



**Graduate Division, School of Arts & Sciences
Center for Organizational Dynamics
Working Paper #04-09 November 15, 2004**

The Unexplained Exit Poll Discrepancy

Steven F. Freeman, PhD

stfreema@sas.upenn.edu

Most Americans who had listened to radio or surfed the Internet on Election Day this year, sat down to watch election night coverage expecting that John Kerry had been elected President. Exit polls showed him ahead in nearly every battleground state, in many cases by sizable margins. Although pre-election day polls indicated the race dead even or Bush slightly ahead, two factors seemed to explain Kerry's edge: turnout was extraordinary high, which is generally good for Democrats¹, and as in every US Presidential election with an incumbent over the past quarter-century², undecided voters broke heavily toward the challenger.³

But then, in key state after key state, counts showed very different numbers than the polls predicted; and the differentials were all in the same direction. The first shaded column in Table 1 shows the differential between the major candidates' predicted (exit poll) percentages of the vote; the next shaded column shows the differential between their *tallied* percentages of the vote. The

¹ Democratic party candidates generally benefit from higher turnout because the increase comes disproportionately from voters in lower socio-economic groups that traditionally vote Democratic, q.v., Frank Newport, "1998 midterm elections: a mixed picture" *The Gallup Poll Monthly*, Oct 1998 i397 p5(3)

² Guy Molyneux, "The Big Five-Oh", *The American Prospect Online*, Oct 1, 2004: There have been four incumbent presidential elections in the past quarter-century. On average, the incumbent comes in half a point below his final poll result; challengers exceed their final poll result by an average of 4 points.

³ Even the final "corrected" exit poll data presented on the CNN website – more on such corrected data later – indicate that those who decided in the last three days chose Kerry over Bush 55% - 42%.
<<http://www.cnn.com/ELECTION/2004/pages/results/states/US/P/00/epolls.0.html>> (Thursday Nov 14, 2004)

final shaded column reveals the “shift.” In ten of the eleven consensus battleground states⁴, the tallied margin differs from the predicted margin, and in every one, the shift favors Bush.

Table 1⁵: Predicted vs. Actual percentages in battleground states

	Sample size	Bush predicted	Kerry predicted	Predicted differential	Bush tallied	Kerry tallied	Tallied differential	Tallied vs. predicted
Colorado	2515	49.9%	48.1%	Bush 1.8	52.0%	46.8%	Bush 5.2	Bush 3.4
Florida	2846	49.8%	49.7%	Bush 0.1	52.1%	47.1%	Bush 5.0	Bush 4.9
Iowa	2502	48.4%	49.7%	Kerry 1.3	50.1%	49.2%	Bush 0.9	Bush 2.2
Michigan	2452	46.5%	51.5%	Kerry 5.0	47.8%	51.2%	Kerry 3.4	Bush 1.6
Minnesota	2178	44.5%	53.5%	Kerry 9.0	47.6%	51.1%	Kerry 3.5	Bush 5.5
Nevada	2116	47.9%	49.2%	Kerry 1.3	50.5%	47.9%	Bush 2.6	Bush 3.9
New Hampshire	1849	44.1%	54.9%	Kerry 10.8	49.0%	50.3%	Kerry 1.3	Bush 9.5
New Mexico	1951	47.5%	50.1%	Kerry 2.6	50.0%	48.9%	Bush 1.1	Bush 3.7
Ohio	1963	47.9%	52.1%	Kerry 4.2	51.0%	48.5%	Bush 2.5	Bush 6.7
Pennsylvania	1930	45.4%	54.1%	Kerry 8.7	48.6%	50.8%	Kerry 2.2	Bush 6.5
Wisconsin	2223	48.8%	49.2%	Kerry 0.4	49.4%	49.8%	Kerry 0.4	No dif

The media has largely ignored this discrepancy (although the Blogosphere has been abuzz), suggesting that the polls were flawed, within normal sampling error, a statistical anomaly, or could otherwise be easily explained away. In this paper, I examine the validity of exit polls, the likelihood sampling error, the possibility of statistical anomaly, and other explanations thus far offered.

Data

The data I use for this paper are those posted on the CNN website Election Night. CNN had the data by virtue of membership in the National Election Pool (NEP), a consortium of news organizations that had pooled resources to conduct a large-scale exit poll (as was done in the 2000 election). NEP, in turn, had contracted two respected firms, Edison Media Research and Mitofsky International⁶, to conduct the polls.

Calibrated and Uncalibrated Exit Poll Data

Part of the reason why the issue went away for the media – and has become fodder for conspiracy theorists on the web – is secrecy and confusion about the data and what exactly is

⁴ These eleven states are classified as battleground states based on being on at least two of three prominent lists: *Zogbys*, *MSNBC*, and the *Washington Post*.

