Surprise and Confusion: Do Pre-Election Polls and Exit Polls Buttress 2004 Fraud Claims?

Mark Lindeman, Bard College

first edition, December 29, 2005

It has been argued that state pre-election polls and state exit polls both point to massive vote stealing in the 2004 election favoring George W. Bush – but do they point in the same direction? It depends on how one looks. My analysis suggests that the pre-election polls and exit polls identify different states as "suspicious," and thus do not mutually support a coherent fraud theory. I further find that independent variables posited to predict pro-Bush fraud tend not to have the hypothesized state-level correlations with respect to pre-election and/or exit polls.

The Great Election Robbery of 2004?

Was the 2004 U.S. presidential election stolen? Some observers have argued that the exit poll results, along with other evidence, points to *massive, widespread* fraud altering the outcome of the election. The NEP national exit poll, prior to incorporation of official vote returns, gave challenger John Kerry approximately a 3-point lead over George W. Bush (E/M 2005, 20). (Although Edison/Mitofsky, who conduct the exit polls for NEP, do not calculate a margin of error for the popular vote, all parties agree that the discrepancy between the exit polls and the official returns is beyond the margin of random sampling error.) Extrapolating NEP figures to the electorate, Steve Freeman argues that the exit poll "numbers indicate… a discrepancy on the order of 10 million votes," with Kerry winning by some 6.5 million votes or 4.6 percentage points (Freeman 2005a, 10; Freeman 2005b). Thus, in this family of hypotheses, fraud is *massive* in that it affects many millions of votes, and *widespread* in that it spans many states.

The pseudonymous Internet political analyst "TruthIsAll" has released an "Interactive Election Model" intended to buttress his assertions that both pre-election polls and exit polls point to a Kerry victory (TruthIsAll 2005c). On the eve of the election, TruthIsAll (TIA) concluded that the latest state pre-election polls gave Kerry approximately a 1-point advantage nationwide before allocating undecideds, and a decisive advantage if undecideds were assumed to break to Kerry (TruthIsAll 2004a).[1] TIA's Monte Carlo analysis indicated that Kerry had a 99.8% chance of winning the Electoral College, with a mean advantage of 3.6% (and 337 Electoral College votes). TIA further reckoned from national polls that Kerry had a mean advantage of 3.25% and a 99.99% probability of winning the popular vote. Thus, TIA and some others interpreted the exit poll results as confirming his projections. Kerry had won – until Bush stole the election.

It is worth emphasizing at the outset that this essay does not, and cannot, answer the question whether the 2004 election was in some sense stolen. A net change of some 120,000 votes in the

[1] TIA believed that undecideds would "break" rather sharply: "For the final projection, the *base case undecided/other allocation assumption* to Kerry has been changed from *60% to 75%*. This is consistent with the opinion of professional political pollsters. To gauge the *sensitivity* of the expected electoral vote and win probability to the allocation, the model calculated five scenarios: *60%, 67%, 75%, 80% and 87%.*" (TruthIsAll 2004a) Apparently the 99.8% figure refers to the 75% allocation, whereas the 60% allocation yielded a 98% Kerry win probability (see http://www.geocities.com/electionmodel/index_files/ElectionModel_9609_image004.png).

state of Ohio would have altered the outcome of the election, and nothing in this essay addresses whether fraud (and/or vote suppression, which might not have shown up in the exit polls) had such an effect. This essay treats the narrower question whether the state pre-election polls and state exit polls mutually support the inference of massive, widespread fraud on the scale of several points nationwide. I argue that they do not. I further address the question whether, as asserted by Freeman (2005a), various "neglected correlations" between exit poll discrepancies and political variables point to massive, widespread fraud. Based on analysis of a 49-state data set, I conclude that they do not. The results do not *rule out* massive, widespread fraud, but in concert with other analytical results (e.g. Mebane and Herron 2005; Liddle 2005; Kyle et al. 2005; Lindeman 2005), they further weaken the case for such fraud.

A note on assumptions

Many political observers are nonplussed by the claims that the pre-election polls gave Kerry a prohibitive advantage and that the exit poll error is most likely due to fraud. Indeed, these claims have often not been perceived as meriting refutation; the emerging political science literature on "why Bush won the 2004 election" overwhelmingly assumes that Bush did win the 2004 election. While many political scientists gave John Kerry a good chance of winning the election, few if any put his chances above 99% based on pre-election polls.[2] Nor have many experts openly embraced the view that exit polls should be assumed bias-free, and that therefore any systematic exit poll error is prima facie evidence of fraud. I myself believe that Bush led in the pre-election polls, and that a 3-point exit poll error in vote share is perhaps strange but not especially suspicious. If one wished to convince me that the 2004 election was stolen, one would be well advised to drawn upon other evidence.

Yet the argument in this paper does not depend upon bickering about how to interpret pre-election polls or exit polls. Instead, I explore the implications of assuming that TIA is fundamentally correct about both the pre-election polls and the exit polls. To explain why I doubt this assumption *a priori* would fill a much longer paper.

Surprise: concepts and expectations

Do the exit poll results confirm TIA's pre-election projections? Let us suppose that TIA's state projections and the state exit poll results are both accurate measures of voters' intentions[3] – subject to sampling error – and that the official returns reflect electoral fraud favoring Bush, at

---

[2] TruthIsAll's estimates appear to favor Kerry in at least three ways. First, TIA preferred "registered voter" results to "likely voter" results, which typically produced lower estimates of Kerry support. Second, as we have seen, he aggressively allocated most undecided voters to Kerry. Third, he assumed that any errors across polls are uncorrelated, thus excluding even the possibility that last-minute events could boost Bush's (or Kerry's) vote share in many states simultaneously.

Some observers might further accuse TIA of "cherry-picking" polls. Indeed, it is hard to tell how TIA managed to reckon a 3-point pre-election advantage for Kerry in Ohio – prior to allocating undecideds! – whereas seven of the last eight state polls listed in the electoral-vote.com database put Bush ahead (http://www.electoral-vote.com/2004/info/allpolls.csv, final surveys by Opinion Dynamics, Zogby, Strategic Vision, Survey USA, Gallup, University of Cincinnati, Rasmussen, and Mason-Dixon). I assume that, apart from his preference for registered voter results, TIA did not intentionally cherry-pick his figures.

