Response to US Count Votes' Working Paper "Patterns of Exit Poll Discrepancies: On the Implausibility of a 'Uniform' Bias Explanation of the 2004 Presidential Election Exit Poll Discrepancies" May 12, 2005

Bruce O'Dell
Vice-President and Co-Founder, US Count Votes
Partner, Digital Agility Incorporated

June 1, 2005

Table of Contents

Overview

After unsuccessfully working within US Count Votes to revise or retract the Working Paper that a minority of the USCV membership recently published,[1] I see no alternative but to publicly challenge the report's methodology and conclusions.

The key argument of the USCV Working Paper is that Edison/Mitofsky's exit poll data cannot be explained without either (1) highly improbable patterns of exit poll participation between Kerry and Bush supporters that vary significantly depending on the partisanship of the precinct in a way that is impossible to explain, or (2) vote fraud. Since they rule out the first explanation, the authors of the Working Paper believe they have made the case that widespread vote fraud must have actually occurred.

However, a closer look at the data they cite in their report reveals that *Kerry and Bush supporter exit poll response rates actually did not vary significantly by precinct partisanship*. Systematic exit poll bias cannot be ruled out as an explanation of the 2004 Presidential exit poll discrepancy – nor can widespread vote count corruption. The case for fraud is still unproven, and I believe will never be able to be proven through exit poll analysis alone.

This paper should not be misinterpreted as an argument against the likelihood of vote fraud. Quite the opposite; I believe US voting equipment and vote counting processes are severely vulnerable to systematic insider manipulation and that is a clear and present danger to our democracy[2]. I strongly endorse the Working Paper's call to implement Voter-Verifiable Paper Ballots and a secure audit protocol, and to compile and analyze a database of election results.

But the fact that I chose not to endorse the USCV Working Paper should be a clear indication that I do not support its central thesis, and in fact believe that the simulation data refutes the Working Paper's conclusions. I am not a statistician, but as a computer systems architect, I create mathematical models to simulate the performance of large-scale computer systems, and mathematical simulation of the cost and efficiency of business processes is a significant part of my consulting practice. My own election simulation results are cited on pp. 9 -10 and in Appendix G of the Working Paper; as the creator of the only USCV simulation which accurately reproduces aggregate Mean WPE, Median WPE and participation rate data from the E/M January report, I feel an obligation to ensure that my work is correctly interpreted.

I will show that several of the USCV election simulation programs are flawed, and that when the Liddle Bias Index is applied to the "USCV O'Dell simulation" data cited in the Working Paper, it produces results consistent with those recently reported by Warren Mitofsky for the E/M data as a whole. Finally, I will present recommendations and conclusions, where I make the case for a more nuanced analysis of exit poll bias and vote fraud.

Bruce O'Dell
Vice President and co-founder, US Count Votes
Partner, Digital Agility Incorporated

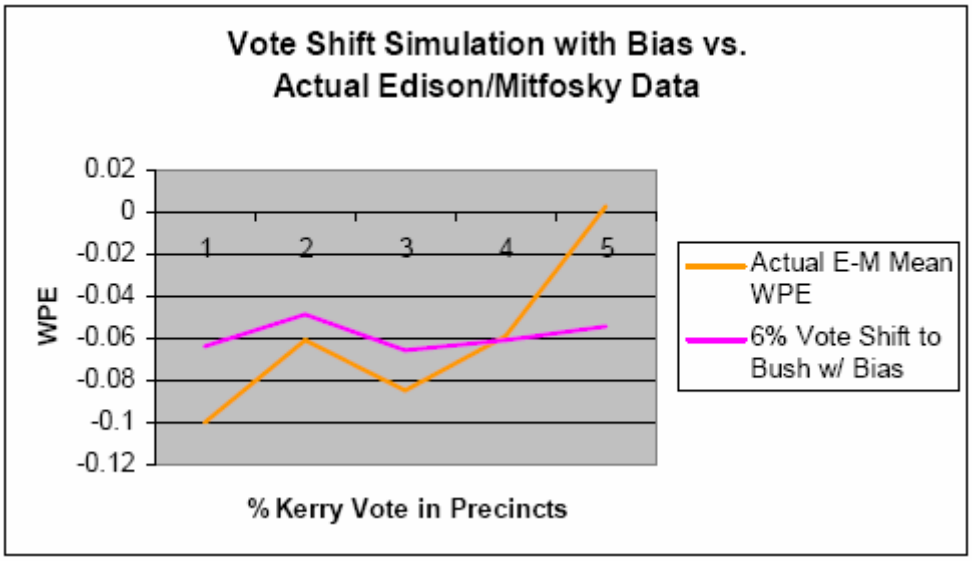
[1] USCV Working Paper at http://uscountvotes.org/ucvAnalysis/US/exit-polls/USCV_exit_poll_simulations.pdf
[2] See Appendix 1, Systematic Risks of Voting Systems below.

Review of the USCV Working Paper Simulation Results and Methodology

The results of a wide variety of election simulations are described in the USCV Working Paper. I will assess several of them to determine their accuracy and ability to confirm or dispute Warren Mitofsky's analysis of response bias as presented at the recent AAPOR convention[3].

1. USCV Vote Shift Simulator is flawed

Relative to E/M actual data, the USCV Vote Shift simulation on page 8 strongly *overestimates* Mean WPE (Within-Precinct Error) for the High Kerry precincts on the right side of the chart (-5.5% simulation versus +0.3% actual, an error of around six percentage points) and significantly *underestimates* Mean WPE for the High Bush precincts on the left side of the chart - by about four percentage points.



USCV Vote Shift Simulation results: Note generally poor fit between simulated and actual Mean WPE (from USCV Working Paper page 8)

Other than sharing a very general similarity in shape with E/M's reported aggregated Mean WPE, the USCV Vote Shift simulation *actually matches only one of five possible data points.* Large inaccuracies in the other four data points undermine the credibility of the model's simulation of Mean WPE data.

The Vote Shift simulation on page 8 does not report any computed Median WPE data, and the simulated participation rates listed in the table on page 8 do not match the aggregate precinct participation rates E/M reported.

The conceptual model of fraud that this simulation is attempting to reproduce is unstated. No reason is given for modeling a "6% vote shift" (the precise definition of which is not specified). There is no rationale for combining a "6% vote shift" with a "6% response bias". There is not even an implicit justification, since the results simply do not match the E/M data very well.

[3] http://uscountvotes.net/docs_pdf/analysis/US/AAPOR-XP-2005%2520(5-14-05)2.pdf

I believe, along with most investigators, that the discrepancy between 2004 exit poll results and official tallies can be explained by either fraud or poll response bias, since both are clearly possible. The explanation for what happened in 2004 undoubtedly lies somewhere between 100% fraud and 100% bias, with both potentially occurring in 0% to 100% of all possible precincts.

In other words, *there are many different hypothetical combinations of fraud and bias that can reproduce the aggregate E/M data*. For example, in Appendix 2 I've included a simulation that shows how a combination of uniform Kerry +8.5 response bias, plus "vote shifting" in just 3% of the total number of precincts clustered in the High-Bush category much more accurately reproduces the E/M aggregate signature than the Vote Shift simulator recipe of "6% response bias" and "6% vote shift" on page 6. This is not to suggest that scenario actually occurred – rather, to emphasize how easy it is to define arbitrary "recipes" of fraud and bias that match E/M's "signature". Since there are so many potential solutions to the same problem and no way to discriminate between them, this does not seem like a productive avenue of investigation.

Mathematically, given the available data, there is an enormous set of combined response bias and vote shifting scenarios scattered across many or a few precincts that collectively reproduce the aggregate E/M Mean WPE signature. Why advocate one over the others?

2. USCV Working Paper analysis of exit poll response bias is based on an inherently-skewed measurement

The simulation of exit poll response bias presented on page 6 of the USCV Working Paper appears mathematically accurate but does not address the limitations of WPE as a measure of response bias. It states on page 6: "Due to the algebraic properties of within precinct error (WPE)[12], any model of exit poll bias produces curves for WPE by partisanship with maximum WPE amounts where Bush and Kerry votes are closest to 50/50 and which go to zero at the endpoints in both Bush and Kerry vote strongholds."

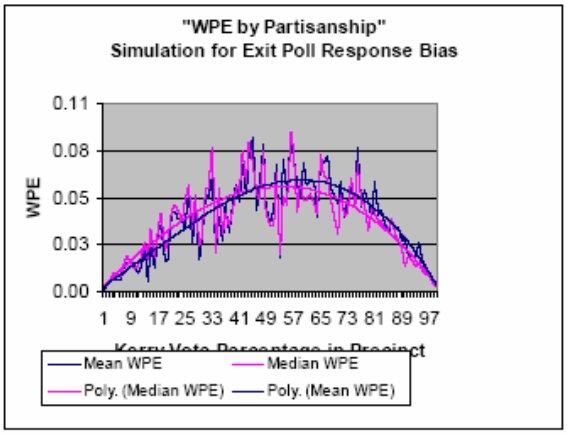
In fact, the important point is not that WPE goes to zero in the partisan strongholds, but that it does so *unevenly*. Liddle[4] showed clearly that although WPE was long thought to be a good measure of exit poll response bias, it actually is not. For a constant response bias, WPE is mathematically distorted (or skewed) in highly-partisan precincts, and so gives misleading results that appear to increase response bias in the highly partisan precincts of one's opponent.

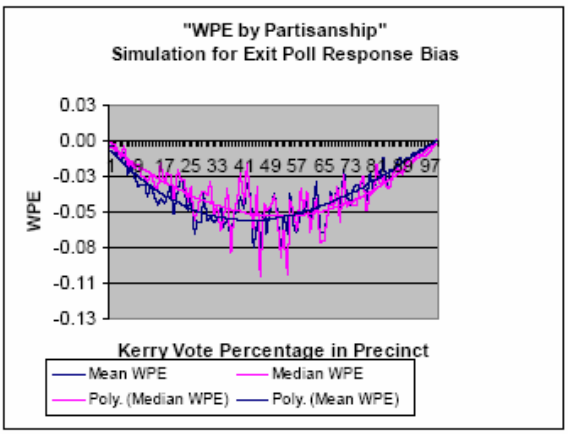
This is due to the mathematics inherent in the definition of WPE, and is not a matter of conjecture. All the USCV simulators can reproduce this effect. *A **constant** +Kerry bias produces **disproportionately more** negative WPE in High Bush precincts than in their counterpart High Kerry precincts.*

The USCV simulations of response bias on page 6 *are* skewed in precisely this manner (see below). However, note that the skew in the Bush bias curve on the left is much more readily apparent than for the corresponding Kerry curve. The Kerry curve *appears* less skewed because

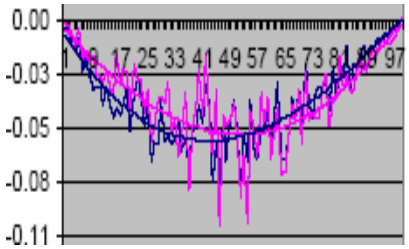
[4] http://www.geocities.com/lizzielid/WPEpaper.pdf

its horizontal scale is compressed relative to the Bush curve. When both curves are presented to the same scale, the tendency for constant + Kerry bias to generate disproportionate WPE in High Bush territory is more evident.