⁵ Source: CNN website. Wednesday November 3, 2004 12:21 am.

⁶ Warren Mitofsky, the founder of Mitofsky International is credited with having invented the exit poll. David W. Moore, Senior Gallup Poll Editor, “New Exit Poll Consortium Vindication for Exit Poll Inventor” 10/11/2003

being characterized as the exit poll. If you go to the CNN website or any other website on which 2004 exit poll data is available, you'll see numbers very different from those that those that were released on Election Day. That's because the survey results originally collected and presented to subscribers were subsequently "corrected" to conform to official tallies.

The pollsters explain this as a natural process: the "uncalibrated" data were preliminary; once the counts come in, they recalibrate their original data on the assumptions that the count is correct, and that any discrepancies must have been due to imbalanced representation in their samples or some other polling error. The pollsters have taken great pains to argue that their polls were not designed to verify election results⁷, but rather to provide election coverage support to subscribers: as one piece of data (among many) that networks could use to "call" states and to explain voting patterns, i.e., who voted for whom, and why people voted as they did.

Whatever the merits of calibrating exit poll data, it confuses the issue about *why* the (uncalibrated) polls were so far off and why always in the same direction. Worse, it suggests that they are not important questions. Although this calibration process may seem perfectly natural to NEP exit-pollsters, it confuses nearly everyone else, even sophisticated analysts intimately involved in voting issues. The MIT-Caltech Voting Project, for example, issued a report concluding that exit poll data were consistent with state tallies and that there were no discrepancies based on voting method, including electronic voting systems. But they used this adjusted data to exonerate the process! In other words, they used data in which the count is assumed correct to prove that the count is correct, and the report has been used to dismiss allegations that anything is awry.⁸

If this issue ever comes to light, those who produced these reports will be embarrassed by them, but it's understandable how one could make that mistake. Few of us realized that this data is corrected. Neither the CNN website, nor any other site of which I'm aware, gives any

⁷ Martin Plissner, "In Defense of Exit Polls: You just don't know how to use them" *Slate* Thursday, Nov. 4, 2004

⁸ Tom Zeller, Jr. "Vote Fraud Theories, Spread by Blogs, Are Quickly Buried" *New York Times* November 12, 2004 (Front page). MSNBC's Keith Olbermann Countdown program November 12, 2004

indication that the data were anything other than what nearly all of us imagine exit poll data to be, which are data based (solely) on subjects surveyed leaving the polling place.

Data Used in This Report

For this report, I use data that apparently are based solely on subjects surveyed leaving the polling place. This data were reportedly not meant to be released directly to the public,⁹ and were reportedly available to late evening Election Night viewers only because a computer glitch prevented NEP from making updates sometime around 8:30 p.m. that night.¹⁰ They were collected by Jonathon Simon, a former political survey research analyst, who corroborates his figures with saved screen shots (see Figure 1). I happened to have ten exit poll internet pages

Figure 1. CNN web page with apparently “uncorrected” exit poll data (12:21 am Wed, Nov. 3, 2004)



9 Martin Plissner, "In Defense of Exit Polls: You just don't know how to use them" *Slate* Thursday, Nov. 4, 2004
 10 Richard Morin, "New Woes Surface in Use of Estimates" *Washington Post*, Thurs, Nov. 4, 2004; Page A29

stored in my computer memory, and in each case, his figures match mine exactly. The numbers are also roughly consistent with those released elsewhere (Appendix C shows *Slate* numbers at 7:28 EST).

To derive the “predicted values” used in Tables 1 & 2, I combining the male and female vote, weighted for their percentage of the electorate. Ohio exit poll data (Figure 1) indicates that 51% of men and 53% of women voted for Kerry. Since the electorate is 47% male/53% female, Kerry’s overall share of the exit poll in was calculated as $(51\% \times 47\%) + (53\% \times 53\%)$ or 52.1%¹¹. Doing the same for calculations for other battleground states and comparing these numbers with final tallies (NY Times, Nov. 7), I completed the columns in Tables 1 & 2.

Some have rejected these data as unweighted, meaning that they have not been adjusted to appropriately weight demographic groups pollsters knowingly under- or over-sampled.¹² But that would seem very unlikely. NEP’s predecessor, Voter News Service warns in bold letters in its 2000 Methodology statement never to use unweighted data (see Appendix A). It makes no sense *ever* to distribute unweighted data to anyone. Pollsters want to get it right. Their customers are depending on it. Broadcasters want to be alerted to probable outcomes, so as to plan their coverage accordingly (e.g., pre-writing stories so they could be completed shortly after poll-closings, assigning top reporters to winners’ headquarters, being prepared for when concession or victory speeches might be forthcoming, etc...). In this case, subscribers were very surprised. Editors and network managers had to scramble, and journalists complained that they had to rewrite their lead stories.