[3] "Voters" here would include some people who cast provisional ballots that were not counted. (In Ohio, approximately 35,000 provisional ballots went uncounted, or about 0.6% of the total presidential vote.)

least in some states. Thus, the official returns are a "**surprise**" in two ways: they do not fit with the pre-election polls (as interpreted by TIA), and they do not fit with the exit poll projections. Each of these surprises yields a measure (or a family of measures) that should be directly influenced by election fraud. The difference between pre-election polls and official returns, I will call "**pre-election surprise**" or "**vote/pre surprise.**" The difference between exit poll projections and official returns, I will call "**exit poll surprise**" or "**vote/exit surprise.**"

How might these two surprise measures be related? The answer broadly depends upon whether fraud is assumed to be relatively uniform across states, or to vary systematically, perhaps depending on differences in motivation and/or opportunity. *If fraud were relatively uniform (or randomly distributed) across states*, then pre-election surprise and exit poll surprise would likely be uncorrelated variables with similar means; assuming that exit polls are inherently unbiased, exit poll surprise should not covary with any other variable of interest. (The variance of pre-election surprise might reflect other variables, such as, perhaps, the relative success of Democratic turnout efforts.)[4] However, fraud theorists generally do not posit that fraud would be uniform. Freeman argues that exit poll discrepancies are larger in swing states, in states with Republican governors, and in states with large African-American populations (Freeman 2005a). Many who suspect fraud have singled out touch-screen voting machines (also known as DREs, Direct Recording Electronic) as a likely vector of fraud because of known security problems and the absence of a paper trail. TruthIsAll has written that "[t]he EAST was the BEAST," i.e., that exit poll discrepancies are largest in the Eastern time zone (TruthIsAll 2005a), perhaps because fraud was targeted to "[j]ump quickly ahead in the popular vote" and win the key states of Ohio and Florida (TruthIsAll 2005b). Thus, it seems reasonable to expect that fraud should vary markedly and systematically across states.

If fraud is systematically larger in some states than in others, then surprise measures should be positively correlated. We expect this correlation because Bush should do surprisingly well in some states' official counts compared with both exit polls and pre-election returns, while in others his performance is close to expectations, with small surprise by either measure.[5] The term "close" is ineluctably fuzzy. For instance, if the pre-election polls tended to underestimate Kerry's performance in some states – perhaps because Democratic turnout was higher than the likely voter models expected – then one would expect to find a substantial pro-Kerry pre-election surprise in some stats. Note, however, that TIA's argument posits a fairly close fit between the pre-election polls and the exit polls. Thus TIA seems to assume that the pre-election polls are at least on average an unbiased measure of Kerry's true electoral support. At any rate, all else equal, if both vote/pre and vote/exit surprise measure fraud, then the surprise measures are likely

---

[4] Fraud theorists often refer to the high overall turnout in the 2004 election, which some assume unambiguously benefits the Democratic candidate – although, as James Campbell has noted, that conventional wisdom has found little support (Campbell 2005). Higher turnout (as a proportion of voting age population) does not appear to predict vote/pre surprise; in fact, the correlation is slightly negative, hinting that higher turnout benefited Bush.

[5] Note that a positive correlation would be expected even if Bush does not literally do surprisingly well by both measures. For instance, suppose that the pre-election polls had a strong (but uniform) bias favoring Bush, such that Bush everywhere did worse than predicted. It might then be the case that that Bush did much worse than predicted in non-fraud states (negative vote/pre surprise), and about the same as predicted in fraud states (nearly zero vote/pre surprise). Thus, vote/pre surprise would still be mathematically greater (less negative) in the fraud states than in the non-fraud states, as would be vote/exit surprise.

to covary (unless, as discussed below, fraud were concentrated in states where Kerry *otherwise would have done* especially well compared to pre-election polls).

Measuring pre-election and exit poll surprise

Pre-election (vote/pre) surprise depends on which pre-election polls one uses and how one interprets them. It is crucial to TruthIsAll's arguments that, properly interpreted, the pre-election polls (both state and national) gave Kerry a decisive advantage, although many analysts have concluded on the contrary that pre-election polls correctly showed Bush ahead (e.g., Pew Research Center 2004). To give, presumably, the fairest test of TruthIsAll's argument, I base my measures of pre-election surprise upon his numbers. Specifically, I begin with the values given in his "Interactive Election Simulation" spreadsheet (TruthIsAll 2005c), StatePreExit tab, columns E and F, which appear generally to coincide with the values used in his pre-election analysis.[6] Measure 1 is based on the raw difference, $P_K - P_B$, where P_K and P_B are, respectively, the Kerry and Bush proportions in the latest available pre-election poll as cited by TIA. Measure 2 is based on a subtler difference measure that (again following the example of the spreadsheet) allocates 67% of estimated undecideds to Kerry.[7] Thus, the surprise measures **TIASURP1** and **TIASURP2** equal the difference between the Democratic margin in the official returns ($V_K - V_B$, Kerry vote proportion minus Bush vote proportion), and the expected margins in TIA's analyses; TIASURP1 disregards undecided respondents, while TIASURP2 allocates them in Kerry's favor. As a check on these results, I have also examined a third vote/pre surprise measure based on E/M's prior estimates (courtesy of Warren Mitofsky); this measure yields substantively similar correlational results.

Exit poll (vote/exit) surprise poses its own measurement issues, because at least five state-level error measures are available. I consider three of them. The Best Geo estimate represents E/M's best weighted[8] projection of the election outcome from exit poll interview data (E/M 2005, 21-22). I consider this the best state-level "exit poll prediction" that does *not* incorporate prior estimates based on pre-election polls. The Best Geo error or surprise (**BGEO04**) is thus the difference between the actual Democratic margin and the margin predicted by the Best Geo estimate. Some analysts have preferred Within Precinct Error measures, which they regard as more directly interpretable because no weights are applied. WPE is the difference between the Democratic vote margin and the exit poll margin within each precinct. The WPE measures are limited by their simplicity: for instance, they simply disregard absentee voters, and they do not

---

[6] However, the spreadsheet cites "weighted" averages that give Kerry a narrower lead than in the pre-election analysis. The spreadsheet averages are 47.50% for Kerry, 47.02% for Bush, whereas the original model gave Kerry a 47.88%-46.89% advantage (TruthIsAll 2004b). The largest individual changes are in Florida (original 50% Kerry / 47% Bush, spreadsheet 47%/47%) and Arizona (incorrectly labeled "AR" in the spreadsheet; original 48%/48%, spreadsheet 46%/48% favoring Bush). These changes and three other smaller changes seem unlikely to account for the observed difference in weighted averages, but the discrepancy will not detain us.

[7] I say "estimated undecideds" because the spreadsheet actually uses vote returns (although apparently not the final certified returns) to estimate the proportion V_O of votes for candidates other than Bush and Kerry, which is used as an estimate of pre-election support for these other candidates. Thus, the proportion of "estimated undecideds" equals $(1 - P_K - P_B - V_O)$, where P_K and P_B are the poll estimates defined in the text. Note also that the default (but adjustable) 67% allocation in the spreadsheet is smaller than the 75% allocation favored in the final Election Model (TIA 2005a).

[8] The weights include adjustments for observed non-response by age, race, and sex, as well as adjustments for selection probability, geography, and absentee voters (E/M 2005, 9).

incorporate non-response adjustments that could potentially reduce spurious error.[9] However, while WPE measures may be noisier than the Best Geo measure, it is not obvious prima facie that they will be biased toward or against indicating fraud. I consider two different WPE measures: **WPE04**, the "WPE" as computed by VNS prior to 2002 (a numerical average with the two high and two low outliers removed for each state), and **IMWPE04**, the "IM WPE," a numerical average with no outliers removed. Freeman has argued that IM WPE (which he calls PLD, Precinct Level Disparity) is the best measure because the outliers could evince intense focused fraud.