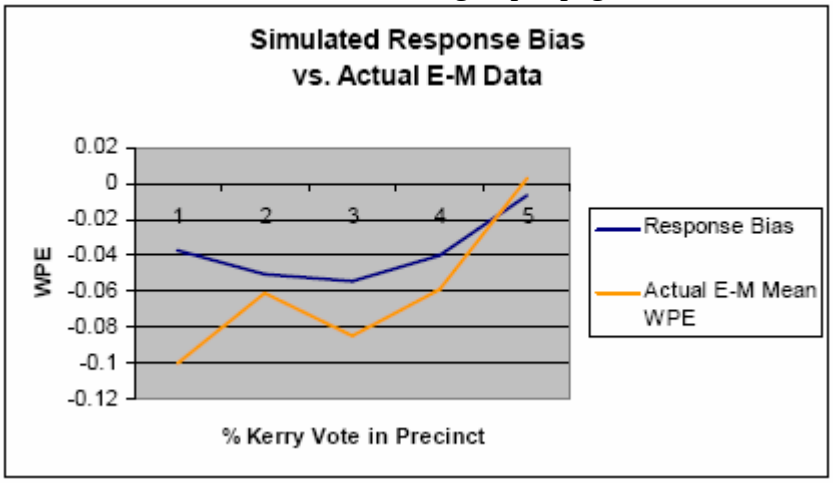


From USCV Working Paper page 6.
Note "skew" in WPE curve on left is more obvious than in the chart on the right, due to compressed scale.
The two curves should have rotational symmetry about the x-axis.
The chart below redraws the Kerry bias curve (above right) in comparable scale to the Bush bias curve (above left) to more clearly show the corresponding "skew" of the Kerry bias curve.



Kerry uniform response bias illustration from USCV Working Paper page 6, scaled to match Bush example.

In fact the "skew" inherent to WPE under conditions of constant response bias is much more clearly shown in this chart from the USCV Working Paper page 7:



USCV Working Paper p.7 – In the top line, note how constant response bias of Kerry + 6% skews WPE more negative in "High Bush" precincts on the left than in corresponding High Kerry precincts on the right

Because this inherent skew is built into the mathematics of WPE in the presence of a *constant* response bias, using WPE to measure response bias is misleading[5]. To address this problem – which she discovered - Liddle invented an "un-skewed" metric of bias called alpha. In doing so, Liddle made an *important contribution to the science of public opinion polling* that is welcomed and appreciated by everyone that understands the mathematics of the problem she solved.

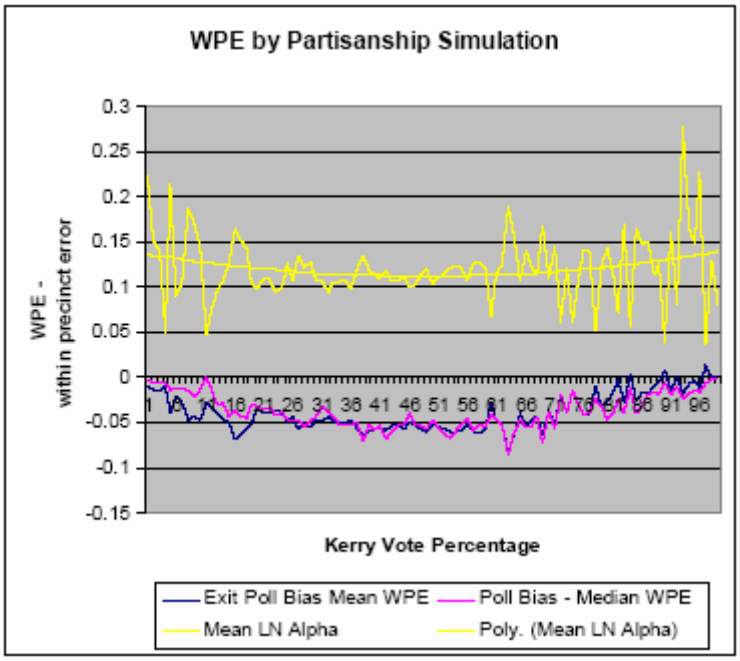
Alpha is simply the ratio of the exit poll participation of one candidate to another. With appropriate sampling protocols this is equal to, for each precinct, the (Number of Kerry poll responders/Number of Kerry voters)/(Number of Bush poll responders/Number of Bush voters). For example, if a precinct has 60 Bush voters and 40 Kerry voters and half of each group participate in the exit poll, alpha = (20/40)/(30/60) = .5/.5 = 1 , and we know that there was no response bias. For technical reasons, Liddle prefers to use the natural logarithm of alpha or ln(alpha) to create a linear, symmetric measurement tool – the "Bias Index".

Although alpha is simple to calculate, there are many ways to interpret its value.

Alpha values less than one are consistent with *any* of the following interpretations: exit poll under-participation by Kerry voters; exit poll over-participation by Bush voters; or net vote count corruption favoring Kerry.

Alpha values greater than one are consistent with *any* of the following interpretations: exit poll over-participation by Kerry voters; exit poll under-participation by Bush voters; or net vote count corruption favoring Bush.

Despite the fact that alpha removes the skew inherent to WPE, the authors of the USCV Working Paper only simulated alpha once, in a spreadsheet on p.4. That simulation produced the following chart (it shows the "Bias Index" or logarithm of alpha):



[5] See Appendix 4 "Why alpha corrects the inherent bias in WPE" below.

Above from USCV Working Paper p.4 – note the *non-linear* curve fit for ln(alpha). This is a consequence of the authors' choice to simulate response rate distributions with "standard deviations which allow *more variation in areas where sample sizes are smaller*." (Emphasis added)

The authors chose to build-in increased response rate variance at both ends of the chart and to use a polynomial regression. These choices *guarantee* a non-linear plot of ln(alpha). Whether this actually corresponds to the characteristics of E/M's aggregate data is left unstated. In fact, I will show later why this chart does *not* correspond to E/M's precinct-level ln(alpha) data.

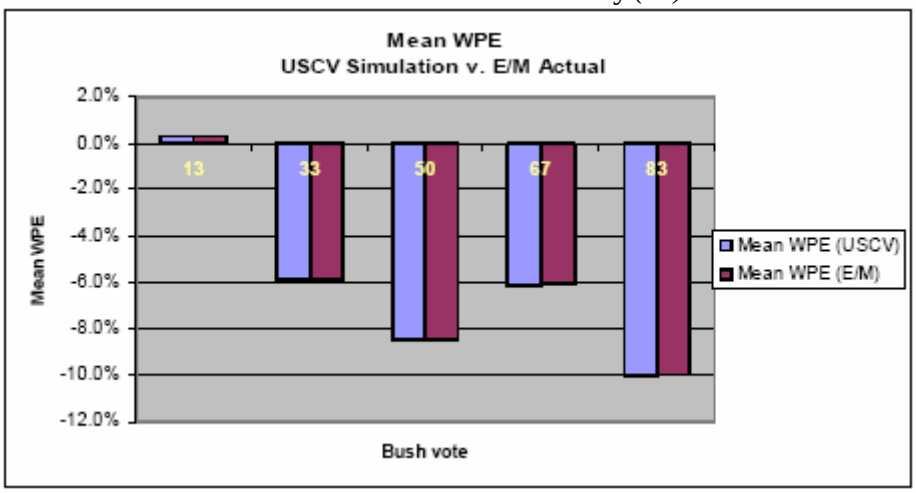
Other than in the chart on p.4, the remainder of the USCV Working Paper attempts to simulate E/M's aggregate WPE.

We know there are limitations to WPE as a measure of response bias. Although the authors of the Working Paper did not do so, it *is* possible to apply the alpha metric to the results of one of the USCV simulations cited in the working paper.

3. USCV "O'Dell Simulator" accurately models aggregate WPE - but also generates alpha consistent with Mitofsky's findings

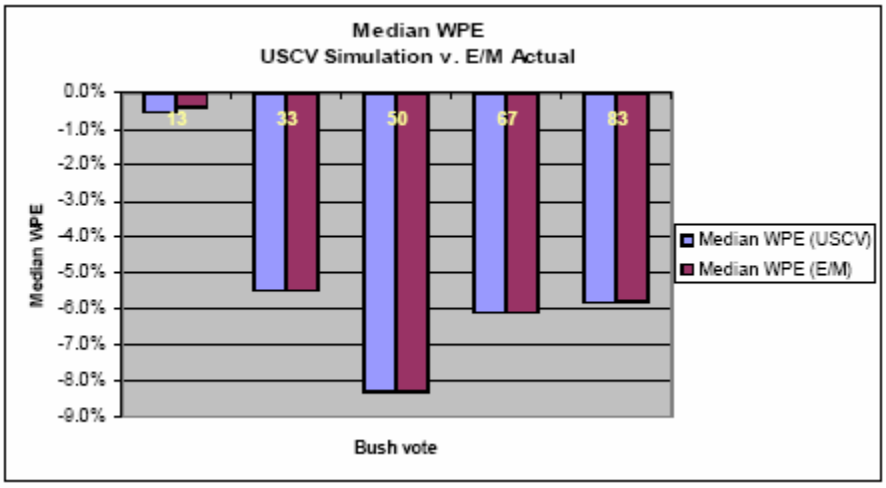
The USCV O'Dell Simulation published in the USCV Working Paper can accurately reproduce E/M Mean and Median WPE and Precinct Participation Rates.

Compare the Mean WPE chart on page 9 of the Working Paper to the chart on page 8 labeled "Vote Shift Simulation with Bias vs. Actual Edison/Mitfosky *(sic)* Data".

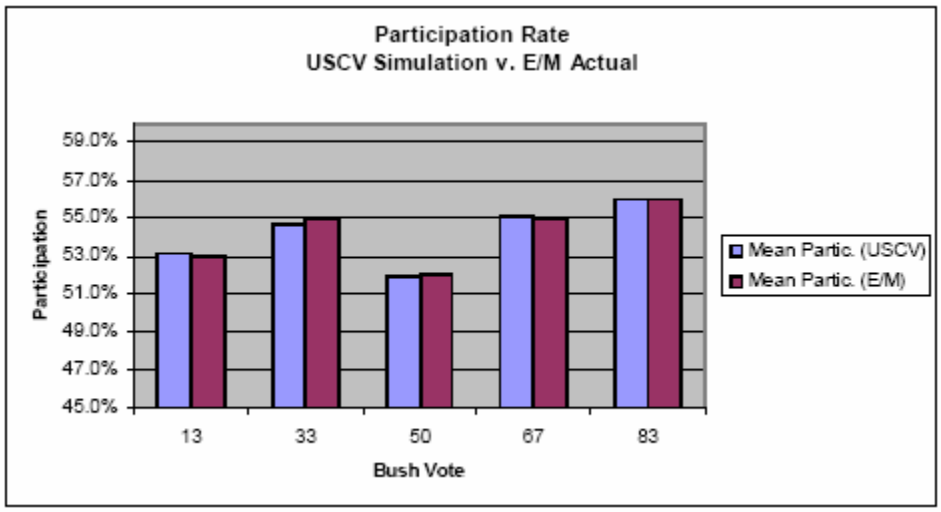


O'Dell USCV Simulation results: Note close fit between simulated and actual Mean WPE in contrast to inaccurate USCV Vote Shift simulation results on Working Paper page 8 (above from USCV Working Paper page 9)

The other two charts on page 9 of the working paper show how the O'Dell USCV Simulation accurately reproduces the E/M aggregate Median WPE and Precinct Participation data:

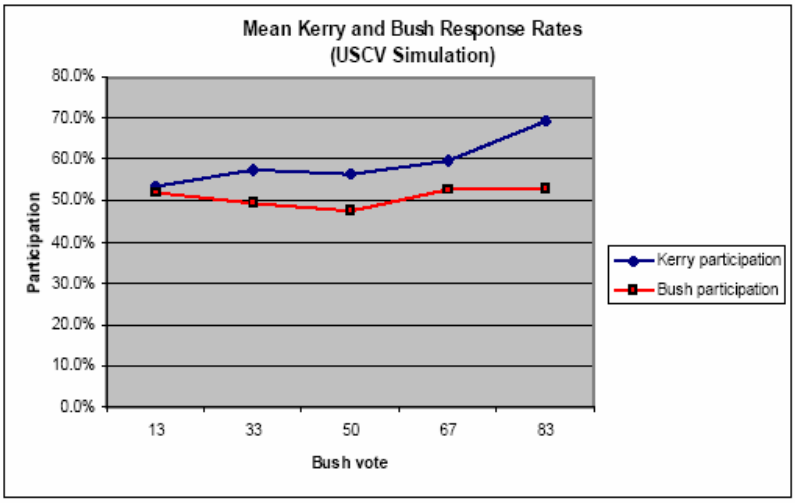


O'Dell USCV Simulation results: Note close fit between simulated and actual Median WPE
(above from Working Paper page 9)



O'Dell USCV Simulation results: Note close fit between simulated and actual Precinct Participation
(above from Working Paper page 9)

The USCV O'Dell simulation computes several other statistics, including alpha, Kerry and Bush supporter poll response rates, and bias (the average difference between the simulated precinct level Kerry and the Bush exit poll response rates). In the Working Paper, the computed response rate results were shown in a chart as follows:



O'Dell simulation results cited in USCV Working Paper p.10

Although the USCV Working Paper does not mention this obvious fact, note that there is an approximately *constant* response rate differential of about 7% plus or minus 1% for the three inner data points that represent about 89% or 1,120 of the overall set of 1,250 precincts.

