshire wards that use Accuvote optical scan machines have voters with different political preferences than wards that use hand-counted paper ballots.

INTRODUCTION

In the immediate aftermath of the 2008 New Hampshire Democratic and Republican presidential primaries there were widespread concerns that the reported vote counts of candidates running in these races were affected by the technologies used in New Hampshire to tabulate votes.¹ Probably the most frequently discussed allegation asserted that a digital method of tabulating votes benefited New York Senator Hillary Clinton at the expense of her chief competitor in the Democratic primary, Illinois Senator Barack Obama.² This allegation appears to be the motivating factor behind the recount that was pushed for by Ohio Congressman Dennis Kucinich.³ As a candidate in the Democratic primary (according to pre-recount results Kucinich received 3901 votes in the race), Kucinich was entitled under New Hampshire election law to request a recount as long as he funded it.⁴ In fact, Kucinich funded a partial Democratic recount, and a complete Republican recount was funded by backers of Texas Congressman Ron Paul.⁵

¹ See, e.g., Dan Tuohy, *Kucinich Says He'll Ask for Recount*, N.H. UNION LEADER (Manchester, N.H.), Jan. 11, 2008, at A7.

² See, e.g., Lauren R. Dorgan, *Primary Votes To Be Recounted*, CONCORD MONITOR (N.H.), Jan. 12, 2008, available at 2008 WLNR 4232046.

³ See *id.*

⁴ N.H. REV. STAT. ANN. §§ 660:2, 660:7 (2008).

⁵ Dorgan, *supra* note 2 (noting that supporters of Congressman Ron Paul financed the Republican recount using online donations); Press Release, William M. Gardener, Sec'y of State of N.H., Statewide Recount of the Republican and Democratic Presidential Primaries

Assuming that the original paper ballots are available, a recount is the only comprehensive method for evaluating the accuracy of a given election's vote tabulations.⁶ However, comprehensive audits of elections remain rare even in the post-2000 presidential election period.⁷ We conduct here a statistical examination of reported election returns to evaluate whether voting technology in New Hampshire affected vote outcomes in the state's 2008 presidential primaries.

Ordinarily, our approach to election forensics would be employed as a means of detecting election irregularities with any identified irregularity being the basis for requesting a recount, i.e., statistical analyses identify problems that necessitate auditing.⁸ In fact, while the New Hampshire presidential primary recounts were already requested at the time we posted our initial findings, they were not complete and their results had not been published at that time.⁹ The subsequent availability of the recount of the New Hampshire Republican primary provides a useful benchmark against which to evaluate the performance of the election forensic methods we demonstrate.

The allegation that there were vote-tabulating problems in the New Hampshire Democratic primary is based on differences in candidate vote shares across voting technologies.¹⁰ These technologies are Accuvote optical scan voting (Accuvote) and hand-counted paper ballots (PBHC).¹¹ No other technology is currently used in New

(Jan. 11, 2008), available at <http://www.sos.nh.gov/recount%20press%20release.pdf> (announcing that Congressman Dennis Kucinich and presidential candidate Albert Howard satisfied the requirements to initiate recounts of the Democratic and Republican primaries, respectively); Recount, <http://www.sos.nh.gov/recountresults.htm> (last visited Nov. 30, 2008) (stating that the Kucinich recount was stopped due to lack of funds on the morning of January 23, 2008).

⁶ See ANDREW W. APPEL, CTR. FOR INFO. TECH. POL'Y & DEP'T OF COMPUTER SCI., EFFECTIVE AUDIT POLICY FOR VOTER-VERIFIED PAPER BALLOTS IN NEW JERSEY (2007), <http://www.cs.princeton.edu/~appel/papers/appel-nj-audits.pdf>.

⁷ Jonathan Wand, *Evaluating Voting Technologies: 2004 New Hampshire Democratic Primary 1* (Stanford University, Working Paper, 2004), available at <http://wand.stanford.edu/elections/us/NH/nh2004primary.pdf>.

⁸ See, e.g., Henry E. Brady et al., *Law and Data: The Butterfly Ballot Episode*, 34 POL. SCI. & POL. 59, 62, (2001) (detailing the authors' decision to act as expert witnesses in a legal challenge to the butterfly ballot used in Palm Beach County, Florida, in the 2000 presidential election based on the unusually large number of votes for Pat Buchanan in Palm Beach).

⁹ We reported our results on the Internet on January 19, 2008, and the recounts were not completed until at least a week later. See *GOP Ballot Recount Begins as Democrat Recount Fizzles*, N.H. UNION LEADER (Manchester, N.H.), Feb. 13, 2008, at A6 (explaining Congressman Kucinich's termination of the Democratic recount on January 24); *No Change in GOP Recount*, N.H. UNION LEADER (Manchester, N.H.), Feb. 13, 2008, at C4 (relating the completion of the Republican recount on February 12).

¹⁰ See Dorgan, *supra* note 2 (relating the "unexplained disparities between hand-counted ballots and machine counted ballots" that prompted Congressman Kucinich to call for a recount).

¹¹ For the 2004 primary and general elections there were two different optical scan technologies used in New Hampshire: ACCUVOTE and Optech. Optech is no longer used in the

Hampshire,¹² and vote totals and percentages by tabulating technology for the top five Democratic and top five Republican candidates are listed in Table 1.¹³ As is evident in the table, Clinton was more successful in New Hampshire Accuvote wards than Obama, receiving 4.3% more of the vote there (40.2% compared to 35.9%).¹⁴ In contrast, Clinton did worse than Obama in PBHC wards, trailing by 6.1% (33.7% compared to 39.8%).¹⁵

If one uses associations like these—where one candidate does better than another in areas with a given vote-tabulating technology—as a metric for identifying vote-tabulating problems, then it follows that the Republican presidential primary in New Hampshire suffered from such problems as well. While Arizona Senator John McCain performed better than former Massachusetts Governor Mitt Romney in both Accuvote and PBHC wards, Romney trailed by 3.6% in wards with Accuvote machines and by 15% in PBHC wards.¹⁶ By this logic, Accuvote machines in the Republican primary worked to the advantage of Romney.¹⁷

state. New Hampshire's ACCUVOTE machines are manufactured by Premier Election Systems which was previously known as Diebold Election System. Wand, *supra* note 7 (discussing the use of hand counts and Accuvote in the 2004 New Hampshire primary); see New Hampshire Municipalities Which Use the ACCUVOTE Voting Machine, <http://www.sos.nh.gov/voting%20machines2006.htm> (last visited Nov. 30, 2008) (recognizing that while many New Hampshire municipalities use ACCUVOTE, many still count ballots by hand).

¹² See New Hampshire Municipalities Which Use the ACCUVOTE Voting Machine, *supra* note 11.

¹³ For sources used in compiling Table 1, see Presidential Primary, <http://www.sos.nh.gov/presprim2008/index.htm> (last visited Nov. 30, 2008) (stating that the 2008 New Hampshire Democratic and Republican presidential primaries were held on January 8, 2008 and providing links to official vote totals by county). Coincidentally, twenty-one candidates participated in the 2008 New Hampshire presidential primary. See Candidates for Upcoming Presidential Primary Election, <http://www.sos.nh.gov/presprim2008/candidatesfiled.htm> (last visited Nov. 30, 2008).