It’s likewise possible that the data was already partially calibrated to tallies, but given the news story and the abrupt change at 1:30 am, that too seems unlikely. If, in fact, the data already

¹¹ Among the limitations of the CNN exit poll data is the lack of significant digits. Rounding errors mean that exit poll numbers for individual state analyses could be off by up to .5. This is unlikely because it comes from two groups, male and female, and it’s unlikely that they are both rounded that much in the same direction. Regardless, the strength of the finding is such that even if all numbers had been rounded the full .5 in an unfavorable direction, the basic finding would still hold.

¹² Pollsters typically oversample minorities so that they have a sufficient sample size of important demographic groups, but then they negatively weight respondents in these groups to adjust for their actual percentage of the electorate.

had been partially calibrated, however, it would mean that the pure exit poll numbers favored Kerry to an even greater extent.

On (Uncorrected) Exit Polls

In general, we have reason to believe that exit polls, by which I mean *uncorrected* exit polls, are accurate survey instruments. Exit polls are surveys taken of representative respondents from the overall voting population. Both the logic behind them and experience suggest that we can use these surveys to predict overall results with high degrees of certainty. It's easy to get a statistically valid representative sample; and there is no problem with figuring out who is going to actually vote -- or how they will vote.

In Germany, the minute the polls close, polling agencies release prognoses that are reportedly extremely reliable.¹³ In the most recent national elections, final results diverged from official counts by only 0.7% on average (Appendix B).

In the US, students at BYU have been conducting Utah exit polls since 1982.¹⁴ They write:

... results are very precise; In the 2003 Salt Lake [City] mayoral race, the KBYU/Utah Colleges Exit Poll predicted 53.8 percent of the vote for Rocky Anderson and 46.2 percent for Frank Pignanelli. In the actual vote, Anderson carried 54 percent of the vote to Pignanelli's 46 percent.

True to their word, predictions in this year's contests were quite accurate. In the Utah presidential vote, for example, they predicted Bush 70.8%, Kerry 26.5%. The actual was Bush 71.1%, Kerry 26.4%. Consistently accurate exit poll predictions from student volunteers, including in this presidential election, give us good reason to expect accuracy from the world's most professional exit polling enterprise.

Several political commentators have remarked on the reliability of exit polls. Dick Morris, Republican consultant and Fox News regular, writes:

13 Thom Hartmann, "The Ultimate Felony Against Democracy" *CommonDreams.org*. Thu, Nov. 4 2004

14 <http://exitpoll.byu.edu/2004results.asp>. As far as I have been able to determine, this was the only other exit poll conducted on the 2004 presidential election. I had thought that Zogby also had an exit poll, but haven't been able to verify this; they may have been using the same National Election Pool data, when they declared Kerry the winner in Ohio on Election Night. See William Douglas, "Early exit polls come under fire" *The Mercury News* (11/3/2004)

Exit polls are almost never wrong ... So reliable are the surveys that actually tap voters as they leave the polling places that they are used as guides to the relative honesty of elections in Third World countries. When I worked on Vicente Fox's campaign in Mexico, for example, I was so fearful that the governing PRI would steal the election that I had the campaign commission two U.S. firms to conduct exit polls to be released immediately after the polls closed to foreclose the possibility of finagling with the returns.¹⁵

Last fall, international foundations sponsored an exit poll in the former Soviet Republic of Georgia during a parliamentary election. On Election Day, the pollsters projected a victory for the main opposition party. When the sitting government announced that its own slate of candidates had won, supporters of the opposition stormed the Parliament, and the president, Eduard A. Shevardnadze, resigned under pressure from the United States and Russia.¹⁶

Statistical Analysis of the Three Critical Battleground States

Three Critical Battleground States

The conventional wisdom going into the election was that three critical states would likely determine who would win the Presidential election -- Ohio, Pennsylvania, and Florida. Sure enough, they did, with Bush winning two of three and ascending to electoral victory. In each of these states, however, exit polls differed considerably from recorded tallies (Table 2).

Table 2: Predicted vs. tallied percentages in the three critical battleground states

	Sample size	Bush predicted	Kerry predicted	Predicted differential	Bush tallied	Kerry tallied	Tallied differential	Tallied vs. predicted
Florida	2846	49.8%	49.7%	Bush 0.1 ¹⁷	52.1%	47.1%	Bush 5.0	Bush 4.9
Ohio	1963	47.9%	52.1%	Kerry 4.2	51.0%	48.5%	Bush 2.5	Bush 6.7
Pennsylvania	1930	45.4%	54.1%	Kerry 8.7	48.6%	50.8%	Kerry 2.2	Bush 6.5

A Statistical Anomaly?