However, the simulation's computed response rates *are* different for the extremes on the right and left of the chart. On the left of the chart, in the 90 High Kerry precincts (or about 7% of the overall sample), computed response rates are roughly *equal*. The Working Paper does not analyze the anomaly in High-Kerry precincts.

On the right of the chart, in the 40 High Bush precincts (about 3% of the total sample), the computed response rate differential is about *twice as big as in the center of the chart*, or about a 16% difference.

That is a large aggregate difference in response rate. It is an inevitable consequence of simulating a set of 40 precincts with Mean WPE = -10, Median WPE = -5.8 and an overall response rate of 56%. A *small number of precincts with very large negative WPE* are required.

The Mitofsky scatterplot is difficult to read, but it appears that there are even more extreme outliers in reality than in the simulation published in the Working Paper – not just in High Bush precincts, but throughout the entire range of partisanship.

The authors of the "Working Paper" conclude "the Kerry and Bush exit poll response rates must be non-uniform in order to match E/M's published numbers", but recall that the response rate differential looks rather uniform in the middle 89% of the sample, and paradoxically, some of the apparent excess WPE in High Bush precincts can be attributed to the inherent "skew effect" of *uniform* mean response bias that we discussed earlier – along with a few extreme outliers.

Aggregate data has its limitations. A precinct-level analysis is required to further investigate the apparent response rate anomalies in the High Bush and High Kerry precincts.

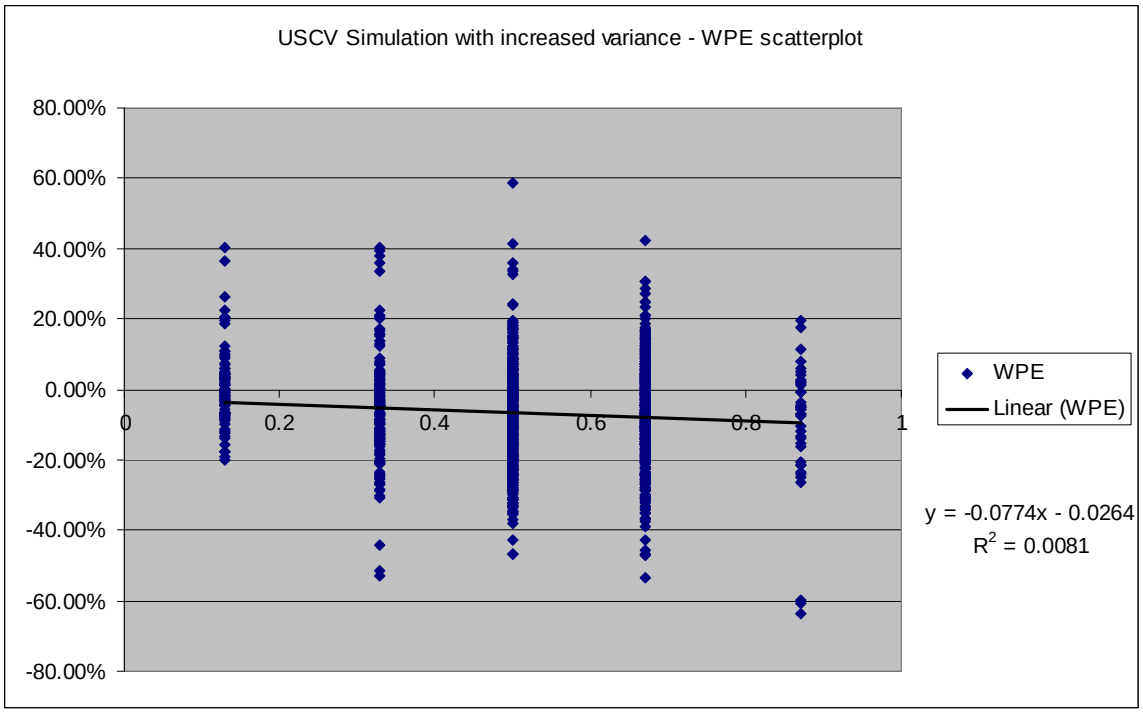
Scatterplot Analysis of USCV O'Dell Simulator

To the extent that my simulated precincts capture the essence of the full E/M sample, I should be able to reproduce Mitofsky's charts.

However, as noted above, when comparing my simulation with the actual E/M WPE scatterplot, I noticed that overall my original simulation (cited in the USCV Working Paper) contained noticeably less variance than E/M's actual data as presented by Mitofsky.

In fact, due to the design of the simulator's randomization method, I saw no simulated precincts with WPE greater than +20%, and I also saw no precincts with very large negative WPEs (as low as -60% in the E/M scatterplot). I therefore manually adjusted my simulated precinct sets to introduce more variance, in line with E/M's actual WPE numbers, while still maintaining the original reported participation, mean WPE and median WPE values[6].

Here are my scatterplots for WPE, alpha, and ln(alpha):



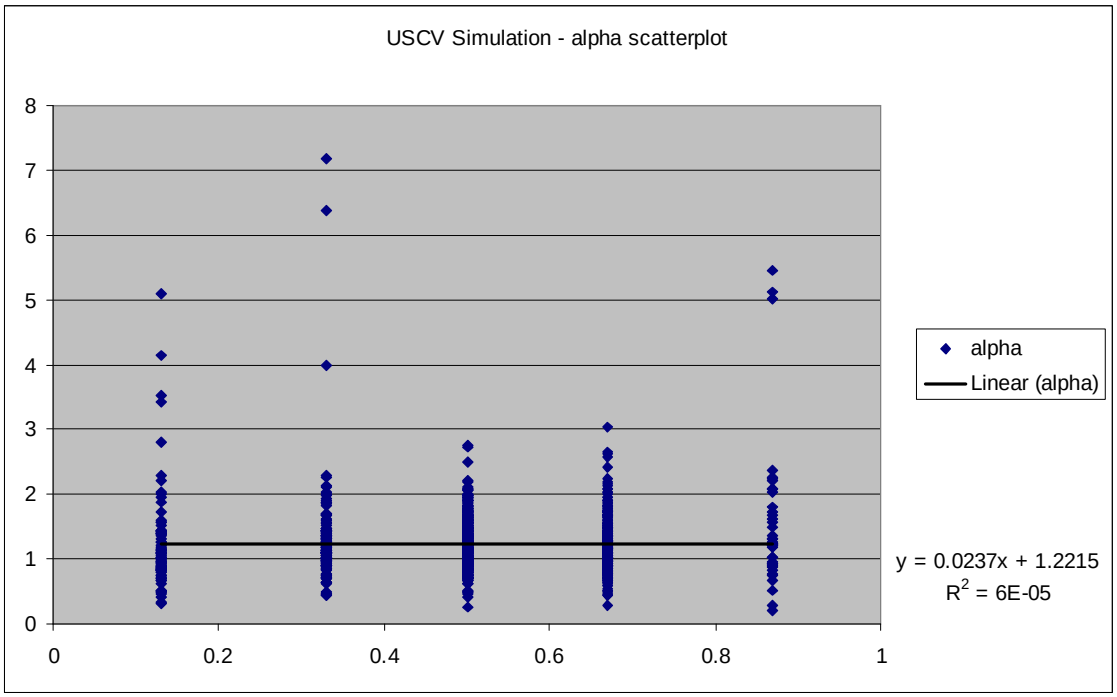
Scatterplot of USCV O'Dell Simulator WPE
(compare to Mitofsky at http://www.mysterypollster.com/photos/uncategorized/wpe_plot.jpg)
Note that by convention, Mitofsky shows Negative WPE above the x-axis while I show it below.

The USCV O'Dell simulator not only accurately models the aggregate mean WPE, median WPE and participation rates reported in the January E/M report, it also reproduces the apparent linear correlation of partisanship with WPE, y = -0.0774x - 0.0264.

[6] This simulation data is available for download from www.digitalagility.com/data/USCV_ODell_Simulation2.xls and the original dataset is available at www.digitalagility.com/data/USCV_ODell_Simulation1.xls .

For this regression, Pearson's correlation coefficient r = -0.09007, t = -3.19506, and for a two-tailed test of significance, p = .0014 (significant). ***This closely corresponds to the WPE regression results presented by Mitofsky at the recent AAPOR conference.***

The following scatterplots show the corresponding results for alpha and ln(alpha).

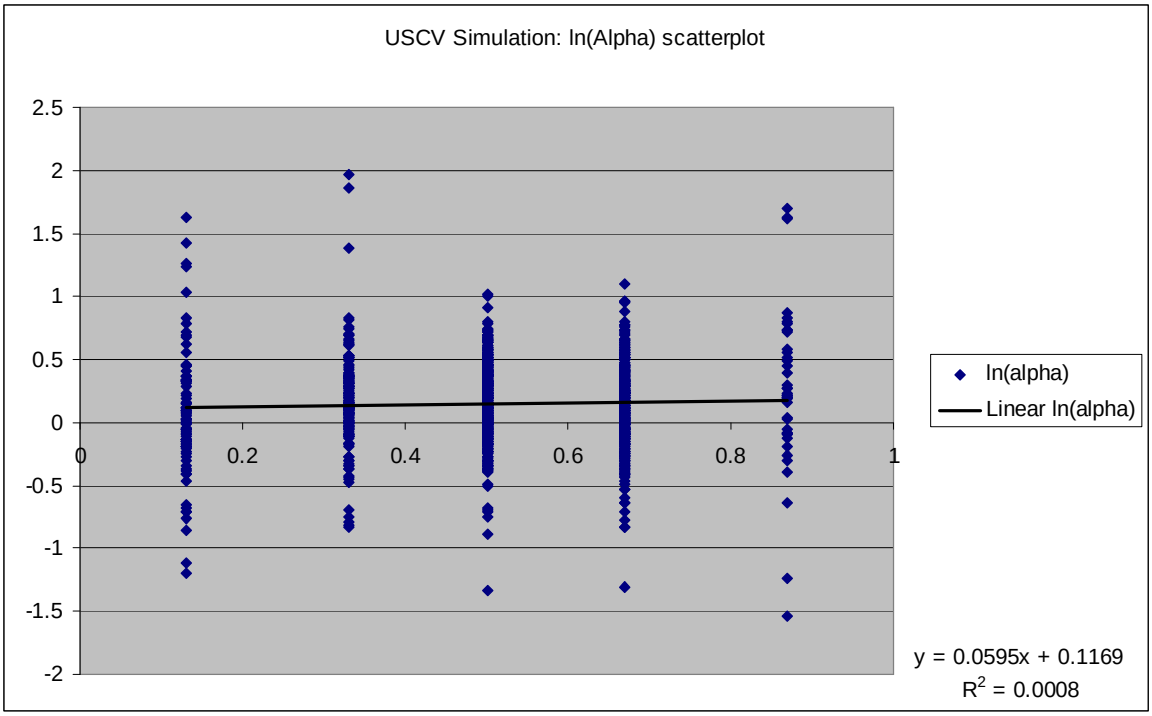


Scatterplot of USCV O'Dell Simulator alpha

This regression shows a visually flat, constant correlation between alpha and partisanship of y= .0.0237x + 1.2215. For the regression of alpha, Pearson's correlation coefficient r = 0.007667, t = 0.270864, and for a two-tailed test of significance, p = .7865 (not significant).