¹⁴ See RealClear Politics, http://www.realclearpolitics.com/epolls/2008/president/nh/new_hampshire_democratic_primary-194.html (last visited Nov. 30, 2008); see *infra* p. 355 tbl.1.

¹⁵ See *infra* p. 355 tbl.1. A New Hampshire ward as discussed here is either a town, e.g., Hanover, or a subdivision, e.g., Manchester. Most towns and cities in New Hampshire are not broken down by ward, but the more populous ones are. See Towns and Wards Districted for Election Purposes, <http://www.sos.nh.gov/electionsnew.html> (follow "Voting Districts" hyperlink; then follow "Towns and Wards" hyperlink) (last visited Nov. 30, 2008) (listing all towns and cities broken down into wards).

¹⁶ See RealClear Politics, *supra* note 14; *infra* p. 355 tbl.1 (subtracting Romney's percentage of the vote in Accuvote and PBHC wards from McCain's percentage of the vote in Accuvote and PBHC wards yields McCain's margin of victory for the two types of voting technology).

¹⁷ Romney, when compared to McCain, performed 11.4% better in Accuvote wards than PBHC wards. See *infra* p. 355 tbl.1.

Table 1: Top Candidate Vote Totals and Percentages (Pre-Recount)

	Total		Accuvote		PBHC	
	Votes	Percent	Votes	Percent	Votes	Percent
<i>Democratic Primary</i>						
Clinton	112,606	39.4	97,388	40.2	15,218	33.7
Obama	105,004	36.8	87,066	35.9	17,938	39.8
Edwards	48,818	17.1	40,871	16.9	7,947	17.6
Richardson	13,239	4.6	10,652	4.4	2,587	5.7
Kucinich	3,901	1.4	3,063	1.3	838	1.9
<i>Republican Primary</i>						
McCain	88,570	37.7	73,684	36.5	14,886	39.9
Romney	75,546	32.2	66,246	32.9	9,300	24.9
Huckabee	26,859	11.4	21,964	10.9	4,895	13.1
Giuliani	20,439	8.7	17,375	8.6	3,064	8.2
Paul	18,307	7.8	14,875	7.4	3,432	9.2

Notes: Candidates are listed in order of total votes, and all vote totals are pre-recount. Edwards refers to former North Carolina Senator John Edwards; Richardson to New Mexico Governor Bill Richardson; Romney to former Massachusetts Governor Mitt Romney; Huckabee to Arkansas Governor Mike Huckabee; Giuliani to former New York City Mayor Rudolph Giuliani; and Paul to Texas Congressman Ron Paul. Vote shares are rounded and do not sum to one because not all candidates are listed.

The difference between pre-election polls and the Democratic primary outcome (Clinton over Obama) has perhaps contributed to skepticism of the accuracy of vote tabulations in this contest.¹⁸ Pre-election polls were consistent in their view that Obama, who won the previously held Iowa Caucuses, would beat Clinton by

¹⁸ See *How Did Pollsters Blow It?*, SEATTLE TIMES, Jan. 10, 2008, at A1.

somewhere between 5 and 13%.¹⁹ Depending on which pre-election poll one consults, there was perhaps a ten- to twelve-point swing between Clinton's pre-election deficit to Obama compared to her observed electoral margin over the Illinois senator.²⁰ In contrast, the victory in the Republican primary of John McCain was predicted by many, if not all, opinion polls that circulated immediately prior to the election.²¹

With this set-up in mind, we investigate the claim that voting technology differentially aided Clinton, and more generally we consider whether the New Hampshire Democratic and Republican primaries were fair in the sense of having accurate vote tabulations. Simply put, we seek to understand whether a facet of election administration—ward choice of voting technology—can explain why a given candidate received either many or few votes. In an ideal world, all such administrative choices would have no effect on election outcomes or at least not systematically bias the results in favor of any particular candidate. Thus, to the extent that we can explain differences between Clinton's and Obama's vote counts (and similarly, between McCain's and Romney's vote counts) without recourse to voting technology, the more that New Hampshire voters can trust that the election outcomes in the recent 2008 presidential primaries were not artifacts produced by a particular method of tabulating votes.

Our objective is similar to prior efforts aimed at evaluating the effect of voting technology and other election administration matters on vote outcomes including studies of the 2004 New Hampshire Democratic presidential primary²² and the 2004 presidential election in New Hampshire.²³ It is also similar to analyses of elections across the United States.²⁴ The question of whether the administration of an election

¹⁹ Dan Balz, *With Echoes of Clinton '92, Another 'Comeback Kid,'* WASH. POST, Jan. 9, 2008, at A1; see RealClear Politics, *supra* note 14 (listing the most recent polling before the election, including a Suffolk/WHDH poll showing a five-point Obama lead and a Reuters C-Span/Zogby poll showing a thirteen-point Obama lead).

²⁰ See RealClear Politics, *supra* note 14, for a summary of the pre-primary polls.

²¹ See RealClear Politics, http://www.realclearpolitics.com/epolls/2008/president/nh/new_hampshire_republican_primary-193.html (last visited Nov. 30, 2008).

²² See Wand, *supra* note 7.

²³ See Michael C. Herron & Jonathan Wand, *Assessing Partisan Bias in Voting Technology: The Case of the 2004 New Hampshire Recount*, 26 ELECTORAL STUD. 247 (2007).

²⁴ See, e.g., HENRY E. BRADY ET AL., COUNTING ALL THE VOTES: THE PERFORMANCE OF VOTING TECHNOLOGY IN THE UNITED STATES (2001); Stephen Ansolabehere, *Voting Machines, Race, and Equal Protection*, 1 ELECTION L.J. 61 (2002); Walter R. Mebane, Jr., *The Wrong Man Is President! Overvotes in the 2000 Presidential Election in Florida*, 2 PERSP. ON POL. 525 (2004); Michael Tomz & Robert P. Van Houweling, *How Does Voting Equipment Affect the Racial Gap in Voided Ballots?*, 47 AM. J. POL. SCI. 46 (2003); Jonathan N. Wand et al., *The Butterfly Did It: The Aberrant Vote for Buchanan in Palm Beach County, Florida*, 95 AM. POL. SCI. REV. 793 (2001).

is without error is one that should be asked after every election,²⁵ but the question has particular salience in the 2008 New Hampshire Democratic primary because of Clinton's surprise victory. All elections should be subject to audits, and particular attention should be paid to elections that surprise most political observers.²⁶

I. BACKGROUND ON NEW HAMPSHIRE WARDS

This section sets the stage for statistical matching results that follow shortly, and here we make two points. First, we present several illustrations of the fact that voting technology in New Hampshire is not distributed across wards independently of ward-level political characteristics. That is, we provide several examples of the fact that the type of voting technology used in a New Hampshire ward is related to ward features. And second, we show that Clinton vote share varied systematically by ward characteristics.

Note that wards are the smallest New Hampshire voting units and that each ward chooses its own voting technology.²⁷ There are ten New Hampshire counties, and no county is uniform in its use of voting technology.²⁸ This should not be considered surprising since county governments in New Hampshire have no role in election administration.²⁹

A. Voting Technology Across Wards

A very simple point about the distribution of voting technology across New Hampshire wards is that, perhaps not surprisingly, counties with small numbers of voters tend to use PBHC and counties with cities Accuvote.³⁰ For example, the

²⁵ See Wand, *supra* note 7.