Note that all the surprise measures have the same basic scale and direction. A surprise of, say, +0.04 would mean that Kerry's net margin was four points better in the vote count than "expected" (in pre-election polls or exit polls); a surprise of –0.04 would mean that Kerry did four points worse in margin than "expected." In the ensuing analysis, I report statistics for 49 states, excluding Oregon, which is anomalous because the vote is conducted entirely by mail, as well as the District of Columbia, which votes very heavily Democratic and would be an influential outlier in some analyses.

The two pre-election or vote/pre surprise measures, TIASURP1 and TIASURP2, are highly correlated with each other (Pearson $r = 0.973$). The unweighted means are -0.033 for TIASURP1, and -0.050 for TIASURP2. Thus, both measures indicate that Kerry did worse in the official returns than in pre-election returns, but TIASURP2 reckons a larger surprise, since it assumes greater pre-election support for Kerry by allocating the undecideds in his favor. (The E/M prior estimate measure has a mean of -0.020, again indicating that Kerry did worse than in pre-election polls, although not enough worse to reverse the popular vote outcome.)

The exit poll (vote/exit) surprise measures are also strongly correlated, although unsurprisingly BGEO04 is least like the others.[10] The means are -0.050 for BGEO04, -0.060 for WPE04, and -0.064 for IMWPE04. Thus, at first glance, all the surprise measures except TIASURP1 point to a discrepancy of 5 or more points favoring Bush.[11]

Correlational results: pre-election surprise versus exit poll surprise?

So, does pro-Bush pre-election surprise coincide with pro-Bush exit poll surprise? Counterintuitively, the two surprises tend to point in opposite directions: *when vote/pre surprise most favors Bush, vote/exit surprise on average least favors Bush.* The correlations are similar regardless of measure, and all are statistically significant (Table 1).

[9] For instance, if women participate at a higher rate than men, the raw tallies used to calculate the WPE will reflect a pro-Kerry bias, in comparison with a weighted result that would reflect observed ratios of male and female voters.
[10] r(WPE04, IMWPE04) = 0.951; r(BGEO04, IMWPE04) = 0.811; r(BGEO04, WPE04) = 0.839.
[11] If the results are weighted by Voting Eligible Population as compiled by Michael McDonald (United States Elections Project 2005), the two WPE measures are even more extreme (means: -0.071 for WPE04, -0.075 for IMWPE04); the means of the other measures change little.

Table 1: bivariate correlations among surprise measures (49 states)

Pre-election surprise measure	Exit poll surprise measure		
	BGEO04 (Best Geo estimate "error")	WPE04 (mean WPE, outliers removed)	IMWPE04 (mean WPE, all precincts)
TIASURP1 (TIA without undecideds)	-0.301 ($p = 0.04$)	-0.376 ($p < 0.01$)	-0.333 ($p = 0.02$)
TIASURP2 (TIA, undecideds 67% Kerry)	-0.293 ($p = 0.04$)	-0.389 ($p < 0.01$)	-0.330 ($p = 0.02$)

Figure 1 depicts the relationship between TIA's second (and favored) pre-election surprise measure and my own favored exit poll surprise measure. (A plot using WPE appears in the appendix.) By hypothesis, the relationship should be positive, but the ordinary least-squares regression line indicates that it is negative. That is, the states where Kerry did worst compared to pre-election expectations (at left) tend to be states where he was close to or even above expectations in the final count (at top). If extreme negative surprises evince fraud, then the strongest case for fraud would be in the little-mentioned state of Mississippi;[12] the much-discussed battleground states Ohio, Florida, and Pennsylvania cluster in the middle of the plot.

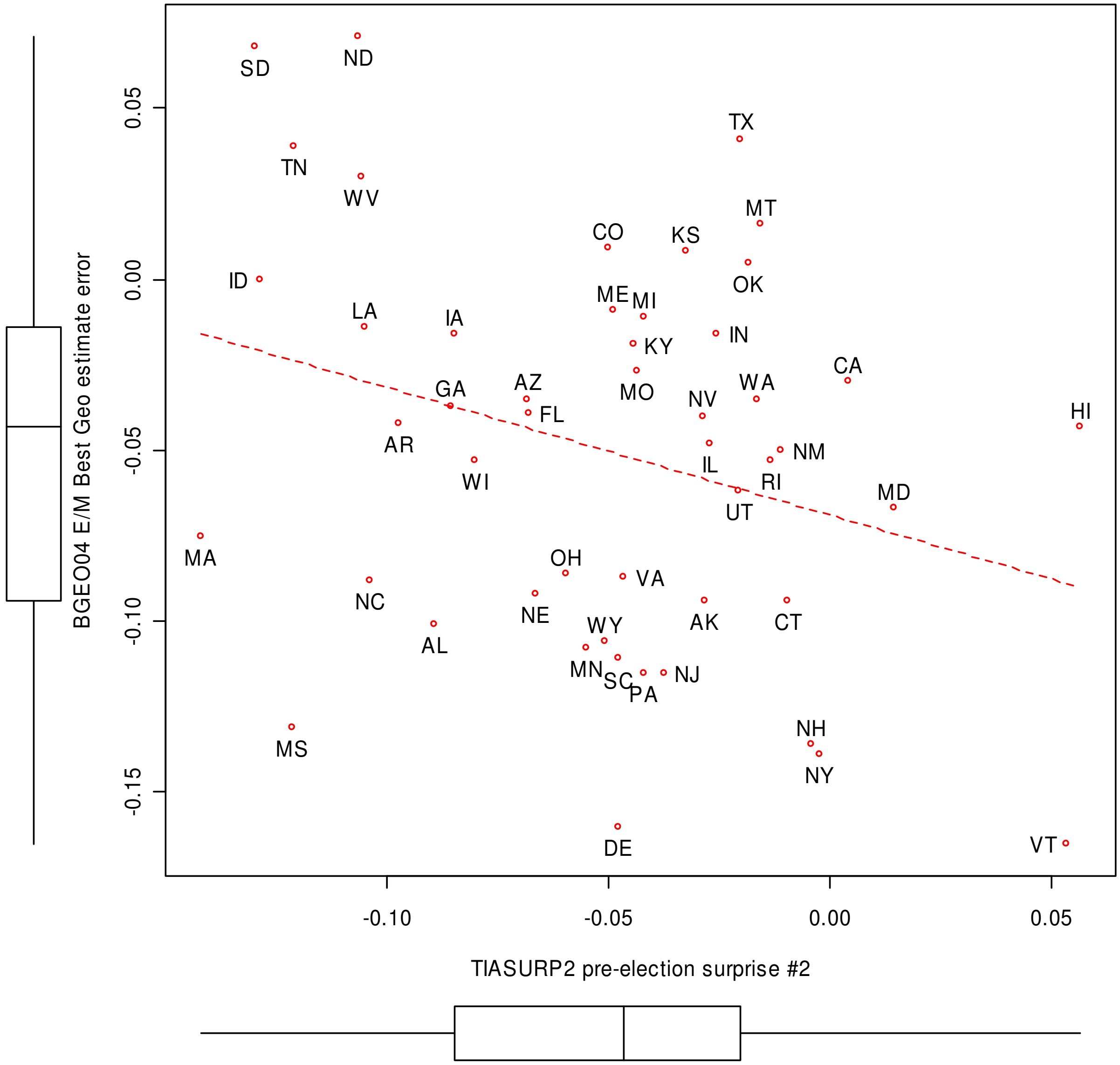
Potentially, these unexpected negative correlations could be unduly influenced by small states. Notice, in Figure 1, the influence of North and South Dakota at upper left and Vermont at lower right. To assess whether the result is affected by the use of smaller samples in many small states' exit polls,[13] we can use the *t* score of the Best Geo estimator in place of the raw percentage error. Perhaps surprisingly, the *t* score yields an even larger negative correlation! While I have no comparable uncertainty measure for the pre-election polls, weights based on the number of surveys completed in each state after September 1 produced similar results.[14] Weighting by the number of votes cast did produce statistically insignificant results for the Best Geo measures (but not for the WPE measures). Thus, the negative correlation between vote/pre and vote/exit surprise, while apparently statistically robust, may not be important for understanding the popular vote. Nonetheless, it poses a riddle.