That implies alpha for B13% = .0237(.13) + 1.2215 = 1.224581 and alpha for B87% = .0237(.87) + 1.2215 = 1.242119, or *only about a 1.4% change across the entire range of partisanship – and this slight correlation is not statistically significant.* It should be obvious that WPE varies much more than alpha across the same range of partisanship, and is a skewed measure of response bias relative to alpha.

The final scatterplot shows the natural logarithm ln(alpha). Taking the natural logarithm of alpha corrects for potential asymmetries in the plot of alpha and reduces geometric changes in alpha to linear ones, for easier analysis. The following chart can be directly compared to Mitofsky's "Bias Index" scatterplot.



Scatterplot of USCV O'Dell Simulator Bias Index
(compare to Mitofsky at http://www.mysterypollster.com/photos/uncategorized/wpe_index_plot.jpg)

The regression shows ln(alpha) = .0595x + 0.1669 as a function of partisanship.

For the ln(alpha) regression Pearson's correlation coefficient = 0.029, t = 1.008, and for a two-tailed test of significance, p=.314 (*not significant)*.

In other words, ***this USCV simulation shows there is no statistically significant correlation between precinct-level ln(alpha) or "Bias Index" and partisanship. This also closely matches Mitofsky's findings.***

USCV O'Dell Simulation - Sensitivity Analysis
Some investigators have questioned the Mitofsky linear regression results and suggest that there are local deviations from overall linear Bias Index worth investigating.

This seems questionable, since if the local deviations are both large enough to affect the outcome of the election and are correlated to partisanship, you would expect to see an overall non-linear trend in alpha. Even so, although Mitofsky did not perform such a sensitivity analysis, I can, via the USCV O'Dell Simulator. The results are in Appendix 3.

The most interesting result came as a complete surprise. The apparent correlation between WPE and partisanship disappears when High-Kerry precincts are removed from the overall sample!

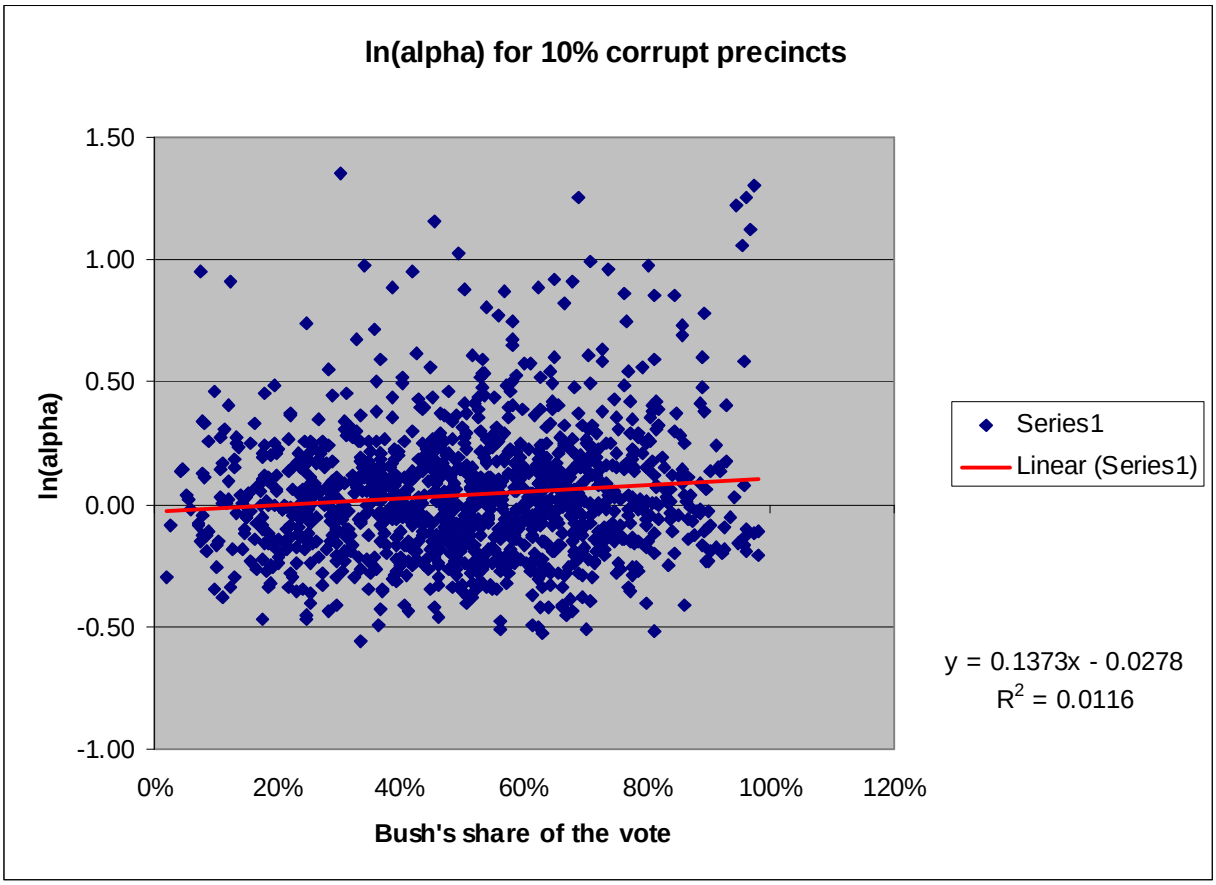
The correlation between WPE and partisanship seems to be driven by "more positive than expected" WPE in the *High Kerry* precincts, which has effect of tilting the left side of the regression line *up*. When High Kerry precincts are eliminated, the statistically-significant

correlation between WPE and partisanship almost completely vanishes. This surprisingly seems the reverse of Bsvcc (the USCV "Bush stronghold vote count corruption" hypothesis), since the statistical correlation between WPE and partisanship *changes very little* when the High-Bush precincts are removed. It would be interesting to see if the same sensitivity analysis applied to the actual E/M data would produce the same results.

4. Simulation of significant and pervasive vote shifting leads to noticeable changes in the alpha scatterplot

To distinguish between fraud and bias, we need to see what effect an accurate simulation of vote shift would have on the alpha bias curve.

Liddle[7] has created a vote shift simulation that shows how widespread uniform vote shifting produces both a noticeable and statistically significant signature in ln(alpha) - as depicted in the following chart.



(Above) Liddle vote shift simulation, ln(alpha) regression of simulated 10% vote shift.
Liddle's simulator recalculates every time the spreadsheet is opened or when f9 is pressed, so your results will vary.

For this regression, Pearson's coefficient of correlation r = .107573, t = 3.82242, and for a two-tailed test of significance, p=.0001 (significant).

[7] Liddle vote shift simulation is available at www.digitalagility.com/data/errorsim.xls

The absence of this signature in the Mitofsky scatterplot implies either that widespread pervasive uniform vote shifting did not occur in 2004, that a large number of smaller vote shifts occurred, or that the statistical signature of pervasive vote count corruption was masked by other factors.

For example, a hypothetical carefully-targeted vote fraud conspiracy could potentially affect the outcome of the Electoral College contest decisively, without requiring votes to be shifted uniformly and pervasively.

Conclusions

1. The USCV Working Paper has not refuted "constant mean exit poll response bias"

The Abstract of the USCV Working Paper states: "New evidence from mathematical simulations conclusively shows that any constant mean exit poll response bias hypothesis such as the 'reluctant Bush responder' (rBr) hypothesis is not consistent with the pattern shown by the Edison/Mitofsky exit polling data."

This statement is not supported by the simulation results cited in the paper.

A significant positive correlation between WPE and partisanship exists, and has been confirmed by this analysis and by Mitofsky. But since we now know WPE is a skewed measurement rod when it comes to assessing response bias, analyses based solely on WPE are literally off-target.

In fact, using either alpha or ln(alpha), the USCV O'Dell Simulator cited in the Working Paper not only accurately reproduces the aggregate E/M Mean WPE, Median WPE and participation rates, at the same time it shows *no statistically-significant correlation between exit poll bias and partisanship* when precinct-level alpha is analyzed.

The fact that the both the E/M data and the USCV O'Dell simulation data appear to support the Working Paper's thesis of highly differential response rates when calculations based on *aggregate WPE* are modeled - but supports a uniform mean response rate bias when *precinct-level alpha* is modeled - indicates that the USCV O'Dell Simulation actually *has* captured essential aspects of the E/M data.

My conclusion is that the USCV Working Paper calculations derived from aggregate WPE are inherently inaccurate in modeling response bias, and are also highly sensitive to the presence of a few extreme outlier precincts; they should be replaced. Calculations based on precinct-level alpha are mathematically sound, and should be used instead when modeling response bias.

2. There is no evidence for the USCV "Bsvcc" hypothesis: disproportionate "Bush stronghold vote count corruption" is not indicated by the currently available data

There is no indication in the Mitofsky scatterplots or in this analysis of an exceptional amount of either WPE or alpha in the High Bush precincts. However: *the data could be consistent with*

widespread vote count corruption, with widespread uniform mean differential response bias, or a (currently unknowable) combination of the two.

3. A much more nuanced assessment of fraud and bias is needed

It is possible to speculate that if a hypothetical group or groups actually have the power to covertly shift votes on a national scale sufficient to change the election outcome, they doubtless would be extremely careful to try to avoid detection – *especially, one would think, by angry statisticians.*

Recall that E/M's published sample is only 1,250 out of the approximately 170,000 precincts in the US[8]. The full E/M exit poll covers less than 1% of all precincts. That is sufficient for their purposes, but there is no reason to presume that it is also necessarily suitable for detecting widespread vote count corruption.

There are conceivable vote fraud scenarios that might totally bypass the E/M sample. Even if some corrupt precincts are included in the E/M sample, it may not be possible to reliably distinguish them. And even if they are identifiable, they may occur disproportionate to their actual incidence in the field. Or the net effects of uniform mean response bias in E/M's exit polls may accidentally correspond to the net effects of vote fraud in precincts outside the exit poll sample.

Even so, I join many others in recommending additional analysis of the extreme outliers in E/M's data – favoring both Bush and Kerry. The underlying explanations may well be innocuous, but extremely large and extremely small values of the bias index *are* suspicious. They are just as consistent with a concentrated pocket of nullified votes or vote switching as they are of unusually poor poll-taking technique.

New regressions of the bias index against voting equipment type, geography, and the other variables analyzed in the E/M January report are also appropriate, since the previous analysis based on WPE is now known to be skewed.

Finally, any coherent model of pervasive manipulation must generate testable hypotheses.

Any hypothetical group of vote manipulators would be faced with a difficult fundamental tradeoff: whether to implement a large number of small vote shifts that are harder to detect individually, but perhaps more difficult to implement collectively in complete secrecy; as opposed to a smaller number of large vote shifts that may be easier to engineer, but also easier to detect - especially via statistics.

Given the conflicting goals of avoiding detection and ensuring their candidates are elected, the hypothetical vote manipulators may choose to concentrate on two geographic scenarios: a primary goal to selectively tip the electoral votes of "battleground" states, with a secondary goal

[8] http://www-vdc.fas.harvard.edu/ROAD/docs/node9.shtml

of driving up the overall popular vote margin of victory in less-closely contested (and closely-watched) ones.

Therefore the hypothetical vote manipulators may prefer a strategy of large numbers of small vote shifts in battleground states, anticipating risk from a high degree of post-election scrutiny. A strategy of fewer, larger vote shifts may be reserved for strongholds of one candidate or the other where they would attract less attention. This could potentially be testable, but not easily by exit poll data. The signature of low-level, pervasive vote shifting would fall in the portion of the scatterplot just above the x-axis, and would be difficult to distinguish from ordinary variance in poll response.

The recent analysis of the Ohio exit poll data reported by Election Science Institute is a start. But as ESI alluded to on pp. 9 - 11 of their presentation[9], and given the small number of precincts surveyed by E/M, it may take in-depth analysis of actual precinct-level results as also proposed by USCV, and others, to rigorously test the various fraud hypotheses and identify areas for further conventional investigation.