²⁶ *Id.*; see, e.g., Brady et al., *supra* note 8 (investigating the 2000 presidential election in Florida because of the large number of votes for Pat Buchanan).

²⁷ N.H. REV. STAT. ANN. § 44:4 (2008) (allowing a city to divide into wards for the purpose of administering elections and stating that wards are equivalent to towns for the purpose of a presidential election); N.H. REV. STAT. ANN. § 656:40 (2008) (permitting mayors or aldermen to authorize the use of voting machines).

²⁸ N.H. REV. STAT. ANN. § 22:1 (2008); see New Hampshire Municipalities Which Use the ACCUVOTE Voting Machine, *supra* note 11 (displaying the voting technology used by each ward); Towns and Wards Districted for Election Purposes, *supra* note 15 (listing every ward's county).

²⁹ See WILLIAM M. GARDENER ET AL., NEW HAMPSHIRE ELECTION PROCEDURE MANUAL: 2006–2007 (2007).

³⁰ Cf. New Hampshire Municipalities Which Use the ACCUVOTE Voting Machine, *supra* note 11; Presidential Primary, *supra* note 13 (providing the number of voters by county and ward). These sources, considered together, show the correlation between counties with large populations and the use of Accuvote by those counties.

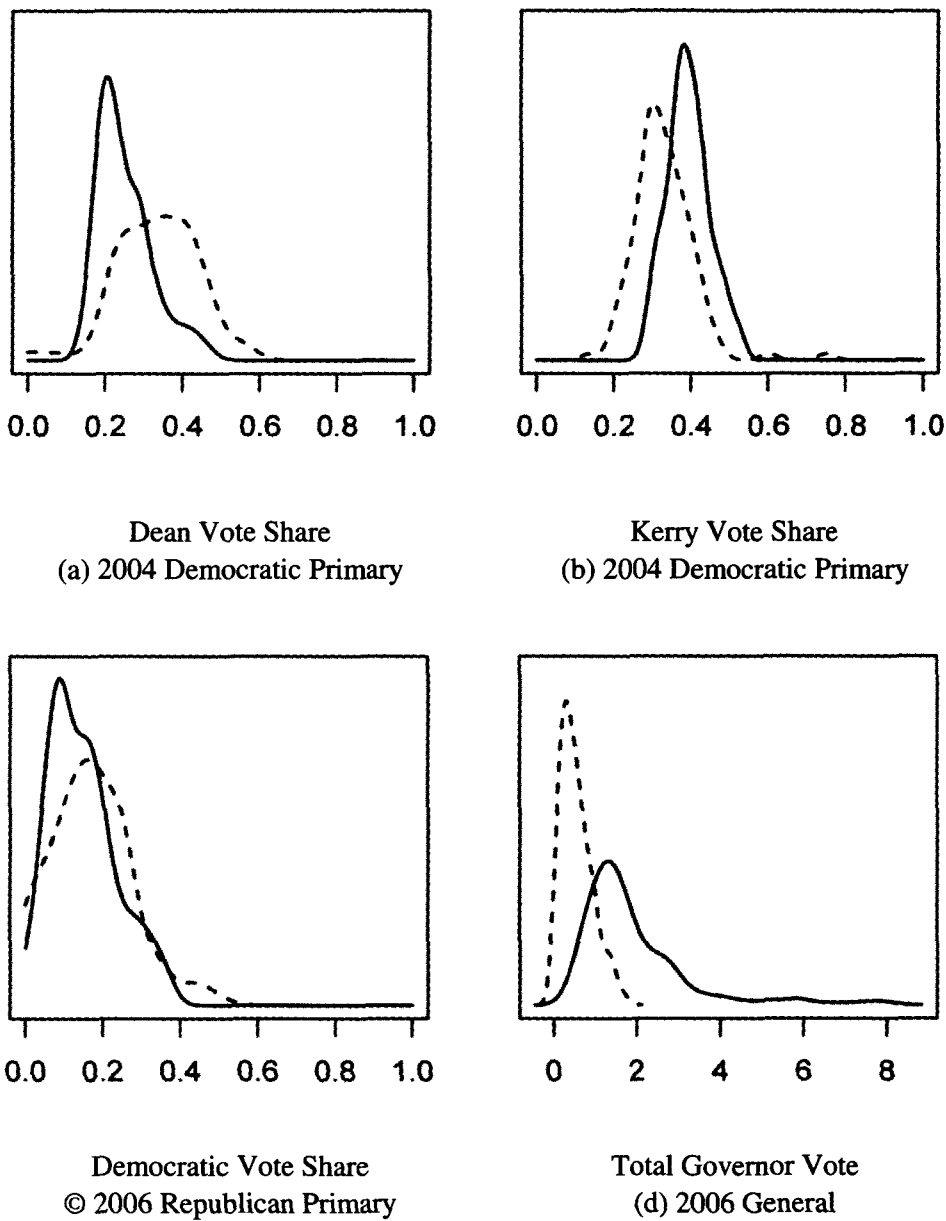
northernmost county in New Hampshire, Coos County, contains forty-six towns that contribute votes to statewide elections.³¹ Of these, forty use PBHC and only six use Accuvote.³² In contrast, Hillsborough County, which contains Manchester, New Hampshire's most populous city, has forty-eight wards of which thirty-four use Accuvote.³³ The point here is that ward size is correlated with ward voting technology.

Consider Figure 1(a); Figures 1(b)–1(d) are qualitatively similar. Figure 1(a) displays smoothed densities based on ward-level Howard Dean vote shares from the 2004 Democratic presidential primary. The solid line density depicts Dean vote share among Accuvote wards, and the dashed line density depicts Dean vote share among PBHC wards. What is notable is that these densities differ: for example, the modal Accuvote ward contained on the order of 20% Dean supporters while the modal PBHC ward contained on the order of 30% to 40% Dean supporters. Explaining patterns of Dean vote share from 2004 is beyond the scope of this analysis and in fact what ultimately motivated Dean voters to support Dean is not germane to the exercise presented here. The point we want to make is that Dean received disproportionately more support in PBHC wards than in Accuvote wards, and in contrast, Figure 1(b) shows that John Kerry received disproportionately more support in the 2004 Democratic presidential primary among Accuvote wards than among PBHC wards.

³¹ See Towns and Wards Districted for Election Purposes, *supra* note 15.

³² See New Hampshire Municipalities Which Use the ACCUVOTE Voting Machine, *supra* note 11 (listing, *inter alia*, the six Coos County wards using Accuvote).

³³ Cf. New Hampshire Municipalities Which Use the ACCUVOTE Voting Machine, *supra* note 11 (listing thirty-four Hillsborough County wards using Accuvote out of a total of forty-eight wards); NEW HAMPSHIRE OFFICE OF ENERGY AND PLANNING, 2007 POPULATION ESTIMATES OF NEW HAMPSHIRE CITIES AND TOWNS (2008), available at http://www.nh.gov/oep/programs/DataCenter/Population/documents/2007_population_estimates.pdf (showing Manchester to have the largest population of any city in New Hampshire at 108,580); Towns and Wards Districted for Election Purposes, *supra* note 15 (listing the wards in Hillsborough County).