[12] Of course I do not rule out double-digit fraud in Mississippi *a priori* – although it should be noted that the poll used by TIA took place in mid-September. (While that poll gave Bush a 9-point lead in Mississippi, Edison/Mitofsky's prior estimate put Bush up 17; he won by 20 points.) The Mississippi exit poll was also among the smallest in the country. My point, however, is not to single out one state, but simply to underscore that the scatterplot does not facially support the hypothesis of fraud influencing both vote/pre and vote/exit surprise.

[13] Sample sizes were not always proportionate to size; for instance, the swing state New Hampshire had almost as large a sample as much larger (but less competitive) New York.

[14] Specifically, I used Weighted Least Squares with two different weights: the count of surveys after September 1, and the square root of the survey count. Data were drawn from the electoral-vote.com database described in footnote 2. The poll count ranges from 1 in five different states (Alaska, Delaware, Idaho, Mississippi, and Wyoming) to 90 in Florida and in Pennsylvania. These counts are somewhat misleading because tracking polls are listed more than once, but under the circumstances it seemed futile to pursue further refinements.

Figure 1: exit poll surprise (Best Geo) as a function of pre-election surprise



Interpreting the "surprise surprise"

Recall that we posited a zero correlation between vote/pre and vote/exit surprise if fraud were uniformly or randomly distributed, and a positive correlation if fraud varied systematically. (If there were no fraud, we would presumably again expect a zero correlation.) Why might there be a *negative* correlation? Also, is there any evidence that various systematic hypotheses predicting variations in fraud find support in either surprise measure?

To address these questions, I gathered data for a number of variables relevant to hypotheses offered by Freeman, TruthIsAll, and other observers. I briefly enumerate these variables here:

- Republican governor in 2004, REPGOV04. "Where Republicans controlled the State government, there might have been greater opportunity to steal votes" (Freeman 2005a, 8).[15]
- Swing state in 2004, SWING04. (I use Edison/Mitofsky's list of 11 states [E/M 2005, 42].) "The first priority was, of course, to win the election. Thus, it would makes sense that votes would be most vigorously coveted in the swing states that would determine the winner of the Electoral College" (Freeman 2005a, 7).
- Eastern Standard Time, EASTTIME. ("The EAST was the BEAST," see discussion above.)
- % of population assigned to electronic voting machines (estimated), EVMPCT.[16] "[T]he GAO now confirms that electronic voting machines provided an open door to flip an entire vote count. More than 800,000 votes were cast in Ohio on electronic voting machines, some seven times Bush's official margin of victory." (Fitrakis and Wasserman 2005) In my analysis (as in the quotation), touch screen and push-button Direct Recording Electronic (DRE) equipment are considered electronic voting machines.[17]
- MYVOTE1 "Voter Alert Line" incident calls per million Voting Age Population (VAP), MYV_VAP. The MYVOTE1 line, sponsored by NBC News and several other organizations, received a total of 55,815 complaint calls on election day, according to the figures posted on MSNBC.com.[18] Freeman found that the rate of MYVOTE1 complaint calls was significantly correlated with WPE, "indicating that those problems did in fact result in corrupted counts" (Freeman 2005a, 9). This interpretation is questionable. Clearly complaints about long lines might correlate with artificially depressed turnout, perhaps contributing to vote/pre surprise – but it is not obvious how long lines would result in corrupted *counts* and vote/exit surprise. At any rate, the variable merits further analysis.

---

[15] Since in most states the elections are administered by the Secretary of State (SoS), arguably SoS party affiilation would be a more useful variable, but I follow Freeman's preference here.

[16] The data used here are drawn from a heterodox source: a Democratic Underground post titled "Pobeka's Complete tabulation: voting methods/vendors by population in state," and presently available only via Google's archive. The percentages are based on the reasonably accurate information about voting methods available on the Verified Voting website (http://verifiedvoting.org/verifier). However, they may misstate electronic voting in some instances because counties with more than one vendor or voting method are divided evenly among the vendors and methods. Thus, for instance, Suffolk County in Massachusetts is reckoned as one-half "paper ballots: hand counted" and one-half "optical scan: precinct based" because Essex uses paper ballots, while the rest of the county – including Boston – uses optical scan. (This example does not affect EVMPCT, since opscan is not considered an electronic voting method.)

[17] Optical scan systems – which utilize paper ballots, electronically counted – are also regarded as vulnerable to fraud. I did derive a measure of optical scan percentage, again relying heavily on "Pobeka's" work with a few adjustments. This measure was *positively* correlated with both vote/exit surprise measures, statistically significantly so with WPE ($r = 0.290$, $p = 0.046$). Thus the exit polls might seem to suggest that Kerry stole votes in opscan precincts! Freeman suggests that the most appropriate comparison is between paper ballots (hand counted, or PBHC) and all other methods: "If the ballots are never manually counted anyway or if there is a lax chain of custody on the ballots, then all machines are as vulnerable to programming fraud as electronic machines. But look at the difference [in mean WPE] between Paper ballots and machines of all types!" (Freeman 2005a, 8) One wonders whether this assertion properly applies to mechanical voting machines such as the lever machines used in most of New York. At any rate, while the mean WPE was indeed small in the 40 PBHC precincts – which are some 3% of the sample – so few states make wide use of PBHC that a state-level comparison would be futile.