As one example, if we assume our hypothetical vote manipulators focus their efforts on the most important and influential elected offices, a detailed analysis of the patterns of ticket-splitting over time by type of election equipment may be worthwhile.

A call to move beyond Exit Polls

The USCV Working Paper authors have yet to make their case that differential response bias is absolutely ruled out nor that widespread vote fraud is *required* – much less, that we can actually estimate the specific combination of each one that occurred. They are, like the rest of us, still in search of a method of reliably allocating the overall exit poll discrepancy to the correct combination of fraud and/or bias that actually occurred in a specific set of precincts in 2004.

Edison/Mitofsky should re-run their WPE-based exit poll regressions using the new Bias Index to determine if there are any significant new findings. The rest of us may have taken our exit poll analysis as far as it can reasonably go.

I strongly recommend that USCV should focus instead on the construction and analysis of our proposed National Election Data Archive.

As Einstein liked to say, the solution to any problem should be as simple as possible, but no simpler. The USCV Working Paper is far too simple a solution to an enormously complex problem.

[9] http://www.electionscience.org/Members/StevenHertzberg/report.2005-05-14.4978140903/report_contents_file/

Appendix 1: Systematic Risks of Voting Systems

Many people believe that widespread vote manipulation is not feasible because it would involve too many participants in too many locations to remain secret for long. They are misinformed.

In the past, since votes were recorded and counted on "analog" systems - paper, punch card, mechanical - vote fraud was "retail"; it generally required hands-on access. And in that era, votes were only fully recounted – that is to say, audited - when the tally was very close. Manual counting and tabulation procedures were presumed to have a small margin of error, and a recount could change the outcome of a close election.

In the modern era, vote fraud can easily be performed "wholesale" and at a distance. There *is* an inherent, systematic vulnerability in voting systems - not from outside hackers, but from malicious insiders. In the financial services industry, we refer to these people as "embezzlers", and we design elaborate auditing mechanisms to deter and detect their activities. A handful of people with the right administrative access and insider knowledge of internal controls and security measures can generate an enormous security breach. Enron is one extreme example.

In computer science terminology, a voting system is a very good example of a geographically-dispersed distributed system. Collectively it consists of thousands of individual devices. All of those devices are programmed (in what amounts to an industrial process) using centrally managed master software copies that are under the control of just *a very few* administrators. Each individual device also contains hardware components the function of which, by visual inspection alone, is impossible to verify.

The only way to know what such a system is actually programmed to do (as opposed to what we *think* it is programmed to do) is to observe the visible inputs and verify that they match the visible outputs in the way we expect.

For example, although your bank keeps your private information confidential, your name is on your account, your checks and credit cards, and you and the other parties the bank does business with receive regular statements. This web of record-keeping vastly simplifies the job of keeping your money safe. But votes are much more difficult to secure than automated teller transactions.

Your vote is a *private* and *anonymous* transaction which nevertheless *must be recorded with complete accuracy*. Think for a moment about how difficult that is to guarantee!

Imagine the potential for fraud at your bank if you had to submit anonymous, private transactions with no independent audit trail.

Suppose that in order to make a deposit, you go behind a screen, take your cash out in private, put your account number on a deposit slip, and drop it and the cash in a barrel. At the end of the day, the bank knows only the number of customers it had. The barrel is opened, the deposits are handed to the tellers, and money is posted to accounts - but no receipts are issued, you never receive a statement, and if challenged, all the bank can ever do is double-check the sum of the

day's deposit slips – they can't actually tell you that *your* deposit was properly credited. And then suppose they won't even do that unless one of their customers asks for an audit - and then pays for it out of their own pocket.

How long would you keep your money in that bank?

Electronic voting systems are not only inherently vulnerable to systematic manipulation, they are uniquely challenging to keep secure, precisely because applying conventional security techniques tend to breach either the privacy or the anonymity of your vote.

The only fully reliable way to ensure that anonymous, private voting transactions are tallied accurately is to capture and store a voter-verified paper ballot – a document with legal standing that we can use to randomly audit the official tally to deter fraud and to audit *en masse* if problems are detected or suspected, to undo the theft of an election.

Thirty percent of the vote in 2004 was tallied on electronic devices which, by design, produce no independently-verifiable audit trail. In fact, Florida even *prohibits* recounting its touch screen voting equipment, since they feel the outcome is accurate by definition.

How did "trust me" ever become an acceptable standard of audit and control for elections?

These are such bad ideas that when US members of the Association of Computing Machinery (ACM) - the world's oldest and largest computer society – were polled last year, *95%* of us agreed that voting equipment should not be deployed unless it featured a "voter-verifiable paper ballot" that could be audited or recounted[10].

There is overwhelming consensus among computer professionals that our election equipment and vote auditing procedures are highly vulnerable to systematic manipulation[11].

This poses a clear and present danger to our democratic way of life.

The solutions are straightforward and the cost is minimal.

We have no excuse not to act.

[10] http://www.acm.org/usacm/weblog/index.php?p=73
[11] See for example http://www.acm.org/usacm/Issues/EVoting.htm

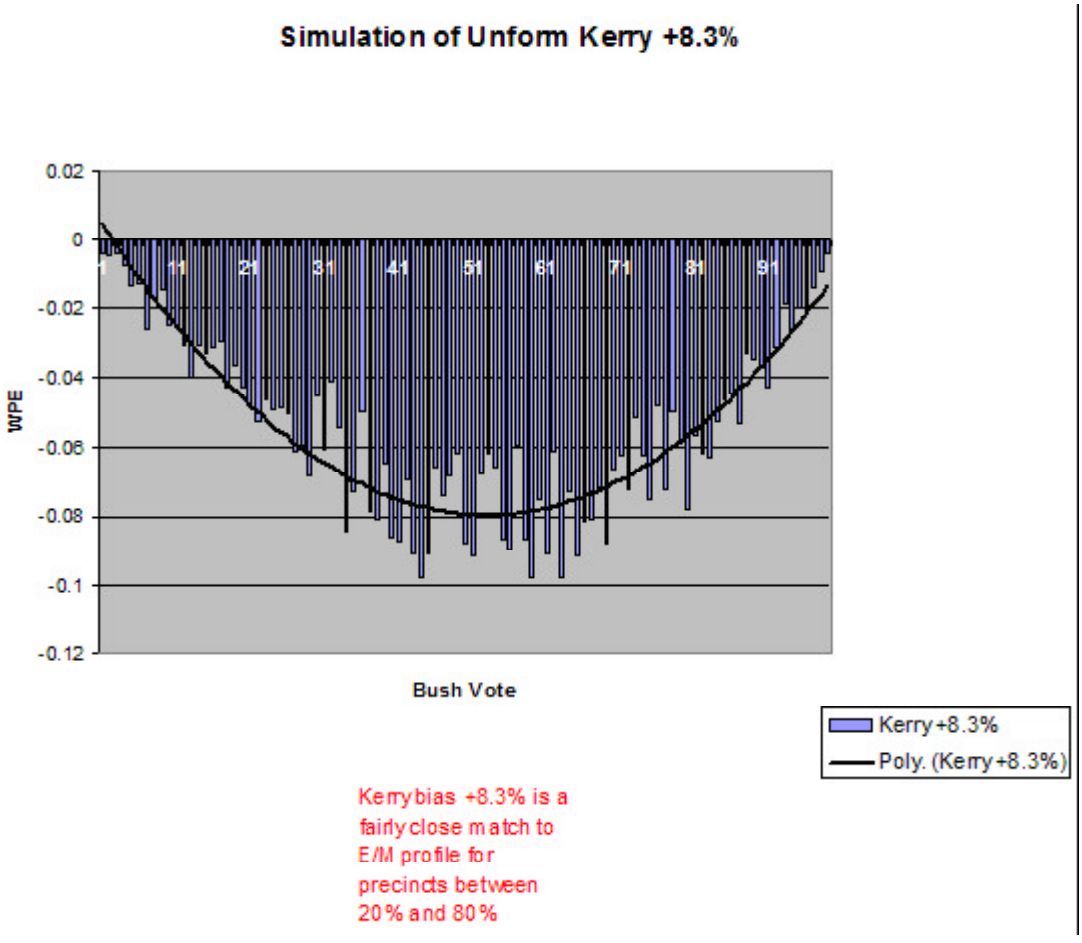
Appendix 2: Arbitrary combination of fraud and bias

A random combination of response bias and vote count corruption, yet it closely mimics the aggregate E/M data

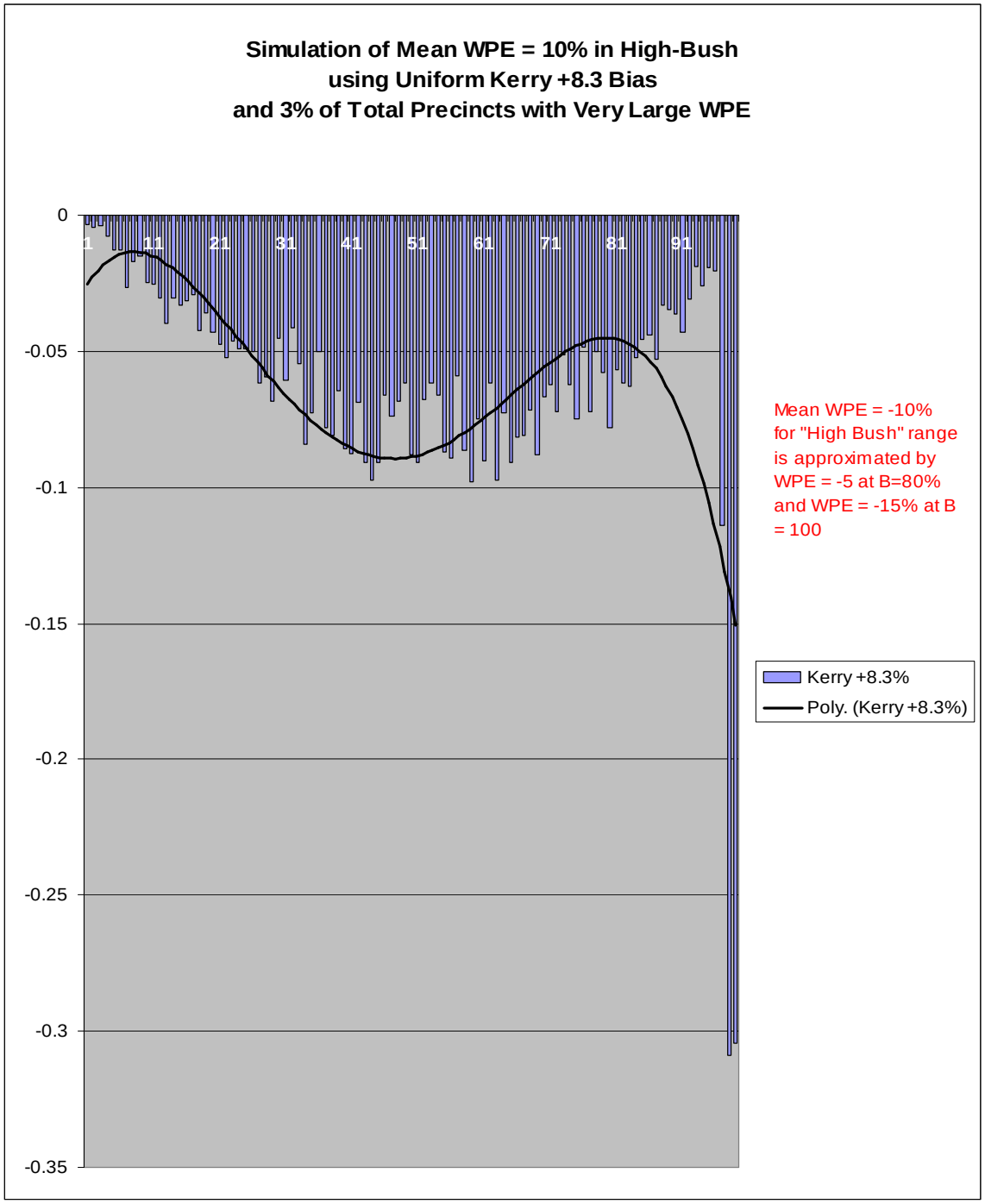
Here is a plot of uniform Kerry + 8.3 response bias.