Note: Each figure displays two smoothed densities based on ward-level variables. Solid line densities are for Accuvote wards and dashed line densities for PBHC wards. In this figure, “2004 Democratic Primary” refers to the 2004 Democratic presidential primary, and “2006 Republican Primary” refers to the 2006 gubernatorial primary. Finally, Total Governor Vote in the 2006 general election is in thousands of voters.

Figure 1: Variance Across New Hampshire Wards by Voting Technology

Continuing with this logic, Figure 1(c) highlights differences by voting technology in the fraction of Democratic votes cast in the 2006 Republican gubernatorial primary. These votes are write-ins and are cast by individuals who were given Republican ballots yet selected a Democratic candidate.³⁴ Whether this behavior reflects moderate policy preferences, confused voters, voters who forgot to change party affiliations before a primary, ward-administration deficiencies, or some other factor is not known. What we can say, though, is that whatever drives this sort of crossover voting varies by type of ward.

Finally, Figure 1(d) highlights the correlation between ward size and voting technology that we commented on earlier when discussing Coos and Hillsborough Counties. The Figure uses the total number of votes in the 2006 gubernatorial election as a measure of ward size. It is clear from the two densities in the Figure that in New Hampshire PBHC is found only in relatively small wards but that some small wards use Accuvote.³⁵

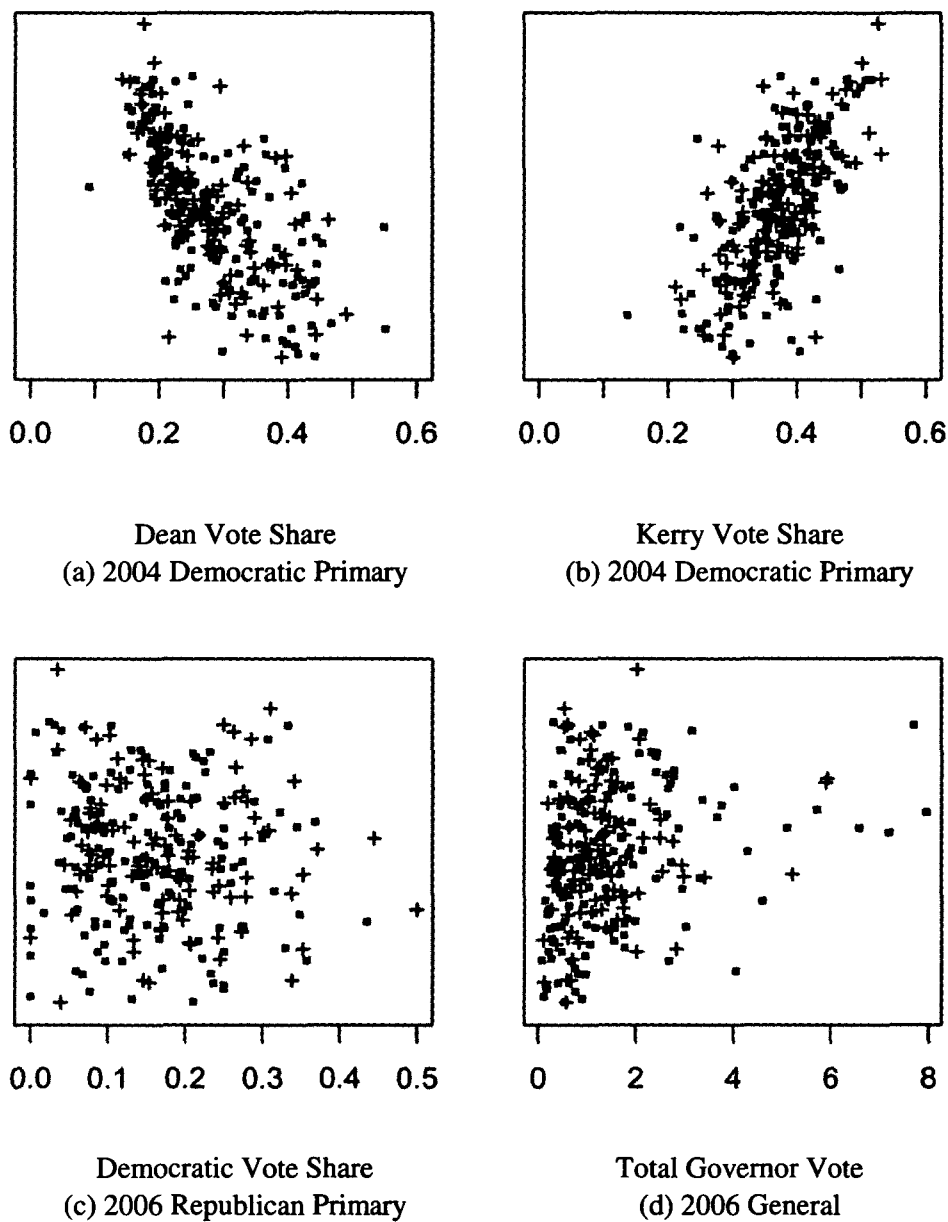
We note that Figure 1 contains only four variables that are correlated with voting technology. Many more such variables exist, and we draw on the four from the Figure plus others in our subsequent matching analysis.

B. Clinton Vote Share across Wards

We now turn to Figure 2, which presents four scatterplots analogous to the four densities highlighted above. Each panel in the figure plots Clinton vote share among all candidates against a ward-level variable. For instance, Figure 2(a) (which parallels Figure 1(a)) plots Clinton vote share from 2008 against Dean vote share from the 2004 Democratic presidential primary. The correlation between these variables is -0.539: wards with many Dean voters in 2004 had fewer Clinton voters in 2008.

³⁴ Such votes were not added by the New Hampshire Secretary of State to the candidates' vote totals in the Democratic primary. See State of New Hampshire Elections Divisions, <http://www.sos.nh.gov/stateprimary2006/rgovsum.htm> (last visited Nov. 30, 2008) (including votes for Democrat Lynch in the Republican primary as votes for the Republican nomination).

³⁵ See *infra* p. 361 fig.2(d). The largest ward using PBHC is slightly over 2000. Some wards with fewer than 100 voters use Accuvote.



Note: Panel labels are identical to those in Figure 1, and each panel plots Clinton share of the total Democratic primary vote against another variable. Within each panel a plus sign denotes a New Hampshire Accuvote ward and a square dot a PBHC ward.

Figure 2: Variance in Clinton Vote Share

In contrast, the correlation between the 2008 Clinton vote share and the Kerry vote share from the 2004 presidential primary is 0.579; these two vote shares are depicted in Figure 2(b). We do not observe a notable bivariate correlation between Clinton vote share and the Democratic vote in the 2006 Republican primary (correlation is -0.119), but we do observe a correlation of 0.239 between Clinton vote share and the size of the governor vote in the 2006 general election. Namely, larger wards were more pro-Clinton.

Why precisely wards that were pro-Dean in 2004 were anti-Clinton in 2008 is beyond our scope. One could conjecture that Dean voters were anti-establishment and that Clinton is seen, among contending Democrats, as a rather pro-establishment figure. Regardless of whether this conjecture is true, the combination of Figures 1(a) and 2(a) suggests a plausible explanation for a correlation between Clinton vote share and voting technology that has nothing to do with vote-tabulating problems.