Perhaps the first question to ask on looking at such a discrepancy is whether it is just a statistical anomaly. It can happen, for example, that a fair coin tossed ten times will land heads each time, but it doesn't happen often (1 out of 256 times to be precise). If we witness this, we

¹⁵ Dick Morris, "Those faulty exit polls were sabotage" *The Hill* Nov. 4, 2004
<http://www.thehill.com/morris/110404.aspx>

¹⁶ Martin Plissner, Exit Polls to Protect the Vote, *New York Times* 10/17/04

¹⁷ Earlier exit polls, including one released by Slate at 7:28 EST, 28 minutes after the Florida polls closed showed Kerry leading 50% to 49% (Appendix C)

will at least suspect that the coin might be adulterated, especially if the stakes are high, and we are not permitted to carefully inspect it.

Statistical significance, which means that the discrepancy is such that it is unlikely to occur by chance, depends on four factors – the size of the discrepancy, the sample size, sample characteristics, and the level of significance (just how unlikely does it have to be?).¹⁸ The NEP state samples are for statistical purposes, quite large. Two thousand or so respondents is roughly the size of most national polls. Without access to the data and methodology, I cannot model the sample characteristics precisely. Nevertheless, we do know the general procedures by which exit polls are conducted; based on these we can make a reasonable approximation.

In the analysis to follow, I assume a normal distribution curve. Exit polls are not random samples, but normality is nevertheless probably a reasonable approximation because of two counterbalancing processes – clustering and stratification. To avoid prohibitive expense, exit poll samples are clustered, which means that precincts (rather than individuals) are randomly selected. This increases variance and thus the margin of error because of the possibility that precinct voters share similar characteristics that differentiate them from the rest of the state in ways that past voting behavior would not predict.

Stratification, on the other hand, decreases variance. Identifying voters by key characteristics that predict voting behavior (race, sex, age, income, ethnicity, religion, party affiliation, etc...) ensures that the sample is representative of the overall population, either by seeking out subjects with specific demographic characteristics and/or weighting groups depending on their representation in the sample compared with that of the overall voting population. By getting samples in which minorities are over-represented (but subsequently negatively weighted), pollsters can ensure adequate sample sizes of each of these representative subgroups. Knowing exactly how much to weight over- or under-represented population depends on an accurate

¹⁸ (Apologies to those who are well versed in statistical inference. Most readers of this paper are not, so I provide much more explanation than I would for a purely academic reader.)

knowledge of overall demographics of the electorate. But historical data, census data, and registration roles, can be used to compliment sampling site counts to allow accurate weighting.

Appendix A describes how the 2000 Presidential election exit polls were conducted. Depending on exactly how states were clustered and stratified in the 2004 election, the findings that I present may be stronger or weaker but, as is the case with the question of rounding, the strength of the finding is such that any modification of this normality assumption would not change the basic conclusions.

Figure 2. Statistical prediction of Kerry's true percentage of the vote in Ohio

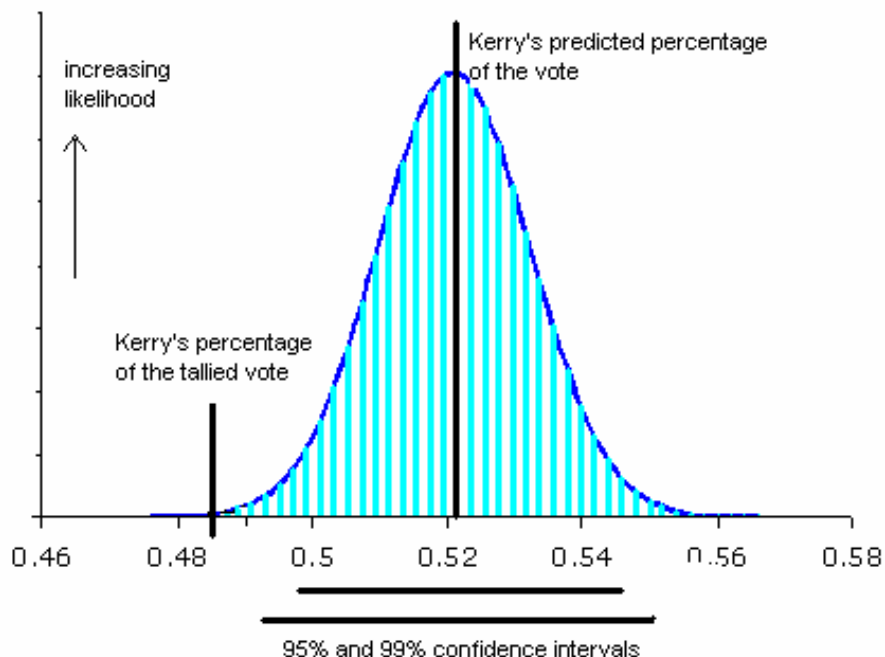


Figure 2 depicts the probability density of a normal distribution curve, showing the relative likelihood, given the Ohio poll data, of the actual percentage of the vote Kerry would be expected to receive in Ohio. The horizontal lines below the curve indicate the poll's statistical margin of error, the corresponding zone of 95 and 99 percent confidence.¹⁹ Given a similar poll indicating