Note that for precincts between 20% and 80%, there is a good visual match to E/M's Mean WPE aggregates of:

Bush = 20% to 40%, Mean WPE = -5.9
Bush = 40% to 60%, Mean WPE = -8.5
Bush = 60% to 80%, Mean WPE = -6.1



On the next page, I manually adjusted 3% of the overall precincts as if there was significant vote count corruption in the High Bush Precincts.

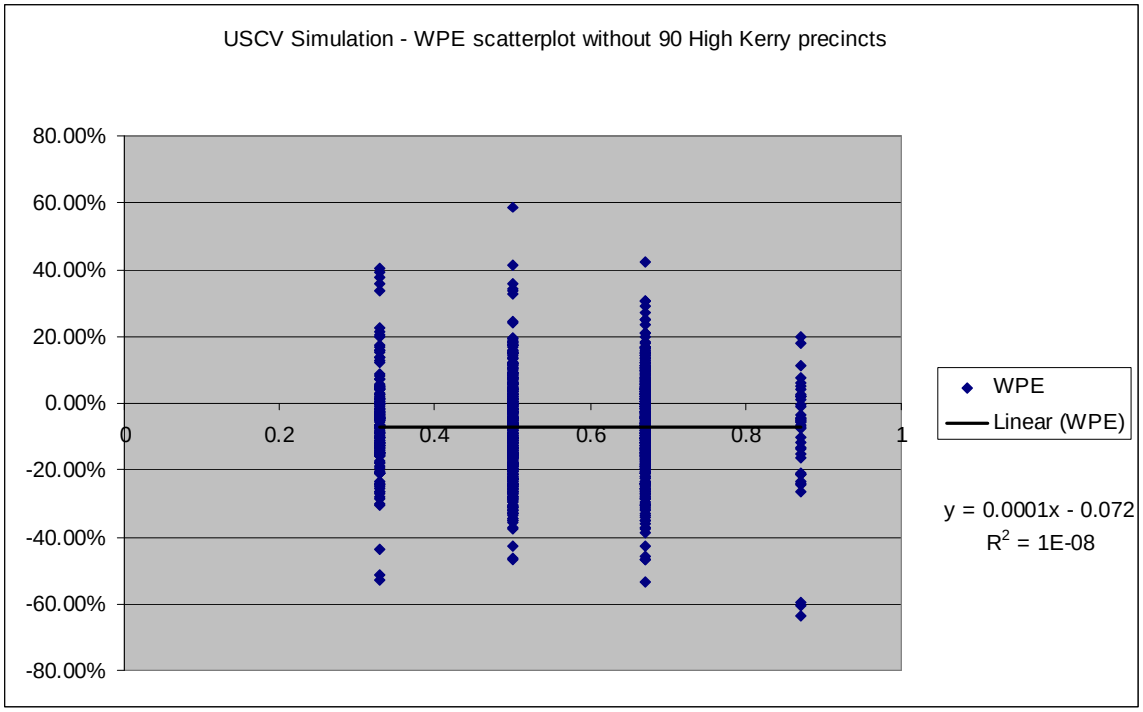


I am not at all suggesting this as a credible model of what happened in 2004. I include it as a case in point to show just how easy it is to arbitrarily combine fraud and bias scenarios to generate the aggregate E/M January numbers. That approach is useful only to the extent it generates hypotheses that can be tested against the actual E/M data.

Appendix 3: Simulation Sensitivity Analysis

Sensitivity Analysis: Remove High-Kerry Precincts

If the High Kerry precincts are eliminated, the following somewhat surprising WPE and alpha scatterplots are produced:

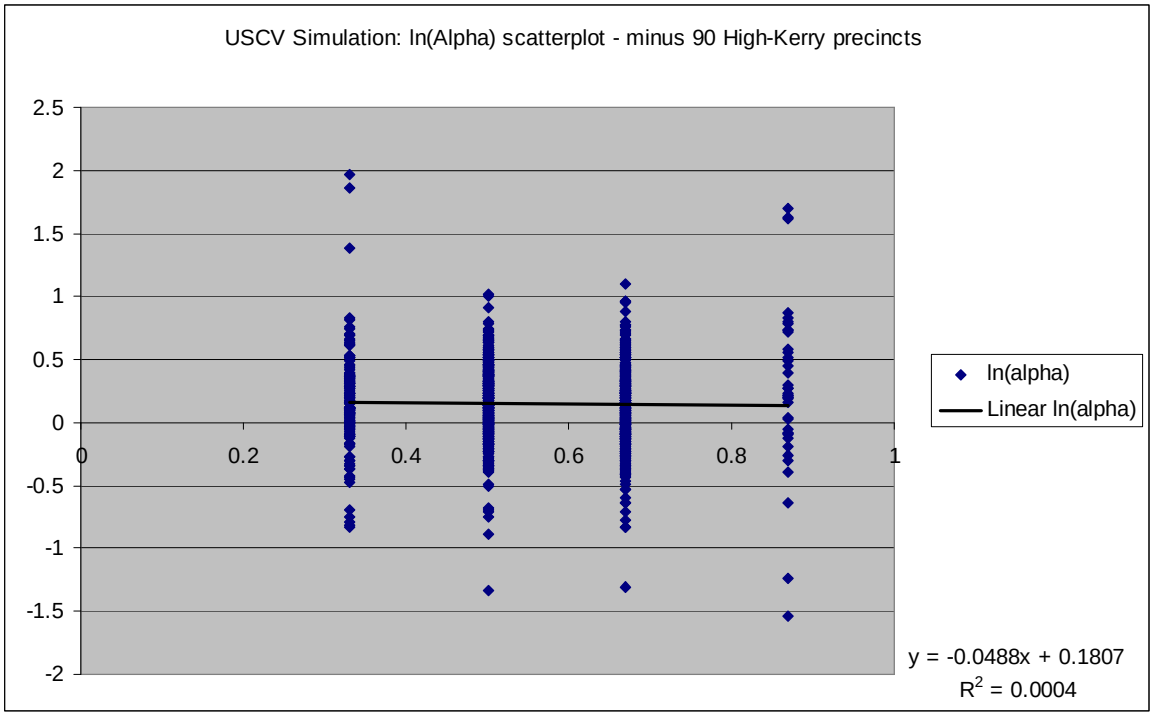


(Above) O'Dell USCV Simulation – WPE regression, excluding 90 High-Kerry precincts.

For this regression, Pearson's correlation coefficient r = 0.000106, t = 0.003603, and for a two-tailed test of significance, p = .971 (not significant).

After removing the High Kerry precincts, the regression of WPE across the other E/M aggregates yields a constant value *of - 7.2%.* This suggests it is "excess positive" WPE in the High Kerry precincts that is the significant influence on the overall WPE trend line.

This is an unexpected and intriguing finding, and it would be interesting to know if this simulation accurately characterizes the behavior of the full E/M dataset when the High Kerry precincts are removed.



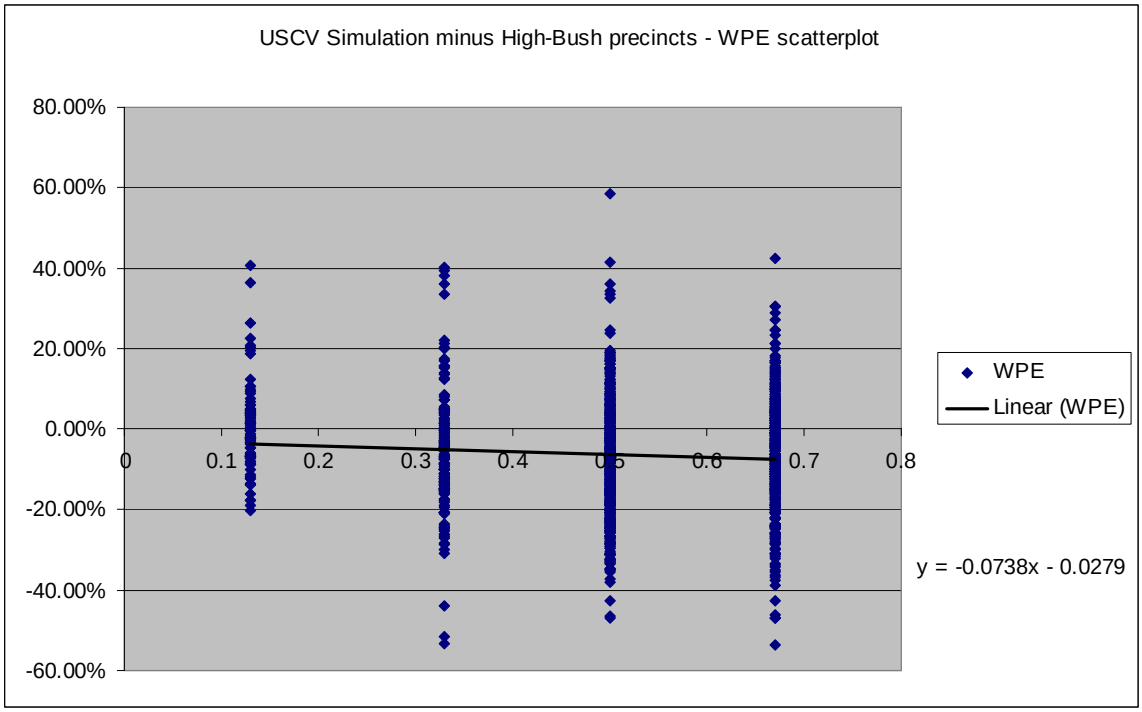
(Above) O'Dell USCV Simulation – ln(alpha) regression, excluding 90 High-Kerry precincts. Although the correlations are not statistically significant, note the slight negative slope in ln(alpha) here, as opposed to the slightly positive slope of ln(alpha) for the full dataset.

In this regression, the Pearson correlation coefficient r = -0.01912, t = -0.65068, and for a two-tailed test of significance, p = .5154 (not significant). These results can be readily reproduced by downloading the simulation results and deleting the 90 simulated High Kerry precincts from the "Scatterplots" tab in the USCV O'Dell simulation data spreadsheet.

Sensitivity Analysis: Remove High-Bush Precincts

When the 40 High Bush precincts are removed, we see the following WPE and ln(alpha) scatterplots.

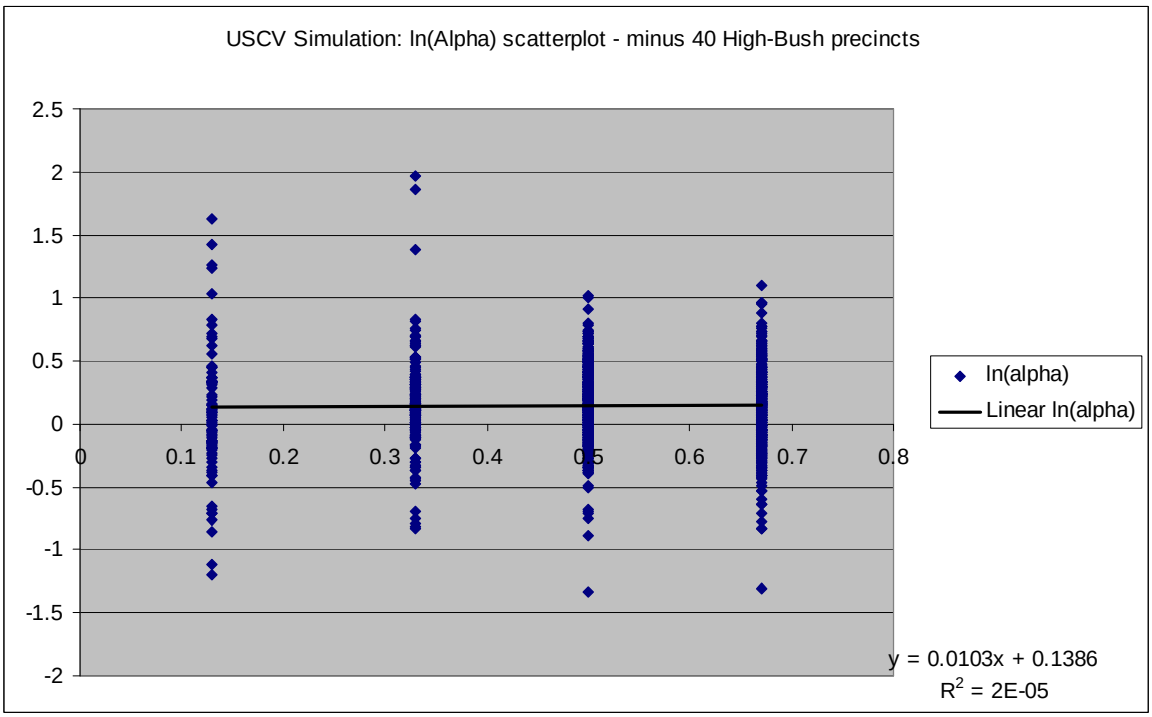
Contrary to my initial expectation, *there is little change in the slope of the regression line.* This indicates that the presence or absence of the High-Bush precincts seems to have little impact on the overall relationship between WPE and partisanship.