II. ESTIMATES OF VOTE-TABULATING TECHNOLOGY EFFECTS

We have now illustrated that, as of January, 2008, the distribution of voting technology across wards in New Hampshire depends on ward characteristics.³⁶ Any analysis of the effect of voting technology in New Hampshire on ward-level election outcomes will be confounded by features of wards that are correlated with both voting technology and ward political profile.³⁷ To estimate how much of the Accuvote-PBHC difference in the candidates' votes is due to something inherent to technology, we need a way to purge the comparison of the multitude of differences between wards that used each technology.

To do this we use genetic matching to find a set of matched pairs of wards.³⁸ Each pair contains one ward that used Accuvote technology and one that used PBHC, and the two wards are similar with respect to a large set of observable characteristics whose values were determined before the 2008 primaries.³⁹ A set of matched wards is deemed to be well-balanced and suitable for making comparisons if each observable characteristic has the same distribution in the two types of wards.⁴⁰ Borrowing the language of experiments, we look at PBHC as the "treatment" and Accuvote as the "control" technology,⁴¹ and we investigate the effect of having PBHC technology on the wards that had PBHC technology. This is formally described as estimating

³⁶ See Wand, *supra* note 7, at 2.

³⁷ See *id.* at 5.

³⁸ See Jasjeet S. Sekhon, *Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching Package for R*, J. STAT. SOFTWARE (forthcoming) (manuscript at 3, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1009044).

³⁹ See Wand, *supra* note 7, at 9.

⁴⁰ See Herron & Wand, *supra* note 23, at 254.

⁴¹ See *id.* at 256 (describing a similar matching exercise).

the Average Treatment Effect on the Treated (ATT).⁴² Because vote-tabulating technology is not randomly assigned to wards, we do not expect and indeed do not observe that there are good matches among the wards that used Accuvote for all of the wards that used PBHC. In fact, of the 110 PBHC wards in our analysis, only twenty-four have a sufficiently good match among the wards that used Accuvote.⁴³ In a strict sense, our inferences about the effects of vote-tabulating technology apply only to this subset of the wards. But these are the only wards about which we can have any confidence that other characteristics we are able to observe—and which we know are both imbalanced across technologies and related to the distribution of votes—are not responsible for any differences that may appear.⁴⁴

The matching method we use is one-to-one matching with replacement.⁴⁵ Table 2 shows the definition of the treatment indicator variable (vs10 = 1 for wards using PBHC, 0 for wards using Accuvote), along with the set of variables used to estimate the probability that each ward uses PBHC, conditional on those characteristics. This probability is known as the estimated propensity to receive the treatment, given the

⁴² Alberto Abadie & Guido W. Imbens, *Large Sample Properties of Matching Estimators for Average Treatment Effects*, 74 *ECONOMETRICA* 235, 239 (2006). Ideally we would also like to estimate the effect of having Accuvote technology on the wards that had Accuvote technology, which would allow us to draw more comprehensive conclusions about the effects of the technology. Unfortunately there is insufficient overlap between the two types of wards to support such an analysis. See *id.* at 237–38. Across New Hampshire more wards used Accuvote technology, and their sizes and urban or suburban characters make too many of them substantially different from the wards that used PBHC. See New Hampshire Municipalities Which Use the ACCUVOTE Voting Machine, *supra* note 11.

⁴³ In all, 148 of New Hampshire's 323 wards used PBHC, but of the PBHC wards eighteen recorded zero votes in the 2008 primary, nineteen recorded zero votes in the 2006 primary, twenty-three were missing 2006 average wage data, and twenty-six were missing poverty rate data. Excluding the wards with zero votes or missing data leaves 110 PBHC wards. Only one of the 175 wards that used Accuvote was excluded for these reasons (one ward is missing 2006 average wage data).

⁴⁴ See Wand, *supra* note 7, at 9 (explaining that comparisons can only be made when all features other than voting technology are controlled).

⁴⁵ See Abadie & Imbens, *supra* note 42, at 240 for a discussion of matching with replacement. The relevant calls to functions in the Matching package have arguments as indicated in the following, where Y is the vector of outcomes of interest (e.g., vote counts or vote shares for a candidate), X is a matrix of characteristics to be matched on, and BM is a matrix of the characteristics on which the genetic matching algorithm attempts to achieve balance. Sekhon, *supra* note 38.

```
genout <
GenMatch(Tr=Tr,X=X, BalanceMatrix=BM, estimand="ATT", M=1,
  pop.size=10000, max.generations=101, wait.generations=100,
  caliper=rep(1,ncol(X)))
Match(Y=Y, Tr=Tr, X=X, estimand="ATT", Weight.matrix=genout,
  caliper=rep(1,ncol(X)))
```

See *id.*

values of the conditioning variables.⁴⁶ The conditioning variables are a collection of vote proportions, vote counts, and functions of proportions and counts from the 2004 New Hampshire presidential primary and 2006 New Hampshire gubernatorial primary. The set of conditioning variables also includes a variable that measures one aspect of the economic profile of each ward (the percent of persons in poverty as of 2000). We use a simple logistic regression model to compute the estimated propensities. The exact formulation for this model using the glm function of R appears at the end of Table 2.⁴⁷ We match on the estimated propensity score along with a subset of the variables used to compute the estimated propensity score.⁴⁸ We set a caliper of one standard deviation for each of the matched variables, and this means that we drop all matches that have more than one standard deviation of discrepancy for any of these variables. Table 3 lists the variables on which the matching algorithm seeks balance. This set includes some additional measures of characteristics of the 2004 and 2006 primaries, some additional measures of each ward’s income profile, and a measure of the number of people living in each ward as of 2002.

Table 2: Treatment and Propensity Variables

Treatment Variable

vs10 vote tabulating technology: 1 if PBHC, 0 if Accuvote

“Propensity” Variables

wpkerry04pd standardized Kerry proportion in 2004 Democratic presidential primary

wdpgov06T standardized difference between proportion voting Democratic in the 2006 gubernatorial primary and Kerry’s proportion in 2004 Democratic presidential primary

Pdean04p Dean proportion in 2004 Democratic presidential primary

Pclark04p Clark proportion in 2004 Democratic presidential primary

⁴⁶ See generally Paul R. Rosenbaum & Donald B. Rubin, *The Central Role of the Propensity Score in Observational Studies for Causal Effects*, 70 BIOMETRIKA 41 (1983).

⁴⁷ See also R DEV. CORE TEAM, R: A LANGUAGE AND ENVIRONMENT FOR STATISTICAL COMPUTING (2008), <http://cran.r-project.org/doc/manuals/refman.pdf> 1. See the R Project for Statistical Computing, <http://www.r-project.org/> (last visited Nov. 30, 2008), for a general description and introduction to R.

⁴⁸ The variables used for matching are the estimated propensity score plus all the covariates used to estimate the propensity score except fdp06R and f06D - f06R. These are the variables included in the matrix X in the invocation of GenMatch.