[18] In an effort to replicate Freeman's measures, I compiled call counts from the digital map at http://www.msnbc.msn.com/id/6364287/. A data appendix released on October 27, 2005 reports slightly different numbers, with a total of 56,024 complaint calls (http://www.sas.upenn.edu/fels/FGRS/MyVote1Appendix.InfoVoterReport.pdf, 2 and passim).

- Election Incident Reporting System (EIRS) election day incidents per million Voting Age Population (VAP), EIRS_VAP. The EIRS initiative, similar in purpose to the MYVOTE1 program,[19] was publicized by various "Election Protection Organizations." EIRS data have been widely cited to buttress charges of election theft (e.g. Miller 2005, 229-237 *passim*). The EIRS reports 42,841 "election incidents" related to the 2004 election, of which over 31,000 were reported on or after Election Day. (The figures here are parsed to remove some duplicate reports.)
- MYVOTE1/EIRS composite election incident score, INCIDENT. If the MYVOTE1 and EIRS data measure the same underlying phenomenon, a measure of their shared variance may perform better than either measure separately. INCIDENT is the first principal component of the natural logarithms[20] of the two indexes MSN_VAP and EIRS_VAP.
- African-American population (natural logarithm of percentage), LNAFRAM. "Traditionally, African Americans are the most disenfranchised group in the country…. And is PLD [WPE] higher in states with more blacks? Yes, it is…" (Freeman 2005a, 8).[21]

I also consider three variables not directly linked to fraud hypotheses.

- New England state, NEWENG. Observationally, exit poll errors appeared to be higher in many New England states for no obvious reason.
- Democratic presidential vote margin, DEMARG04. I include this as something of a nuisance variable. Freeman notes that WPE is apparently larger in Republican precincts, and adds, "If fraud were afoot, it would make sense that the president's men would steal votes in GOP strongholds…" (Freeman 2005a, 7). Be that as it may, WPE is larger (more negative) on average in *Democratic states* as measured by Kerry vote percentage.
- Truth Is All's pre-election predicted margin with undecideds allocated, TIAMARG2 (the same variable used to calculate TIASURP2).

If these variables are positively correlated with fraud, we would expect them to be negatively correlated with both vote/pre and vote/exit surprise. Table 2 indicates otherwise. In fact, African-American population (logged) is the only variable negatively correlated with both vote/pre and vote/exit surprise – and the vote/pre correlation falls well short of statistical significance. Paradoxically, every other variable is *positively* (but generally insignificantly) correlated with vote/pre surprise. This result indicates that Kerry did better than expected in the states assumed to be more subject to fraud! (We might argue that fraud could perfectly well have been concentrated in states where Kerry otherwise would have done best relative to pre-election expectations. This might be true, but there is no prior evidence for it. Surely it is far from obvious that Kerry would have increased his margins in New Hampshire, New York, and Vermont by double digits if concentrated fraud had not intervened.) Overall, the table offers

---

[19] The MYVOTE1 program used the telephone number 1-866-MYVOTE1; the EIRS used 1-866-OUR-VOTE.

[20] Empirically, the distributions of the indexes are skewed, and the logarithms of the two indexes are marginally more strongly correlated than the raw indexes.

[21] As with election incident calls, the causal theory here is far from clear. Freeman enumerates various forms of vote suppression, which are unlikely to contribute to the exit poll discrepancy because they would typically prevent people from either voting or participating in the exit poll. He then notes that exit poll error is correlated with African-American population proportion, "so one must at least consider the possibility that… blacks are being disenfranchised by yet other means as well" (Freeman 2005a, 8). One hypothesis would be that African-American voters experience higher rates of ballot spoilage (Klinkner 2001).

little support to theories of massive, widespread fraud. Nevertheless, some of the significant coefficients deserve further attention.

Table 2: bivariate correlations of surprise measures with systematic variables

		TIASURP2 TIA pro-Dem surprise	BGEO04 Best Geo error	WPE04 Mean WPE
REPGOV04 Republican governor	Pearson *r* *p*	0.095 *0.517*	-0.111 *0.448*	-0.157 *0.280*
SWING04 swing state (E/M list of 11)	Pearson *r* *p*	0.031 *0.830*	-0.080 *0.586*	-0.211 *0.146*
EASTTIME Eastern Standard Time	Pearson *r* *p*	0.069 *0.637*	**-0.282** *0.050*	-0.230 *0.112*
EVMPCT % of vote on electronic voting machines	Pearson *r* *p*	0.101 *0.491*	-0.014 *0.923*	-0.112 *0.442*
NEWENG New England	Pearson *r* *p*	0.192 *0.187*	-0.254 *0.079*	-0.273 *0.058*
MYV_VAP MYVOTE1 calls per million VAP	Pearson *r* *p*	0.116 *0.429*	**-0.304** *0.034*	**-0.329** *0.021*
EIRS_VAP EIRS election incidents per million VAP	Pearson *r* *p*	0.045 *0.761*	-0.054 *0.712*	-0.206 *0.156*
INCIDENT MSNBC/EIRS composite factor score	Pearson *r* *p*	0.123 *0.399*	**-0.283** *0.049*	**-0.426** *0.002*
LNAFRAM African-American pop (natural logarithm)	Pearson *r* *p*	-0.133 *0.361*	-0.245 *0.090*	**-0.303** *0.034*
DEMARG04 Dem vote margin	Pearson *r* *p*	**0.351** *0.013*	**-0.337** *0.018*	**-0.440** *0.002*
TIAMARG2 TIA pre-election predicted Dem margin #2	Pearson *r* *p*	0.094 *0.520*	**-0.359** *0.011*	**-0.329** *0.021*

Because the African-American population variable comes closest to fitting theoretical expectations, a plot of *unlogged %* African-American versus vote/pre surprise appears below as Figure 2 on page 17.[22] The plot suggests the generalization that Kerry did worse than anticipated in states of the Old South. By TIA's vote/pre surprise measure, Kerry ran an average 4.1 points worse than expected outside the South, but an average 8.2 points worse in the Southern states ($p < 0.01$). In fact, if we control for Southern versus non-Southern states, states with higher proportions of African-Americans tended to have insignificantly *higher* – rather than insignificantly lower – vote/pre surprises favoring Kerry. However, this result does not appear to evince fraud concentrated in the Southern states, because Southern states did not have

[22] The unlogged African-American measure happens to be more strongly correlated with vote/pre surprise than is the logged version, although neither correlation is statistically significant.

significantly different levels of vote/exit surprise.[23] Nor is the relationship between race and vote/exit surprise robust. Not only is the relationship (using the logged variable LNAFRAM) not statistically significant if Best Geo error rather than WPE is considered ($p = 0.09$, as reported in Table 2), but the t score of Best Geo error yields a result even further from significance ($r = -0.115$, $p = 0.43$). Fraud and/or other vote count error may disproportionately affect African-American voters, but – contrary to Freeman – correlational analysis at the state level does not support the conjecture.