Note that the slope of WPE is quite similar to the value found for the overall dataset (-.0774x)
Apparently the presence or absence of the High-Bush precincts has little effect on overall WPE.

In the regression of WPE against partisanship for the simulation minus the 40 High Bush precincts, Pearson's correlation coefficient r = -0.08162, t = -2.84619, and for a two-tailed test of significance, p = .0045 (still significant but close to the 95% confidence level).

Here is the ln(alpha) scatterplot, minus the High-Bush precincts:



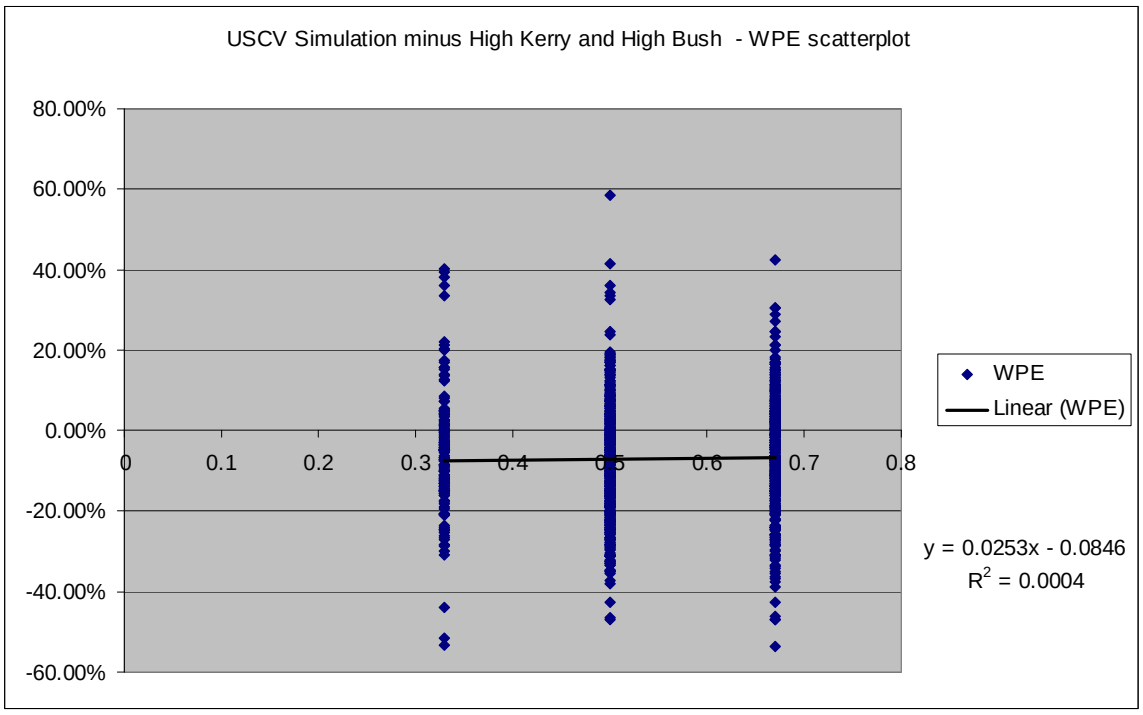
(Above) O'Dell USCV Simulation – ln(alpha) regression, excluding 40 High-Bush precincts.

For the regression of ln(alpha) against partisanship for the sample minus the 40 High Bush precincts, Pearson's correlation coefficient r = 0.004853, t = .168689, and for a two tailed test of significance, p = .8661 (not significant).

These results may be readily reproduced by removing the data for the 40 High-Bush precincts from the "Scatterplots" tab in the USCV O'Dell simulator spreadsheet.

Sensitivity Analysis: Remove High Bush and High Kerry Precincts

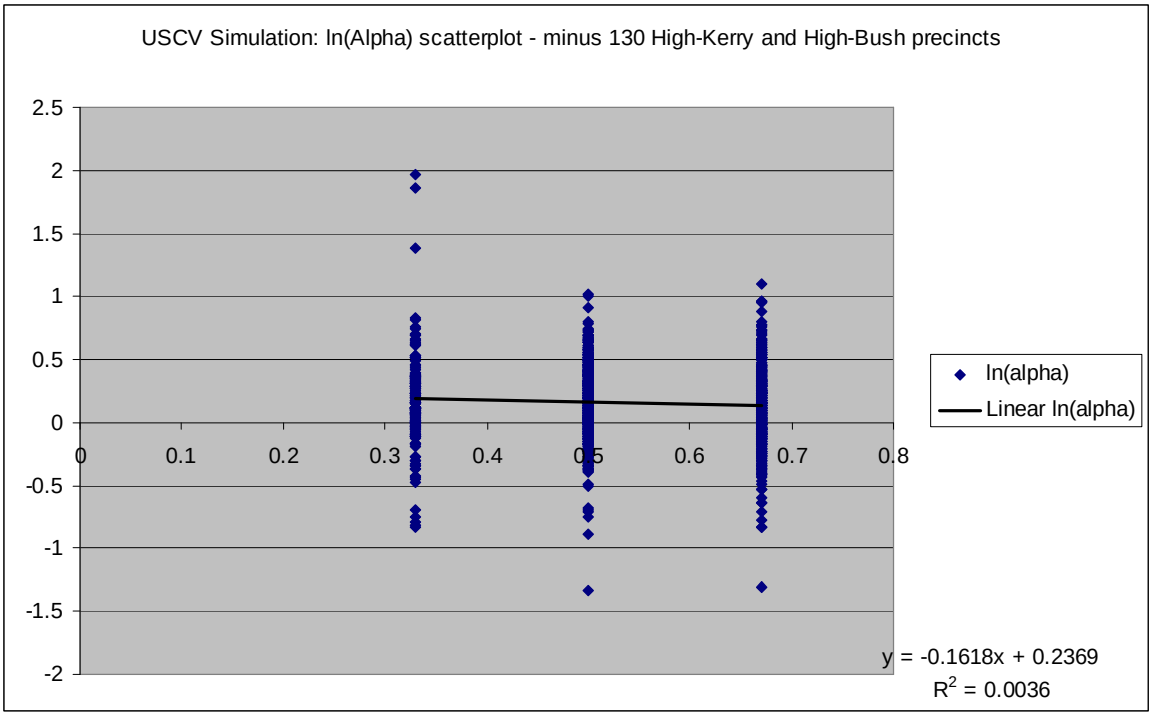
When the 130 High Bush and High Kerry precincts are removed, we see the following WPE and ln(alpha) scatterplots. The WPE regression line is visually flat, and the alpha regression line shows slightly lower response bias with increasing partisanship. Since this data includes 89% of the overall sample, the lack of a strong correlation one way or the other is of interest.



(Above) O'Dell USCV Simulation WPE regression excluding 130 High-Bush and High Kerry precincts

For the regression of WPE against partisanship excluding the 130 High Bush and High Kerry precincts, Pearson's correlation coefficient r = .020917, t = .669533, and for a two tailed test of significance, p = .5033 (not significant).

The corresponding ln(alpha) scatterplot:



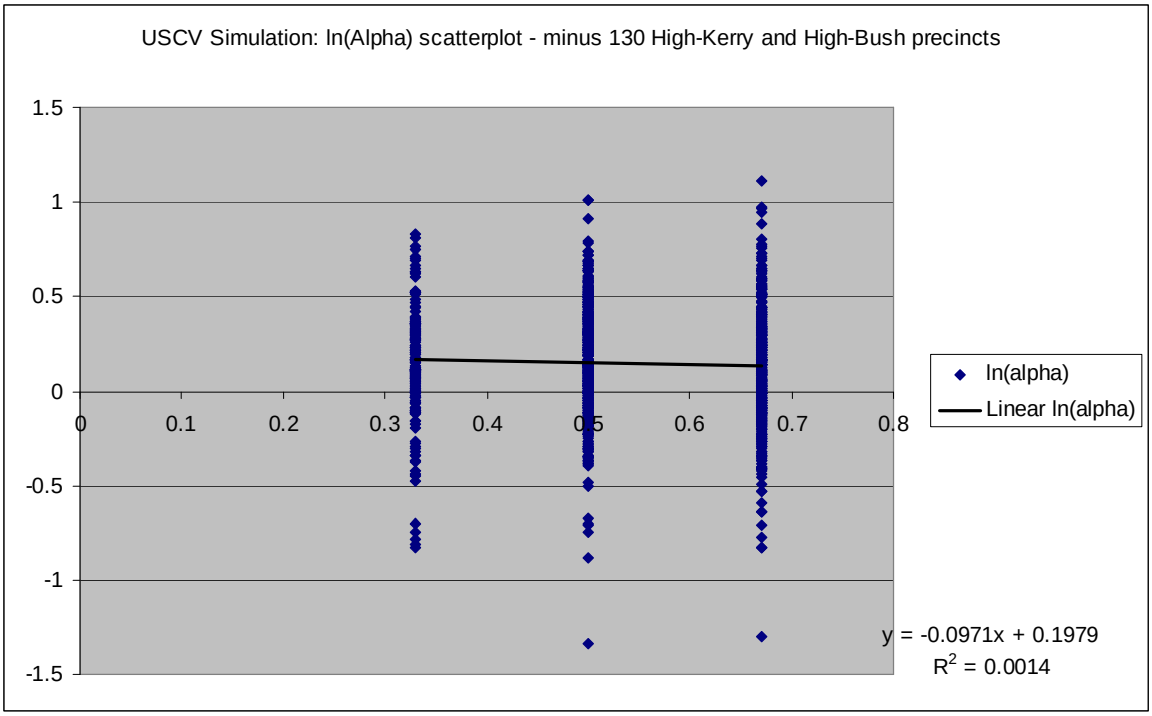
(Above) O'Dell USCV Simulation ln(alpha) regression excluding 130 High-Bush and High Kerry precincts ln(alpha) for B=33% is (-.1618)(.33)+.2369 = .1835 and ln(alpha) for B=66% is (-.1618)(.67)+.2369 = .1301. At the midpoint B=50% it is (-.1611)(.5)+.2369 = .1564

For the regression of ln(alpha) against partisanship excluding the 130 High Bush and High Kerry precincts, Pearson's correlation coefficient r = -0.06019, t = -2.01627, and for a two tailed test of significance, p = .044 (significant but close to the 95% confidence level).

There is a small statistically significant relationship between ln(alpha) and partisanship in the 1120 precincts between B=20% and B=80%. where ln(alpha) decreases with increasing Bush vote.

These results may be readily reproduced by removing the data for the 130 High-Bush and High Kerry precincts from the "Scatterplots" tab in the USCV O'Dell simulator spreadsheet.