Plieberman04p	Lieberman proportion in 2004 Democratic presidential primary
Pkerry04p	Kerry proportion in 2004 Democratic presidential primary
pforRinD04	proportion voting for a Republican candidate in the 2004 Democratic presidential primary
pforDinR04	proportion voting for a Democratic candidate in the 2004 Republican presidential primary
frp06D	proportion voting for a Democratic candidate in the 2006 Republican gubernatorial primary
fdp06R	proportion voting for a Republican candidate in the 2006 Democratic gubernatorial primary
f06D - f06R	difference between proportion of all voters voting in the Democratic 2006 gubernatorial primary and proportion voting in the Republican 2006 gubernatorial primary
povrate	percent of persons in poverty: persons are classified as being below the poverty level by comparing their total 1999 income to an income threshold. Refer to http://www.census.gov for more detailed information. Source: 2000 Census SF-3

Note: The propensity variable pfit is computed as follows.

```
z <- glm(vs10 ~ wpkerry04pd + wdpgov06T + povrate + Pdean04p + Pclark04p +
Plieberman04p + Pkerry04p + pforRinD04 + pforDinR04 + frp06D + fdp06R +
I(f06D - f06R) , family="quasibinomial")
pfit <- as.real(fitted(z))
```

Table 3: Variables Used to Measure Balance When Matching

wpkerry04pd	standardized Kerry proportion in 2004 Democratic presidential primary
wdpgov06T	standardized difference between proportion voting Democratic in the 2006 gubernatorial primary and Kerry’s proportion in 2004 Democratic presidential primary

Pdean04p	Dean proportion in 2004 Democratic presidential primary
Pclark04p	Clark proportion in 2004 Democratic presidential primary
Plieberman04p	Lieberman proportion in 2004 Democratic presidential primary
Pkerry04p	Kerry proportion in 2004 Democratic presidential primary
pforRinD04	proportion voting for a Republican candidate in the 2004 Democratic presidential primary
pforDinR04	proportion voting for a Democratic candidate in the 2004 Republican presidential primary
povrate	percent of persons in poverty: persons are classified as being below the poverty level by comparing their total 1999 income to an income threshold. Refer to http://www.census.gov for more detailed information. Source: 2000 Census SF-3
pop2002e	estimated population in 2002. Source: NH Office of Energy and Planning
pfim	median family income: total income received in 1999 by all family members fifteen years of age and older. Source: 2000 Census SF-3
pci	per capita income in 1999. Source: 2000 Census SF-3
pballots.abs06	proportion absentee ballots in the 2006 gubernatorial primary elections
pchecklist.undl06	proportion of voters with partisanship undeclared in 2006 gubernatorial primary election
pchecklist.dem06	proportion of voters with Democratic partisanship in 2006 gubernatorial primary election
gov06T	number of votes recorded in 2006 Democratic and Republican primaries
dpgov06T	number of votes recorded in 2006 Democratic gubernatorial primary

rpgov06T	number of votes recorded in 2006 Republican gubernatorial primary
f06D - f06R	difference between proportion of all voters voting in the Democratic 2006 gubernatorial primary and proportion voting in the Republican 2006 gubernatorial primary
fp06D	proportion voting for a Democratic candidate in either Democratic or Republican 2006 gubernatorial primary
frp06D	proportion voting for a Democratic candidate in the 2006 Republican gubernatorial primary
fdp06R	proportion voting for a Republican candidate in the 2006 Democratic gubernatorial primary
T2004D	number of votes recorded in 2004 Democratic presidential primary

Note: These are the variables included in the matrix BM in the invocation of GenMatch.

Table 4 reports the mean values of several variables in the wards that used PBHC and the wards that used Accuvote, before and after matching.⁴⁹ Even though we are not concerned with any kind of hypothesis testing at this stage, the p-value from a t-test of the formal hypothesis of no difference between these means provides a useful summary for how well both the mean and variance of the distributions in the respective subsets of wards correspond. Large values for these p-values do not guarantee that the achieved level of balance is sufficient for reliable inferences about the effects of the treatment,⁵⁰ nonetheless it is noteworthy and reassuring that none of the p-values in the matched subset of data are small.

⁴⁹ These statistics were computed using the MatchBalance function in the Matching package. Sekhon, *supra* note 38 (manuscript at 8).

⁵⁰ See Alexis Diamond & Jasjeet S. Sekhon, *Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies* 26–27 (Inst. of Governmental Stud., Working Paper No. 35, 2006); Jasjeet S. Sekhon, *Alternative Balance Metrics for Bias Reduction in Matching Methods for Causal Inference* 5 (unpublished manuscript), available at <http://sekhon.berkeley.edu/papers/SekhonBalanceMetrics.pdf>.

Table 4: Balance Tests Before and After Matching

variable	<u>Before Matching</u>			<u>After Matching</u>		
	mean		<i>t</i> -test	mean		<i>t</i> -test
	PBHC	Accuvote	p-value	PBHC	Accuvote	p-value
pfit	0.84874	0.095623	< 0.00001	0.54114	0.47607	0.13273
wpkerry04pd	3.7192	8.1903	< 0.00001	5.2914	5.567	0.18664
wdpgov06T	3.7885	5.6025	0.0019328	4.6297	4.6594	0.9542
povrate	0.064598	0.068807	0.31696	0.055721	0.053367	0.53088
Pdean04p	0.34102	0.25129	< 0.00001	0.31170	0.31363	0.82156
Pclark04p	0.12407	0.12478	0.87565	0.12037	0.12030	0.98418
Plieberman04p	0.058903	0.086213	< 0.00001	0.07099	0.073491	0.4497
Pkerry04p	0.32683	0.39369	< 0.00001	0.34959	0.35209	0.69207
pforRinD04	0.00058559	0.0013068	0.0011216	0.00050928	0.00051322	0.97256
pforDinR04	0.10735	0.12720	0.02611	0.10433	0.10146	0.7454
pop2002e	1584.4	6248.1	< 0.00001	2693.1	2984.9	0.18405
pfim	51807	57333	0.000042	55403	55680	0.89085
pci	21961	23367	0.010485	23231	23193	0.9529
pballots.abs06	0.06533	0.055516	0.026040	0.056382	0.056036	0.96201
pchecklist.undl06	0.50046	0.43456	< 0.00001	0.48623	0.47197	0.55152
pchecklist.dem06	0.20656	0.27398	< 0.00001	0.21323	0.23334	0.16148
gov06T	597.31	1910.2	< 0.00001	986.5	1055.5	0.26347
dpgov06T	67.273	204.24	< 0.00001	105.08	109	0.68952
rpgov06T	56.845	164.28	< 0.00001	89.125	94.958	0.61281
f06D	0.74752	0.75467	0.35224	0.75259	0.75254	0.99533
fp06D	0.59197	0.62149	0.054287	0.61254	0.6089	0.8105
fdp06R	0.0021841	0.00091978	0.36341	0.00090086	0.0012821	0.71886
frp06D	0.17569	0.14901	0.023768	0.16112	0.15386	0.43305
T2004D	303.83	1056.0	< 0.00001	511	548.92	0.29725

Note: Before matching there are 110 PBHC and 174 Accuvote observations. After matching there are 24 matched pairs. Before matching we use the two-sample *t*-test, and after matching we use the paired *t*-test.