Election "incidents" and surprise

The MYVOTE1 incident variable is rather strongly correlated with the vote/exit surprise measures, and the composite INCIDENT measure fares about as well or better. The EIRS variable has weaker correlations, conceivably because the EIRS toll-free line was less widely publicized. All three incident variables have insignificant and positive correlations with vote/pre surprise – that is, Kerry does *better*, relative to expectations, in states with higher incident rates. The unexpected outcome for vote/pre surprise infirms the theory because, as noted above, some election incidents seem likely to induce a stronger negative correlation with vote/pre surprise than with vote/exit surprise. (For instance, people who leave a long line without voting, or never join the line to begin with, are unlikely to be represented in an exit poll, but might well contribute to a deviation from pre-election expectations.) Still, one should not hasten to conclude that election incidents have no discernible impact on election outcomes, and so I subjected both EIRS and MSNBC incident reports to further analysis.

What sort of "election incident" should be correlated with vote/exit surprise? Observers have suggested two in particular: the malfunction and/or rigging of touchscreen DREs to favor George W. Bush, and people being forced to cast provisional ballots that later went uncounted. Other election incidents – undelivered absentee ballots, registration purges, and long lines – might disproportionately affect vote/pre surprise. (Note that absentee and registration problems both could contribute to provisional ballot problems.) Both EIRS and MYVOTE incidents are categorized by type, although the categorizations are imperfect and unwieldy.[24] I derived a

[23] The state average WPE was trivially more negative in Southern states (-6.5 versus -5.9); the average Best Geo error even more trivially negative (-5.2 versus -5.0); and the average t score of Best Geo error was *less* negative (-1.03 versus -1.26), partially reflecting smaller samples in many Southern states.

[24] The EIRS categories include "Absentee-ballot related problem," "Criminal status-related problem," "Disability access problem," "Early closing," "Identification-related problem," "Insufficient number of ballots," "Late opening," "Long lines," "Machine problem," "Non-English lang. assistance problem," "Polling place inquiry," "Problem based on student status," "Provisional ballot problem," "Registration-related problem," "Unable to read ballot," "Voter intimidation," "Other ballot-related problem," "Other polling place problem," and "Other." A call could be attributed to multiple categories. I limited my analysis to the following eleven categories, in order of frequency: registration-related (15,518); polling place inquiries (8,941); absentee (4,052); "other" (3,834); "other polling place" (3,421); machine problems (2,356); provisional ballot problems (2,076); voter intimidation (1,979); long lines (1,570); identification (1,417); and other ballot problems (1,090). The remaining categories were much less common. (One should not hasten to infer, for instance, that active voter intimidation was more common than long lines. These categories are rather amorphous, as in the case of incident number 051812 in Cuyahoga County, Ohio, classified as voter intimidation: "Republican challenger allowed daughter to help elderly mom vote. confrontation, police called." Regardless of categorization, many of the reports make distressing reading.

The MYVOTE1 system is a bit simpler, comprising nine complaint categories not including polling place inquiries. The nine categories, in approximate order of occurrence, are registration, absentee, poll access, ballot/screen, mechanical, coercion, identification, provisional, and other. Over 16,000 calls were manually coded variety of measures and indexes to examine various specifications of how different forms of reports predicted vote/pre and vote/exit surprise.

The most robust result is that, contrary to expectation under the fraud hypothesis, reports of absentee voting problems were the most strongly negatively correlated with vote/exit surprise (both WPE and Best Geo error).[25] The result is anomalous because most absentee voting problems would be "resolved" either by not voting or by casting provisional ballots that were counted, in neither case inducing vote miscount. However, extensive absentee voting can exacerbate both WPE and Best Geo error for reasons unrelated to fraud. Absentee votes contribute to WPE when they are tabulated at the precinct level, since they affect the official precinct results and cannot be reflected in the exit poll results for the precinct. E/M did supplement the exit poll with twelve state telephone polls of early and absentee voters, attempting to mitigate their effect on the statewide results. To the extent that these surveys were inaccurate, improperly weighted, or (in some states with large absentee voting) nonexistent, absentee voters also contributed to Best Geo error. It is unlikely that these influences primarily account for the correlation between reports of absentee voting problems and vote/exit surprise, but they may explain why this correlation is most robust.

As for machine/touch screen and provisional ballot problems, they do not emerge as especially strong predictors of vote/exit surprise. Indeed, none of the provisional-ballot measures is significantly correlated with any vote/exit surprise measure, nor is the EIRS "machine problem" measure. The MYVOTE1 coefficients for machine problems are more equivocal. On the "ballot/screen" measure, the autocoded measure is substantially negatively correlated with all three vote/exit surprise measures (on the order of $r = -0.4$) – but also verges on significant *positive* correlation with vote/pre surprise. (None of the coefficients approaches significant *negative* correlation with vote/pre surprise; if certain types of election incidents are more conducive to fraud, that effect is not apparent here.) In contrast to the autocoded version, the human-coded ballot/screen measure seems utterly unrelated to vote/exit surprise (all $|r| < 0.1$). The MYVOTE1 "mechanical" category, in both human-coded and autocoded variants, does yield significant or nearly significant negative correlations with the vote/exit measures. However, this category was probably quite expansive, especially in its autocoded variant. Consider this list of words and phrases that would lead to an autocoding as "mechanical": 'the machine,' 'machine,' 'jammed,' 'number of voting machines,' 'number of machines,' 'machines,' 'voting machine,' 'voting machines,' and 'voting booth' (Fels Institute 2005, 32).

It thus seems likely that incident reports primarily measure something, or some things, distinct from impacts upon the election outcome. In general, the various incident categories appear to tap a single dimension. In principal component analysis, a single component accounted for 53% of the variance in the human-coded MYVOTE1 reports (or more specifically in the logarithms of their rates per million VAP); 78% of the variance in the autocoded MYVOTE1 reports; and 81%

---

by humans; over 56,000 calls were autocoded after the fact by "audio mining analysis" of the recordings. The analyses agreed that the first three categories above were most common; however, the autocoding indicated that 13% of complaints referred to "ballot/screen" issues, whereas human coders set this figure at only 4%. This disparity is unexplained, although a contributing factor may have been that the word "electronic" or the phrase "sample ballot" were indicators to categorize a complaint as ballot/screen related (Fels Institute 2005, 32).

[25] The largest observed correlation is between WPE and the logarithm of MYVOTE1 human-coded absentee problems per million VAP: $r = -0.440$.

of the variance in the EIRS reports excluding place inquiries (or 79% including place inquiries). For whatever reason, incident reports were more common (on average) in "blue" states. Kerry's prior predicted margin is substantially positively correlated with the incident index ($r = 0.36$ using TIA's predicted margin, or 0.39 using E/M's). Prior predicted margin is likewise substantially negatively correlated with vote/exit surprise: that is, exit poll error overstating Kerry's lead was larger (more negative) in the states where he was expected to do best. Once we control for prior predicted margin, the summary INCIDENT variable is no longer significantly correlated with Best Geo error (beta = -0.211, $p = 0.16$). INCIDENT remains significantly correlated with WPE (beta = -0.341, $p < 0.02$).