¹⁹ To determine the margin of error, calculate the standard error using the formula $\sqrt{\frac{p(1-p)}{N}}$ where p = Kerry's poll percentage and N is the sample size. Sixty-eight percent of the time the true percentage of the population will be within one standard error of a random sample (.01137 or 1.14% in this case). Ninety-five percent of the time, it will be within two standard errors (2.27% in this case). This range within which we expect to find the true percentage is also called the 95% confidence interval. The 99% confidence interval is 2.58 standard errors on either side of the estimate (+/- 2.93% in this case).

Candidate Q received 52.1% of the vote, 95 percent of the time we would expect that the true percentage Q received was between 49.8% and 54.4%. And because half of the 1 in 20 cases that fall outside the interval would be high rather than low, 97.5% of the time we would expect that the true percentage Q received was at least 49.8%. Given this exit poll result, 99.5% of the time we would expect that the true percentage Q received was at least 49.2%. It turns out that the likelihood that Q would have received only 48.5% of the vote is less than one in one thousand (.0008).

Conducting the same analysis for Florida, we find that Kerry's 47.1% of the vote is likewise outside the 99% confidence interval. The likelihood of his receiving only 47.1%, given that an exit poll with 2846 respondents indicated he obtained 49.7%, is less than three in one thousand (.0028). Kerry's count is also outside the 99% confidence interval in the third critical battleground state, Pennsylvania. Although he did carry the state, the likelihood of his receiving only 50.8% given that the exit polls indicated 54.1% is less than two in one thousand (.0018).

The likelihood of any two of these statistical anomalies occurring together is on the order of two-in-a-million. The odds against all three occurring together are 250 million to one. As much as we can say in social science that something is impossible, it is impossible that the discrepancies between predicted and actual vote counts in the three critical battleground states of the 2004 election could have been due to chance or random error.

Official Explanations

The New York Times tells us that they obtained a report issued by the pollsters that debunked the possibility that the exit polls are right and the vote count is wrong²⁰, but does not explain beyond that declaration how the possibility was "debunked." In fact, no evidence at all is presented of skewed survey data or any problem at all with the polls except that "uncorrected" data were released to the public. *Slate* reports that Mitofsky and Lenski insist that the polls were

20 Jim Rutenberg, "Report Says Problems Led to Skewed Surveying Data" *New York Times*, Nov. 5, 2004

perfectly fine.^{21 22} Mitofsky himself only says, “We are still investigating the source of the exit poll problem.”²³

One of the few people close to the pollsters to offer an explanation early on was Martin Plissner, former executive political director of CBS News (and self-described close friend of the pollsters), who identifies three problems with the polls:

The pollsters who work outside the polling stations often have problems with officials who want to limit access to voters. Unless the interviews have sampled the entire day's voters, the results can be demographically and hence politically skewed. Finally, it is of course a poll, not a set of actual recorded votes like those in the precinct samples collected after the polls close.²⁴

Regarding the first problem, Jack Shafer of *Slate* observes:

... exit pollsters always encounter overzealous election officials enforcing electioneering laws. Can we really believe that a significant number of the 1,480 exit poll precincts in 50 states and the District of Columbia that Edison/Mitofsky surveyed on Election Day were so affected? And in sufficient numbers to bend state-by-state exit polls in Kerry's favor?²⁵

If in fact, pollsters have a limited sample of voters from a given precinct, they can weight the voters who are polled from that precinct or demographically similar precincts accordingly.

Regarding time of day variation, this paper does not refer to mid-day reports, but rather end-of-day data, which happened to still be available at midnight. But even if there were an early voter bias in the CNN data that I have used, is there any reason to believe that early votes would be skewed Democrat?

Regarding the last ditch argument that it's just a poll, its true (of course), but, as I have documented, the evidence and logic on exit polls suggest that we have reason to believe they are accurate within statistical limits.