In addition to the estimated average effects that using PBHC tabulation had on the votes recorded in wards that used PBHC tabulation, we report three estimators for the uncertainty in those estimates. We apply each estimator both to the raw vote counts recorded for several candidates and to the proportion of the total votes cast in the appropriate primary for each of the candidates. The candidates we focus on are Clinton, Edwards, Kucinich, Obama, and Richardson in the Democratic primary and Giuliani, Huckabee, Paul, Romney, and McCain in the Republican primary. Hill and Reiter discuss the rationale for and performance of the estimation uncertainty estimators we use,⁵¹ and they directly study one estimator used here, namely the Hodges-Lehmann aligned rank test.⁵² This estimator gives directly for each outcome variable a 95% confidence interval for the average treatment effect measured on the scale of the outcome variable.⁵³

Our other two estimators are variations of the weighted least squares estimators Hill and Reiter discuss.⁵⁴ We apply the weights they define to generalized linear models that are suitable for either counts or proportions.⁵⁵ For the vote counts this is an overdispersed Poisson model, and for the vote proportions this is an overdispersed binomial model.⁵⁶ One version of these estimators incorporates a regression adjustment for bias reduction: the variables we matched on are included as regressors along with the treatment indicator variable (vs10) and a constant (for the intercept) in either a weighted overdispersed Poisson model (for the vote counts) or a weighted overdispersed binomial model (for the proportions) for the outcome variable in the matched data set.⁵⁷ The estimated average treatment effect is represented as the coefficient of the treatment indicator variable in the regression.⁵⁸ If the average treatment effect is zero, then the coefficient is zero, but it is not straightforward to translate a nonzero coefficient into an expression of the treatment effect on the original scale of the data. We will be content to determine whether a symmetric 95% confidence interval

⁵¹ See generally Jennifer Hill & Jerome P. Reiter, *Interval Estimation for Treatment Effects Using Propensity Score Matching*, 25 STAT. IN MED. 2230 (2006).

⁵² *Id.* at 2236.

⁵³ *Id.* at 2239 tbl.1.

⁵⁴ *Id.* at 2232.

⁵⁵ To define the weights, for sample size N , let c_j be the number of times observation $j = 1, \dots, N$ is included in the matched sample. Let n_m denote the number of distinct observations in the matched sample. The weight for each observation is $w_j = n_m c_j / \sum_{j=1}^N c_j$. *Id.*

⁵⁶ See *infra* text accompanying p. 371 tbl.5. In R, these models are estimated using the glm functions with respectively the “quasipoiss1n” and “quasibinomial” family arguments. R DEV. CORE TEAM, *supra* note 47, at 1049.

⁵⁷ See *infra* pp. 371–72 tbls.5 & 6.

⁵⁸ See Hill & Reiter, *supra* note 51, at 2233.

contains zero.⁵⁹ The other version of these estimators does not feature any bias adjustment. For this estimator the treatment indicator variable and a constant are the only variables included in the regressions. For both of these estimators we use the sandwich estimator⁶⁰ to obtain the coefficients' standard errors.

Table 5 shows that all of the estimates for the effect that using PBHC tabulation had on votes recorded in wards that used PBHC tabulation are small, and for all but one of the effects the estimated confidence intervals suggest the effects do not differ significantly from zero. The column labeled ATT in the table reports the point estimates for the average treatment effects. For the vote counts these estimates are on the scale of votes, and for the vote proportions they are on the scale of the proportions. The estimates for vote counts range from an average 30.5 vote deficit for Clinton due to having PBHC tabulation to an average 4.8 vote gain for Richardson. These small counts correspond to proportional effects ranging from a 2.1% deficit for Edwards due to having PBHC tabulation up to a 2.5% gain for McCain. The Hodges-Lehmann 95% confidence interval estimates suggest, however, that among the ten candidates only the effects for Edwards differ significantly from zero. The upper bounds of the Hodges-Lehmann intervals are less than zero both for Edwards's vote counts and for his vote proportions. For the other candidates the Hodges-Lehmann intervals include zero. The 95% confidence intervals for the coefficient of the treatment indicator variable in the weighted regression models for vote counts mostly confirm these inferences. All of the intervals for the bias-adjusted models include zero; this confirms the Hodges-Lehmann results for all the candidates except Edwards. The 95% confidence intervals for the vote count models without bias adjustment show significant negative effects for both Edwards and Huckabee. The intervals for the coefficient of the treatment indicator variable in the weighted regression models for vote proportions show significant negative effects for Edwards and significant positive effects for both Kucinich and Richardson. In view of the small vote shares the two latter candidates received, it is likely that the vote proportion results for these candidates stem from differences not in votes for them but in the aggregate of votes for all the other candidates.

⁵⁹ Using $\hat{b}$ to denote the coefficient estimate of interest and $SE(\hat{b})$ to denote its estimated standard error, we compute the 95% confidence interval using $\hat{b} \pm SE(\hat{b})$.

⁶⁰ See Peter J. Huber, *The Behavior of Maximum Likelihood Estimates Under Nonstandard Conditions*, 1 FIFTH BERKELEY SYMP. ON MATHEMATICAL STAT. & PROBABILITY 221 (1967); see also R.J. CARROLL ET AL., THE SANDWICH (ROBUST COVARIANCE MATRIX) ESTIMATOR 1 (1998), <http://www.stat.tamu.edu/ftp/pub/rjcarroll/sandwich.pdf>.

Table 5: Estimated Average Treatment Effects: Original (Pre-Recount) Official Count

Vote Counts		Weighted Regressions					
		Hodges-Lehmann		Bias-adjusted		Not adjusted	
		lower	upper	lower	upper	lower	upper
variable	ATT						
Clinton	-30.5	-72.5	19.5	-0.361	0.053	-0.304	0.067
Obama	-19.8	-59.5	16.5	-0.279	0.122	-0.260	0.080
Edwards	-24.9	-43.0	-3.5	-0.450	0.021	-0.401	-0.016
Kucinich	1.2	-3.5	5.0	-0.174	0.682	-0.176	0.532
Richardson	4.8	-5.5	14.5	-0.142	0.334	-0.045	0.383
Giuliani	-3.9	-13.5	6.0	-0.326	0.163	-0.340	0.044
Huckabee	-14.5	-31.5	4.0	-0.520	0.088	-0.480	-0.017
Paul	-7.4	-19.5	5.0	-0.459	0.107	-0.450	0.047
Romney	-14.4	-36.5	8.0	-0.326	0.123	-0.322	0.022
McCain	-12.9	-47.0	23.0	-0.267	0.121	-0.274	0.052

Vote Proportions		Weighted Regressions					
		Hodges-Lehmann		Bias-adjusted		Not adjusted	
		lower	upper	lower	upper	lower	upper
variable	ATT						
Clinton	-0.0127	-0.0400	0.0183	-0.185	0.063	-0.156	0.085
Obama	0.0148	-0.0083	0.0419	-0.068	0.186	-0.080	0.140
Edwards	-0.0211	-0.0328	-0.0063	-0.230	-0.013	-0.219	-0.057
Kucinich	0.0022	-0.0026	0.0066	0.031	0.718	0.007	0.580
Richardson	0.0156	0.0024	0.0224	0.051	0.394	0.160	0.500
Giuliani	0.0045	-0.0078	0.0159	-0.111	0.185	-0.108	0.131
Huckabee	-0.0137	-0.0334	0.0098	-0.299	0.069	-0.268	0.039
Paul	-0.0103	-0.0263	0.0090	-0.250	0.118	-0.210	0.100
Romney	-0.0024	-0.0326	0.0281	-0.140	0.180	-0.119	0.134
McCain	0.0246	-0.0063	0.0538	-0.037	0.176	-0.014	0.161

Note: The weighted regression is an overdispersed Poisson model for the vote counts and an overdispersed binomial model for the vote proportions. ATT point estimates and Hodges-Lehmann confidence intervals are in vote count or vote proportion units.