These results lend themselves to two unproven conjectures. First, voters in Democratic states (possibly but not necessarily Democratic voters) may have been more inclined to phone in complaints about election incidents, independent of any propensity for these incidents to disproportionately affect Democratic voters. (Nothing in the data allows us to conclude that election incidents *did not* disproportionately affect Democratic voters, although it does appear that Kerry did not consistently do worse than expected in states with high incident rates.) It would not be surprising were concerns about election fraud and/or awareness of the election incident lines higher among Democratic voters.[26] Second, chaotic conditions that triggered election incident calls may also have exacerbated exit poll error, while not systematically benefiting either candidate.

Surprise and expectations

The one variable in this analysis that is *significantly* correlated with all three surprise variables is Democratic vote margin. The correlation with vote/pre surprise obviously is partially spurious: loosely speaking, Kerry did better (DEMARG04) in states where he did better than expected (TIASURP2). However, this result suggests the possibility that prior *expected* margin might account – statistically, at least – for much of the odd negative correlation between vote/pre and vote/exit surprise measures. To check this hypothesis, I regressed WPE, Best Geo, and *t* score of Best Geo on two sets of prior expectations and surprise measures: TIA's measures (TIAMARG2 and TIASURP2), and the equivalent E/M measures. As you can see in Table 3 below, the prior margin has little effect on the coefficient of TIA surprise (compare Table 1), but substantially attenuates the coefficients of E/M surprise.

Table 3: prior expectations and vote/exit surprise

Independent variables	**-------------------- Dependent variable --------------------**		
	Best Geo	*t* score of Best Geo	WPE
TIA expected	-0.251 ($p = 0.07$)	-0.289 ($p = 0.03$)	-0.325 ($p = 0.01$)
TIA surprise	-0.269 ($p = 0.06$)	-0.386 ($p < 0.01$)	-0.358 ($p < 0.01$)
E/M expected	-0.262 ($p = 0.10$)	-0.312 ($p = 0.04$)	-0.335 ($p = 0.03$)
E/M surprise	-0.134 ($p = 0.39$)	-0.181 ($p = 0.23$)	-0.186 ($p = 0.21$)
E/M surprise (only)	-0.249 ($p = 0.08$)	-0.318 ($p = 0.03$)	-0.333 ($p = 0.02$)

* Coefficients are betas from multivariate OLS regression (*p* values in parentheses).

[26] In particular, awareness of the EIRS hotline would be higher among Democratic voters to the extent that the Election Protection Organizations that promoted it concentrated their efforts in Democratic precincts.

Conclusion

Overall, the variation between the pre-election measure and the actual returns looks much like random noise. To be sure, the vote/pre surprise measures favored by TIA differ from random noise in a crucial respect: their means are well below zero, supposedly indicating massive fraud. But this purported systematic fraud does not seem to fit any pattern that has been proposed so far. If anything, Kerry tended to fare better relative to prior expectations (not necessarily better *than* prior expectations, but at least less worse) in states where the exit polls supposedly pointed to fraud. I certainly am not in a position to rule out massive fraud favoring Bush in Vermont or New York, but I see little evidence for it. Bear in mind that the conjecture of widespread, massive fraud requires millions – perhaps ten millions – of votes to be shifted from Kerry to Bush.

Contrary to Freeman (2005a), the state-level variation between exit polls and actual returns also looks a great deal like noise. Some of Freeman's variables of interest are not statistically significant predictors of "Precinct Level Disparity" in this 49-state dataset (Republican governor; swing state as a state dummy; electronic voting as a statewide percentage). The incident measure and proportion of African-American voters are indeed significantly correlated with vote/exit error. However, upon close inspection, the hypotheses supporting these correlations are underdeveloped, and other variables seem likely to account for the correlations.

The search for a "smoking gun" in the exit poll data increasingly resembles the proverbial pursuit of wild geese. As such, it is not without interest – and it may yet generate fruitful hypotheses – but at this point it seems to contribute little to our understanding of possible electoral fraud. One might argue, reasonably enough, that state-level analysis was never a promising avenue of inquiry. Yet many confident declarations that the 2004 election was stolen have depended on state-level data. One wonders how long these declarations can persist.

References

Campbell, James E. "Why Bush Won the Presidential Election of 2004: Incumbency, Ideology, Terrorism, and Turnout." *Political Studies Quarterly* 120, 2 (Summer 2005): 219-241.

Edison Media Research and Mitofsky International (E/M). 2005. "Evaluation of Edison/Mitofsky Election System 2004." January 19, 2005. Published on the Internet at http://exit-poll.net/election-night/EvaluationJan192005.pdf (last accessed 11/7/2005).

Fitrakis, Bob, and Harvey Wasserman. 2005. "Powerful Government Accountability Office report confirms key 2004 stolen election findings." October 26, 2005. Published on the Internet at http://www.freepress.org/departments/display/19/2005/1529 (last accessed 11/7/2005).

Fels Institute of Government. 2005. "The MyVote1 Project Final Report (2004-04)." November 1, 2005. Published on the Internet at http://www.sas.upenn.edu/fels/FGRS/Final%20MyVote1%20Report%2011.1.05.pdf (last accessed 11/21/2005).

Freeman, Steve. 2005a. "Polling Bias or Corrupted Count? Accepted Improbabilities and Neglected Correlations in 2004 US Presidential Exit Poll Data." Presentation to the Fall Meeting of the American Statistical Association, Philadelphia Chapter, October 14, 2005. Published on the Internet at http://www.appliedresearch.us/sf/Documents/ASAP-Improbabilities%20&%20Neglected%20Correlations%20051014%20text.pdf (last accessed 11/9/05).

Freeman, Steve. 2005b. "The Election Outcome Based on Exit Poll Reported Voting." Handout accompanying presentation at the Fall Meeting of the American Statistical Association, Philadelphia Chapter, October 14, 2005. Published on the Internet at http://www.appliedresearch.us/sf/Documents/Election%20Outcome%20based%20on%20EP%20Respondent%20Reported%20Voting.pdf (last accessed 11/7/2005).

Klinkner, Philip A. 2001. "Whose Votes Don't Count?: An Analysis of Spoiled Ballots in the 2000 Florida Election." Published on the Internet at http://www.hamilton.edu/news/florida/Klinkner%20Analysis.pdf (last accessed 12/27/05).

Kyle, Susan, Douglas A. Samuelson, Fritz Scheuren, and Nicole Vicinanza with Scott Dingman and Warren Mitofsky. 2005. "ESI Brief – Analysis of the 2004 Ohio Exit Polls and Election Results." June 6, 2005. Published on the Internet at http://electionscience.org/Members/stevenhertzberg/report.2005-07-19.7420722886/report_contents_file/ (last accessed 11/7/2005).