21 Martin Plissner “In Defense of Exit Polls: You just don't know how to use them. *Slate* Thursday, Nov. 4, 2004

22 Jack Shafer, “The official excuses for the bad exit poll numbers don't cut it.” *Slate* Friday, Nov. 5, 2004, 9:23 PM PT

23 Warren Mitofsky, personal communication, November 10, 2004

24 Martin Plissner “In Defense of Exit Polls: You just don't know how to use them. *Slate* Thursday, Nov. 4, 2004

25 Jack Shafer, “The official excuses for the bad exit poll numbers don't cut it.” *Slate* Friday, Nov. 5, 2004, 9:23 PM PT

Under-representation?

Other explanations put forth by the *Washington Post* charge that samples may have included too many women, too few Westerners, not enough Republicans, etc ...” Regarding the first part of this critique, Morris writes:

The very first thing a pollster does is weight or quota for gender. Once the female vote reaches 52 percent of the sample, one either refuses additional female respondents or weights down the ones one subsequently counted. This is, dear Watson, elementary.

Moreover, the issue of male/female ratio is irrelevant. CNN and others released data presenting male and female preferences separately, thus automatically weighting sex appropriately.

Other potential imbalances are part of normal sampling error. A random sample would result in the poll precision and confidence intervals that I reported. Under such conditions, Republicans, westerners, etc., are equally (un)likely to be over- or under-represented. Imprecise representation is incorporated within the margin of error. That’s why we have the concept of probability densities, margin of error, etc.... If you could choose a perfectly representative sample, you could predict outcomes precisely. Techniques to ensure sample representativeness make polls more accurate than a purely random sample, thus making the observed discrepancies even more unlikely than my analysis indicates.

Bush voter unwillingness to participate and other “explanations”

More recently, Senior *Gallup Poll* Editor David W. Moore, report that Mitofsky and Lenski say that,

Kerry voters apparently were much more willing to participate in the exit poll than were Bush voters. The interviewers at each of the sample voting locations are instructed to choose voters as t=] exit from the voting booth -- every third, or fifth, or eighth voter -- some sequence of this sort that is determined ahead of time. Many voters simply refuse to participate in the poll. If the refusers are disproportionately for one candidate or another, then the poll will be biased....²⁶

The first problem with this explanation is that the pollsters make adjustments for voters who were missed or refused to be interviewed, based on their observed age, race, and sex (see Appendix A). But, OK, it is true that if Bush voters of all stripes disproportionately refuse to

26 David W. Moore, Senior Gallup Poll Editor, “Conspiracies Galore” *Gallup News Service*: November 9, 2004.

participate, that *could* explain exit poll error. The question is whether we have evidence to support this.

It is *conceivable* that Kerry voters were much more willing to participate in the exit poll than were Bush voters, but although it's not difficult to imagine why some Bush voters might not want to participate, it's also not difficult to imagine why some Kerry voters might not want to participate either. Likewise for the proposed early voter bias. Although it is possible to imagine Republicans not voting early ("they have to work"), it is equally possible to imagine orderly Republicans rising with the sun and doing the civic duty, while sinful Democrats sleep late and procrastinate until near poll closing time.

The problem with these explanations or even one that might have considerably more face validity (which means that it makes sense on the face of it), such as the possibility that absentee/early voters disproportionately favored Bush²⁷, is that it is not an explanation, but rather a hypothesis. It's *apparent* that "Kerry voters were much more willing to participate in the exit poll than Bush voters" *only given several questionable assumptions*. An explanation would require independent evidence. In this case, the hypothesis *could* be tested. One possibility would be to examine discrepancies in precincts that were exit polled and in which the counts are unimpeachable, e.g., paper ballots placed in a tamper-proof box, which at the end of the day was counted in plain site of any observer who wished to see (assuming that some such precincts were in fact exit-polled). If the discrepancy in such precincts is consistent with the statewide discrepancy, then we have an explanation. If not, then we check that possibility off the list and move on to test other hypotheses.

The Role of the Exit Poll

The pollsters have made clear that the purpose of their poll was not to verify the integrity of the election; they were hired by the AP-Network consortium to provide supporting data for

²⁷ To the best of my knowledge, the pollsters have not offered absentee/early voters as an "explanation," presumably because they were able to predict any disproportionate support based on telephone interviews and patterns from previous elections.

election coverage.²⁸ In this election, however, concerns about this exit poll-count discrepancy are compounded by concerns about voting technologies, both new (especially electronic voting machines without paper trails) and old (punch card ballots still widely in use). Allegations about miscount and worse have been lighting up the Internet. In at least three cases, vote count errors have been acknowledged and corrected.²⁹ Additional sources of concern include mistabulation through “spoilage,” (as we saw in Florida in 2000, large numbers of votes are commonly lost due to imperfections in the voting process), overuse of provisional ballots, and limited access by observers to some vote tallies.³⁰

Verifying the integrity of elections is important not only in Mexico, Venezuela, and (the former Soviet Union Republic of) Georgia. Whatever the original purpose of this particular exit poll, it could be used to help verify the integrity of this election if its data and methodology could be reviewed by independent investigators.