Weighted regression confidence intervals are for the coefficient of the treatment indicator variable.

Table 6: Estimated Average Treatment Effects: Republican Recount

Vote Counts

variable	ATT	Weighted Regressions					
		Hodges-Lehmann		Bias-adjusted		Not adjusted	
		lower	upper	lower	upper	lower	upper
Giuliani	-5.8	-14.5	3.0	-0.348	0.115	-0.370	-0.019
Huckabee	-13.2	-29.5	3.5	-0.512	0.099	-0.469	0.002
Paul	-8.1	-20.5	4.5	-0.466	0.108	-0.459	0.034
Romney	-14.1	-36.0	9.5	-0.327	0.123	-0.323	0.021
McCain	-14.4	-48.5	19.5	-0.276	0.118	-0.283	0.045

Vote Proportions

variable	ATT	Weighted Regressions					
		Hodges-Lehmann		Bias-adjusted		Not adjusted	
		lower	upper	lower	upper	lower	upper
Giuliani	0.0020	-0.0109	0.0114	-0.134	0.132	-0.142	0.064
Huckabee	-0.0121	-0.0323	0.0114	-0.291	0.082	-0.256	0.059
Paul	-0.0107	-0.0275	0.0084	-0.258	0.119	-0.218	0.090
Romney	-0.0020	-0.0324	0.0289	-0.136	0.174	-0.121	0.131
McCain	0.0221	-0.0089	0.0517	-0.048	0.168	-0.029	0.150

Vote Count Changes

Vote Proportion Changes

variable	ATT	Hodges-Lehmann		ATT	Hodges-Lehmann	
		lower	upper		lower	upper
Giuliani	-1.8	-30.0	3.0	-0.0025	-0.0280	0.0053
Huckabee	1.2	-1.5	11.5	0.0016	-0.0020	0.0117
Paul	-0.7	-1.5	2.0	-0.0005	-0.0011	0.0028
Romney	0.3	-1.5	1.5	0.0003	-0.0021	0.0029
McCain	-1.5	-3.5	1.0	-0.0025	-0.0064	0.0010

Note: The weighted regression is an overdispersed Poisson model for the vote counts and an overdispersed binomial model for the vote proportions. ATT point estimates

and Hodges-Lehmann confidence intervals are in vote count or vote proportion units. Weighted regression confidence intervals are for the coefficient of the treatment indicator variable.

Table 6 is similar to the previous table, but it is based on Republican recount data. As previously discussed, the Republican recount covered all of New Hampshire unlike the partial Democratic recount.⁶¹ Results based on Republican recount figures parallel those based on pre-recount figures.⁶²

Table 6 also contains estimates of the effect of PBHC in wards that used PBHC on changes between original (pre-recount) and certified (post-recount) figures. What is apparent from this table is that the confidence intervals for the effects of PBHC on such changes are sufficiently wide so as to make associated estimates statistically negligible.⁶³ Finding otherwise, say that the effect of PBHC in PBHC wards on Giuliani vote share was significantly positive, would have been troubling.

CONCLUSION

We find no significant relationship between a ward's use of vote-tabulating technology and the votes or vote shares received by most of the leading candidates who competed in the 2008 New Hampshire presidential primaries. Among Clinton, Edwards, Kucinich, Obama, and Richardson in the Democratic primary and among Giuliani, Huckabee, Paul, Romney, and McCain in the Republican primary, we observe a significant average effect of using PBHC technology on the wards that used PBHC technology only in the votes counted for Edwards, and that difference is small. The effects for Edwards also do not appear to be significant when a regression-based bias adjustment is applied.⁶⁴

Our analysis of post-recount Republican primary election returns did not uncover any systematic problems relating to vote-tabulating technology for the 2008 New Hampshire primary election, and our analysis of changes due to the recount in Republican candidate vote counts is similarly benign. Thus, if one were to view the conclusions we reached using the pre-recount returns as a kind of prediction that the recount would not uncover issues associated with vote-tabulating technologies, then that prediction would be confirmed.⁶⁵

⁶¹ See Dorgan, *supra* note 2.

⁶² See *No Change in GOP Recount*, *supra* note 9.

⁶³ See *supra* p. 372 tbl.6.

⁶⁴ See *supra* p. 371 tbl.5.

⁶⁵ This is not to say that the recount did not uncover any discrepancies. See, e.g., David Brooks, *Results Largely the Same After Recount*, NASHUA TELEGRAPH, Jan. 22, 2008, <http://www.nashuatelegraph.com/apps/pbcs.dll/article?AID=/20080122/NEWS08/572585355&SearchID=73306340666122>. Nonetheless, the discrepancies that were uncovered are not consistent with the allegations of systematic vote-tabulating problems that motivated the recount.

The particular set of variables used for the matching analysis in this study does not exhaust the range of observable ward attributes. It is possible that another set of matching variables and matched pairs of wards would produce even better balance among observables across technologies than we have found. It is also possible that some observables we have not examined in this study remain imbalanced, contributing to bias in our estimates of the average treatment effects. The observable features we have examined, however, include variables that measure many aspects of the preceding primary elections in the state, as well as many demographic features of wards in the state.⁶⁶

The biggest limitation of this study is that the matching exercise on which it is based produced only twenty-four matched wards.⁶⁷ For most of the remaining hundreds of wards in New Hampshire it is not possible to obtain a direct estimate of the effects of vote-tabulating technology that is not confounded with known extraneous factors.⁶⁸ There are other approaches one might use for an analysis that would use many more or perhaps all of New Hampshire's wards, but such methods depend on strong and unverifiable assumptions about features of abstract analytical models.⁶⁹ No analysis is free from assumptions, but the matching approach we use has the virtue of remaining very close to the observable data. The key assumption we need is that the inherent relationship between the two vote-tabulating technologies is roughly the same throughout New Hampshire. In that case, the average treatment effect we estimate in the subset of matched wards is telling us about what is true throughout the state.

If one suspects that either the 2008 Democratic primary or the 2008 Republican primary election in New Hampshire was affected by irregularities, then naturally one may believe that such irregularities are haphazard or perhaps even artfully disguised. This study is not intended to address such suspicions. But with respect to Hillary Clinton's surprise victory in the Democratic primary and the notable differences across vote-tabulating technologies in Clinton's and others' levels of support, our results are consistent with these differences being due entirely to the fact that New Hampshire wards that use Accuvote optical scan machines typically have voters with different political preferences than wards that use hand-counted paper ballots.

⁶⁶ See Wand, *supra* note 7, at 9.

⁶⁷ See *supra* p. 368 tbl.4.

⁶⁸ See Wand, *supra* note 7, at 5.

⁶⁹ See Sekhon, *supra* note 38.