Liddle, Elizabeth. 2005. "Edison/Mitofsky Exit Polls 2004: differential non-response or vote-count corruption?" April 27, 2005. Published on the Internet, http://www.geocities.com/lizzielid/WPEpaper.pdf (last accessed 12/27/05).

Lindeman, Mark. 2005. "Bias and swing in the 2004 NEP exit polls: do the exit polls indicate fraud?" Third edition, November 3, 2005 (original 10/22/05). Published on the Internet, http://inside.bard.edu/~lindeman/slides.html (last accessed 12/27/05).

Mebane, Walter R., and Michael C. Herron. 2005. "Ohio 2004 Election: Turnout, Residual Votes and Votes in Precincts and Wards." June 9, 2005. Section 6 (pages 47-145) of Democratic National Committee Voting Rights Institute, *Democracy at Risk: The 2004 Election in Ohio.* Published on the Internet, http://a9.g.akamai.net/7/9/8082/v001/www.democrats.org/pdfs/ohvrireport/section06.pdf (last accessed 12/27/05). Entire report available at http://a9.g.akamai.net/7/9/8082/v001/www.democrats.org/pdfs/ohvrireport/fullreport.pdf.

Pew Research Center for the People and the Press. 2004. "Pre-Election Polls Largely Accurate." November 23, 2004. Published on the Internet at http://people-press.org/commentary/display.php3?AnalysisID=102 (accessed 10/24/05).

TruthIsAll. 2004a. "The Election Model." http://www.geocities.com/electionmodel/ Dated November 1, 2004 (accessed 10/23/05). [The page has been updated to add links at the top of the page, but the main text of the page appears to be unchanged.]

TruthIsAll. 2004b. "The Election Model [detailed polling and analytic reports]." [This page does not have a unique name, but the description in brackets appears on the main page – TruthIsAll 2004a – as part of the hyperlink to this page.] Dated 1 November (2004). http://www.geocities.com/electionmodel/index_files/ElectionModel.htm (accessed 10/23/05).

TruthIsAll. 2005a. "The EAST was the BEAST: Time Zone Exit Poll Discrepancy Analysis." Posted on Democratic Underground, February 5, 2005, http://www.democraticunderground.com/discuss/duboard.php?az=show_mesg&forum=203&topic_id=318693&mesg_id=318693 (accessed 10/23/05).

TruthIsAll. 2005b. "Implications." Posted on Democratic Underground, February 5, 2005, http://www.democraticunderground.com/discuss/duboard.php?az=show_mesg&forum=203&topic_id=318693&mesg_id=318710 (accessed 10/23/05).

TruthIsAll. 2005c. Interactive Election Simulation. Microsoft Excel spreadsheet. Published on the Internet at http://www.geocities.com/electionmodel/InteractiveElectionSimulation.xls (accessed 10/23/05).

United States Elections Project. 2005. "2004 Voting-Age and Voting-Eligible Population Estimates and Voter Turnout." Published on the Internet at http://elections.gmu.edu/Voter_Turnout_2004.htm (accessed 10/24/05).

Figure 2: vote/pre error (TIASURP2) as a function of African-American population

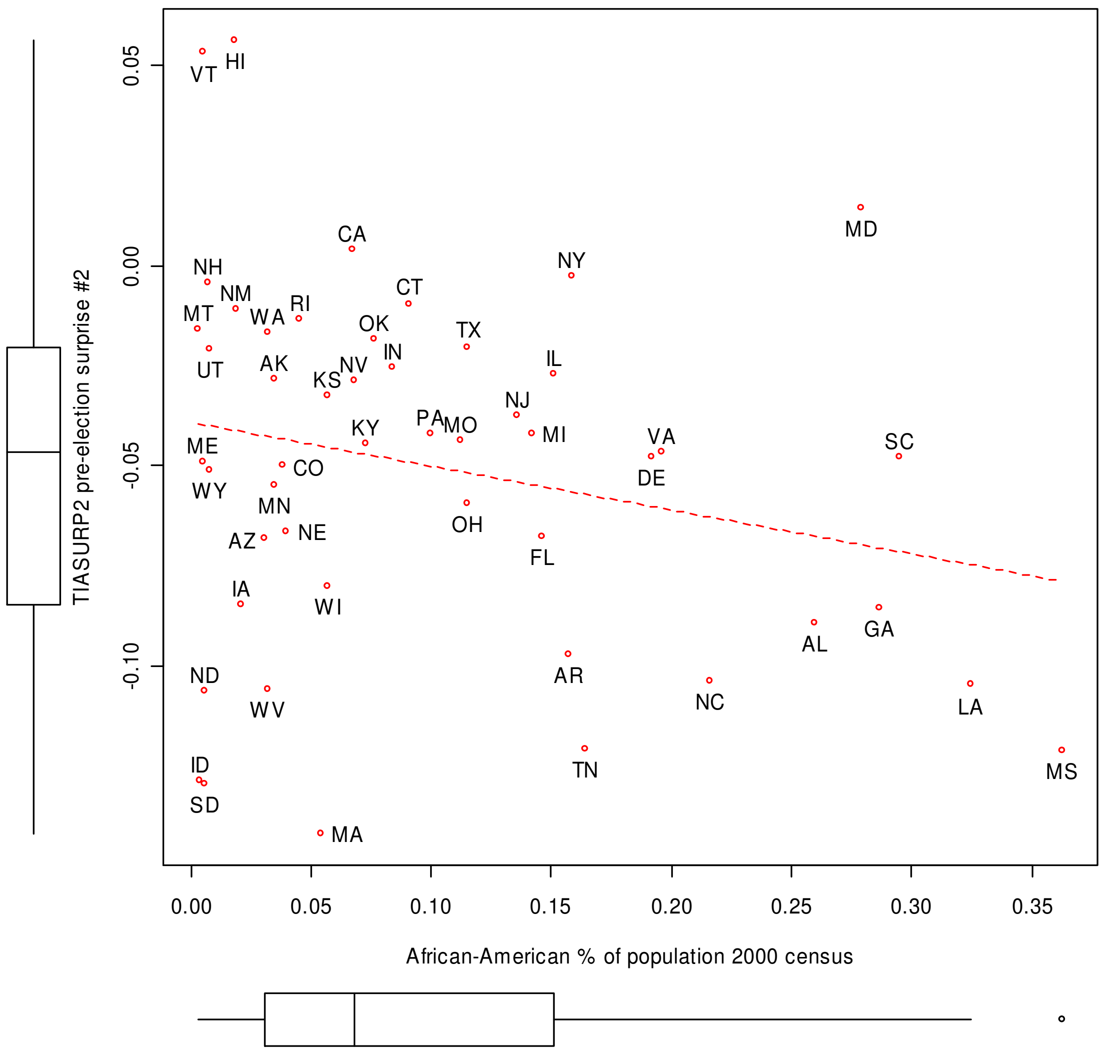


Figure 3: vote/exit error (WPE04) as a function of vote/pre error (TIASURP2)