Summary and Implications

In this paper, I have tried to demonstrate that:

- exit poll data are fundamentally sound,
- the deviations between exit poll predictions and vote tallies in the three critical battleground states could not have occurred strictly by chance or random error, and that
- no explanations of the discrepancy have yet been provided.

The unexplained discrepancy leaves us with two broad categories of hypotheses: the exit poll data are wrong or misleading in ways that have not yet been adequately explained, or the count is wrong. It's important that we review exit poll data and methodology and soon while the evidence

28 It's ironic that the same news organizations that have unquestioningly accepted that the exit polls “screwed up” in predictions on the Presidential election, have nevertheless drawn from those same exit polls such far reaching conclusions about moral values, the Democratic party's disconnect with the Heartland, etc ...

29 I'm sure I'm hopelessly behind in keeping up with these, but I happened to see these three: “Glitch gave Bush extra votes in Ohio” [cnn.com 11/05/04. http://www.cnn.com/2004/ALLPOLITICS/11/05/voting.problems.ap/](http://www.cnn.com/2004/ALLPOLITICS/11/05/voting.problems.ap/) . There were segments on Indiana and North Carolina errors on *MSNBC's* Keith Olbermann Countdown program: http://home.comcast.net/~karl.kaufman/countdown_on_voting_irregs_Nov12.wmv

30 Erica Solvig, “Warren's [Warren County, Ohio] vote tally walled off” *Cincinnati Enquirer* Friday, November 5, 2004

is fresh, but such analyses require NEP's raw data.³¹ Particularly useful statistical analyses would compare the "shift" in states, counties and precincts where safeguards are strong vs. those where they are suspect. Even more important, however, are investigation into the multitude of allegations and concerns about the count itself.

Mass media lack of interest in these concerns has been truly mystifying.³² I sat down to write this paper on November 7, after waiting five days for the press to offer an even plausible explanation. On November 10th, I circulated to colleagues a very early draft with big bold letters on every page, saying

DRAFT – Do not circulate, reproduce, post, or cite without the express consent of the author.

Despite the warning, it was circulated all over, and I released a paper with the qualifier that it was a VERY early draft, and to let me know if it is posted or cited. (My thinking was that I would at least then know where to send revisions). For the past 10 days I've been receiving email from all across the country – 1500 emails and counting – most of them extending thanks for documenting this discrepancy, drawing the conclusions, and asking the (obvious) questions that the media and everyone else with a public voice had been ignoring. But many have been allegations of mistabulation and worse. Three precinct workers from the Appalachian section of Ohio, for example, wrote:

360 people signed the book and 33 absentee ballots were cast for a total of 393 votes
The Board of Election is reporting 489 votes cast in that one precinct
WE HAVE A COPY OF THE ENTIRE POLL BOOK for this precinct.
(other totals were hand checked)

³¹ I contacted Warren Mitofsky, asking for access to NEP raw exit poll data; he seemed sympathetic to my request and said he would take it up with the NEP members; more recently, however, Joe Lenski wrote, "All of the exit poll data will be archived at the Roper Center at the University of Connecticut some time in early 2005. You will be able to access of the exit poll data at that point."

³² Lack of interest has been interrupted only by belittling dismissal: Manuel Roig-Franzia and Dan Keating, "Latest Conspiracy Theory -- Kerry Won -- Hits the Ether" *Washington Post*, November 11, 2004 and Tom Zeller, Jr. "Vote Fraud Theories, Spread by Blogs, Are Quickly Buried" *New York Times* November 12, 2004 (Front page)

They went to the FBI, who referred them to the Secretary of State's office, despite the fact that the precinct workers believe that the Secretary of State's office is culpable. (Ohio Secretary of State Kenneth Blackwell served as the state's Bush-Cheney campaign chairman this year.³³)

My purpose in this paper, however, has not been to allege that Bush did not legitimately win the election, let alone explain how it could have happened. I'm frankly suspicious of these charges (what ever happened to those 20,000 Democratic lawyers?), but I would investigate it if I could. Unfortunately, I cannot even look at but a small fractions of the emails I've received (sorry to those who have written), and have had to get a new email account so that my students can reach me. The question is why isn't the media asking questions? Because no one has provided solid explanations to the public, and no one seems to be investigating, suspicion of mistabulation, fraud, and even election theft is running rampant and unchecked.

Widespread assumption of misplay undermines not only the legitimacy of the President, but faith in the foundations of the democracy. That the President did not legitimately win the election is a very premature conclusion, but the election's unexplained exit poll discrepancies make it an unavoidable hypothesis. In this paper, I have tried to make the case that the media, academia, polling agencies, and law enforcement agencies should investigate it with a much greater sense of urgency and responsibility than they have thus far shown.

----- END -----

Appendix A: 2000 Voter News Service Exit Poll Methodology³⁴

METHODOLOGY STATEMENT

The VNS exit poll was developed and conducted by Voter News Service. The exit poll results are based on interviews with a probability sample of voters exiting polling places on Election Day, 2002.

Sampling

The samples were selected in two stages. **First**, a probability sample of voting precincts within each state was selected that represents the different geographic areas across the state and the vote by party. Precincts were selected with a probability proportionate to the number of voters in each precinct. Each voter in a state had the same chance to have his or her precinct selected. There is one exception. In some states, precincts that have large minority populations were sampled at a higher rate than other precincts.

³³ <http://www.cnn.com/2004/ALLPOLITICS/11/03/ohio.blackwell/>

³⁴ <http://webapp.icpsr.umich.edu/cocoon/ICPSR-STUDY/03527.xml> or http://www.ropercenter.uconn.edu/usvns2002_2.html

The sample weighting (described below) adjusts the representation of these precincts to their correct share of the total vote. **Second**, within each precinct, voters were sampled systematically throughout the voting day at a rate that gives all voters in a precinct the same chance of being interviewed.

The National sample is a subsample of the state sample precincts. The probability of selecting these precincts was the same as if the sample had been selected at a uniform rate nationwide, with the exception that minority precincts were again selected at a higher rate.

Weighting

The exit poll results are weighted to reflect the complexity of the sampling design. That is, the weighting takes into account the different probabilities of selecting a precinct and of selecting a voter within each precinct. For example, states that were selected at a higher rate receive a smaller weight than other precincts of the same size. There is also an adjustment for voters who were missed or refused to be interviewed, which is based on their observed age, race and sex.

NOTE: THE EXIT POLL RESULTS ARE ONLY MEANINGFUL IF THEY ARE WEIGHTED. UNWEIGHTED TABULATIONS MAY BE SERIOUSLY MISLEADING AND SHOULD NOT BE USED FOR ANY REASON.

Appendix B: German Exit Poll Results³⁵

Election German Parliament 22.09.2002

Party	Prognosis I Time 18:00	Prognosis II 18:00	Average of the two polls	Official Final Result	official vs average prognosis 1800
SPD	37	38	37.5	38.5	1.00
CDU/CSU	39	38	38.5	38.5	0.00
GRÜNE	9.5	9	9.25	8.6	-0.65
FDP	7	7.2	7.1	7.4	0.30
PDS	4.3	4.5	4.4	4	-0.40

Election European Parliament 13.06.2004 (only German part)

Party	Prognosis Time 18:00	First Projection 18:16	Official Result 04:04	official vs 18:00 prognosis
CDU/CSU	46.5	46.3	44.5	-2.0
SPD	23	22.9	21.5	-1.5
GRÜNE	10.5	10.6	11.9	1.4
PDS	6	6	6.1	0.1
FDP	6	6	6.1	0.1

Results in per cent. Polls close at exactly 18:00.

³⁵ This is the only data I've as yet been able to collect and review. (I'll be investigating past US and German elections, and any other exit poll data I can collect and as time permits over the next weeks.)

Appendix C: Slate Election Day Exit Poll Numbers

press box Media criticism

Updated Late Afternoon Numbers

Mucho flattering to Kerry; plus Nader makes an appearance.

By Jack Shafer

Updated Tuesday, Nov. 2, 2004, at 4:28 PM PT

Florida

Kerry 51
Bush 49

Ohio

Kerry 51
Bush 49

Michigan

Kerry 52
Bush 46
Nader 1

Pennsylvania

Kerry 53
Bush 46

Iowa

Kerry 50
Bush 49

Wisconsin

Kerry 51
Bush 48
Nader 1

Minnesota

Kerry 52
Bush 46
Nader 2

New Hampshire

Kerry 54
Bush 44
Nader 1

New Mexico

Kerry 50
Bush 48
Nader 1

Colorado

Kerry 49
Bush 50
Nader 1

Arkansas

Kerry 45
Bush 54
Nader 1

Missouri

Kerry 47
Bush 52

New York

Kerry 62
Bush 36
Nader 2

Nevada

Kerry 49
Bush 48
Nader 1

New Jersey

Kerry 54
Bush 44
Nader 1

West Virginia

Kerry 45
Bush 54
Nader 1

Why is *Slate* running these numbers? See this morning's [piece](#)³⁶.

³⁶ <http://slate.msn.com/id/2109053/#Post1#Post1>