This observed correlation may to be due to a few very large outliers in the leftmost group of precincts. When the most extreme three outliers in that group of 165 precincts are removed, the following chart is obtained:



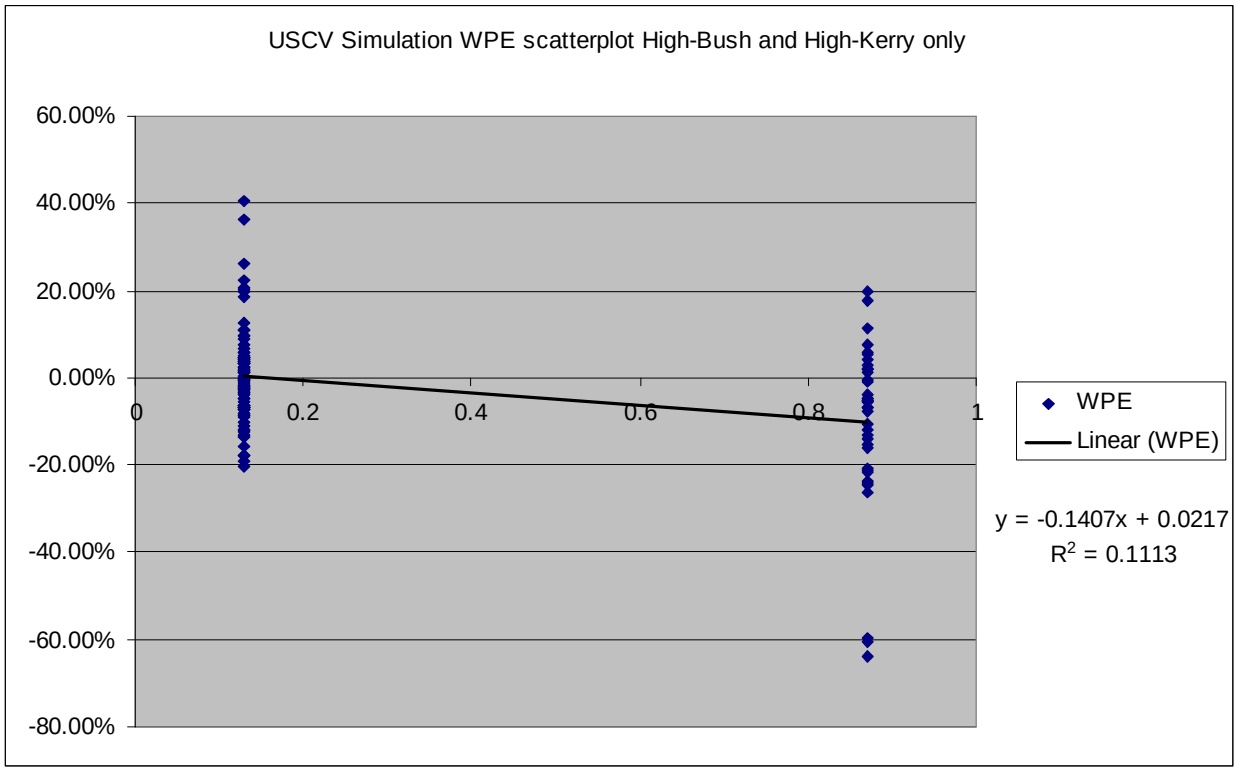
(Above) O'Dell USCV Simulation ln(alpha) regression excluding 130 High-Bush and High Kerry precincts
Note that *when the three extreme outliers in the leftmost group of precincts* are removed,
the correlation is no longer statistically significant.

After removing the three outliers with the most extreme positive alpha in the Bush = 33% category, the regression over the remaining 1117 precincts is no longer statistically significant. In the regression of ln(alpha) with those 1117 precincts, Pearson's coefficient of correlation r = -0.03731, t = -1.24686, and for a two-tailed test of significance, p = .2127 (not significant).

One final sensitivity analysis is called for: removing the precincts in the middle and examining only High Bush and High Kerry precincts.

Sensitivity Analysis: Include only High Kerry and High Bush Precincts

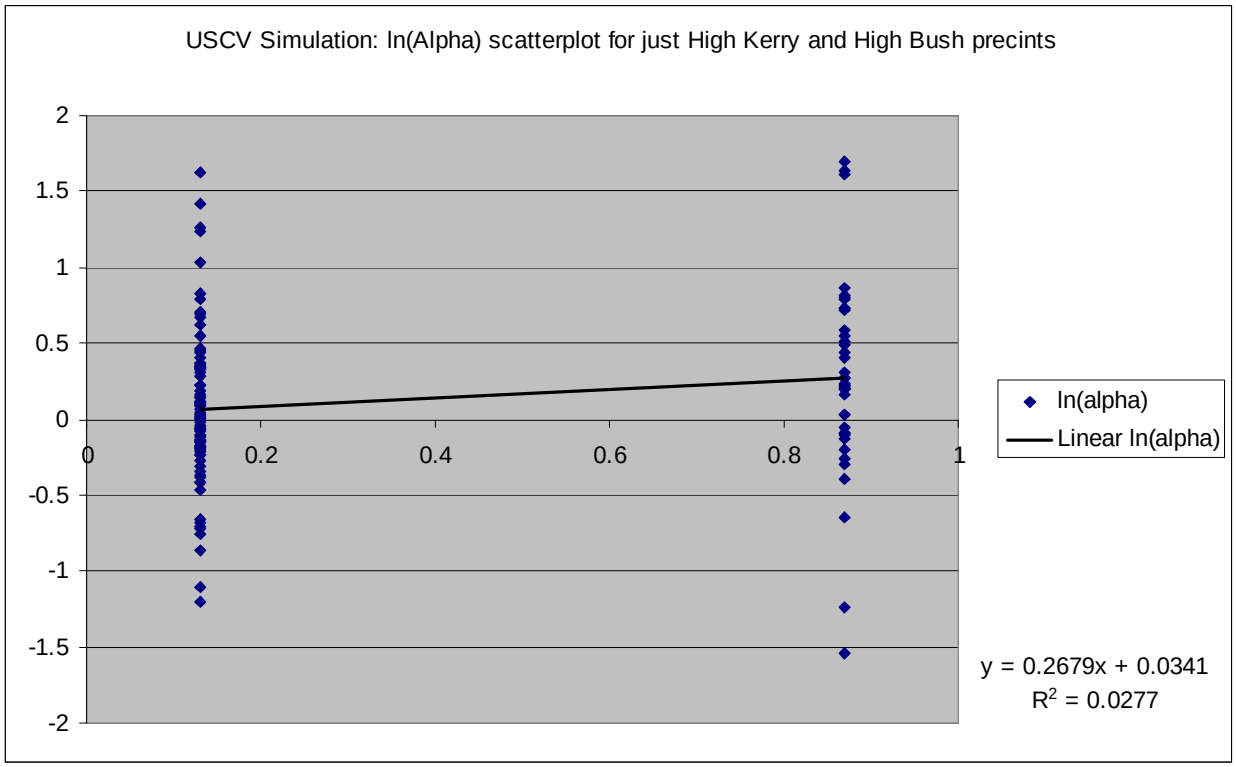
After removing the middle 1120 precincts, the following scatterplot includes the regression of WPE against partisanship for just the 130 High Bush and High Kerry precincts.



USCV Simulation – Simulated WPE by partisanship for just the 130 High Bush and High Kerry precincts
Note steep slope about twice the magnitude of the overall sample, and that the regression line joins Mean WPE for High Kerry precincts (+0.3%) with Mean WPE for High Bush precincts (-10%)

Note that the WPE regression line's endpoints are the aggregate mean WPE for High Kerry (+0.34%) and aggregate mean WPE for High Bush (-10.0%). Of course, this is statistically significant: Pearson's coefficient of correlation r = -0.33362, t=-4.00382 and for a two tailed test of significance, p = .0001 (significant).

On the following page, the regression of ln(alpha) for the 130 High Bush and High Kerry precincts is shown. A marginally statistically significant relationship exists, with Pearson's coefficient of correlation r = .166378, t = 1.908953, and for a two-tailed test of significance, p = .0585 (technically not significant, but close to the p=.05 cutoff for the 95% confidence level).

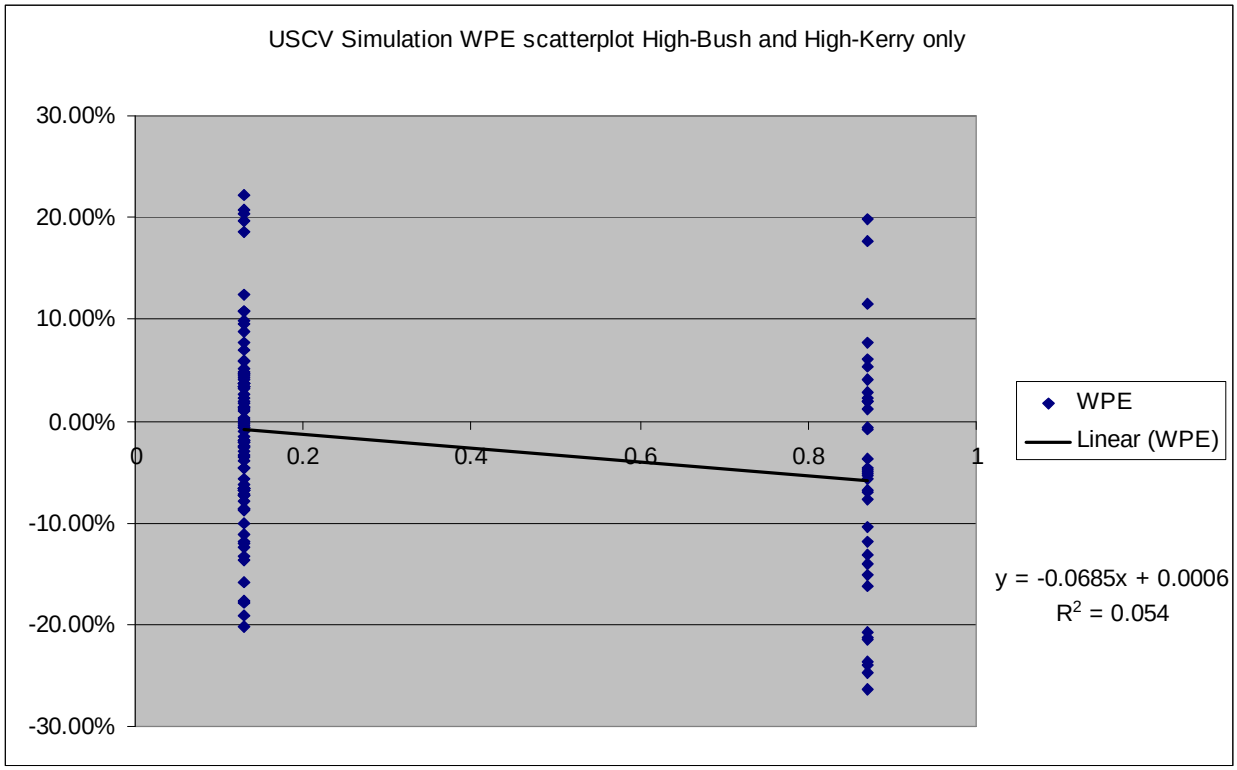


USCV Simulation – Simulated WPE by partisanship for just the 130 High Bush and High Kerry precincts
There is a borderline statistically-significant positive correlation for ln(alpha)
p = .0585 for n=130

However, that there are very positive WPE outliers in High-Kerry territory, and very large negative outliers in High Bush territory.

It is interesting to see what happens if an equivalent number of the most extreme outliers[12] are removed from both populations.

[12] The outliers in this simulation dataset are only a visual approximation of the number and magnitude of the outliers shown in the Mitofsky scatterplot. However, since the simulation dataset, imperfect as it is, still matches the aggregate mean WPE, median WPE and participation rates reported in the January E/M report, the overall characteristics of the simulation sample appear to match the E/M data in aggregate. It would of course be interesting to see whether the corresponding sensitivity analyses, if performed on the actual E/M data, produces the same results as shown on the next page.

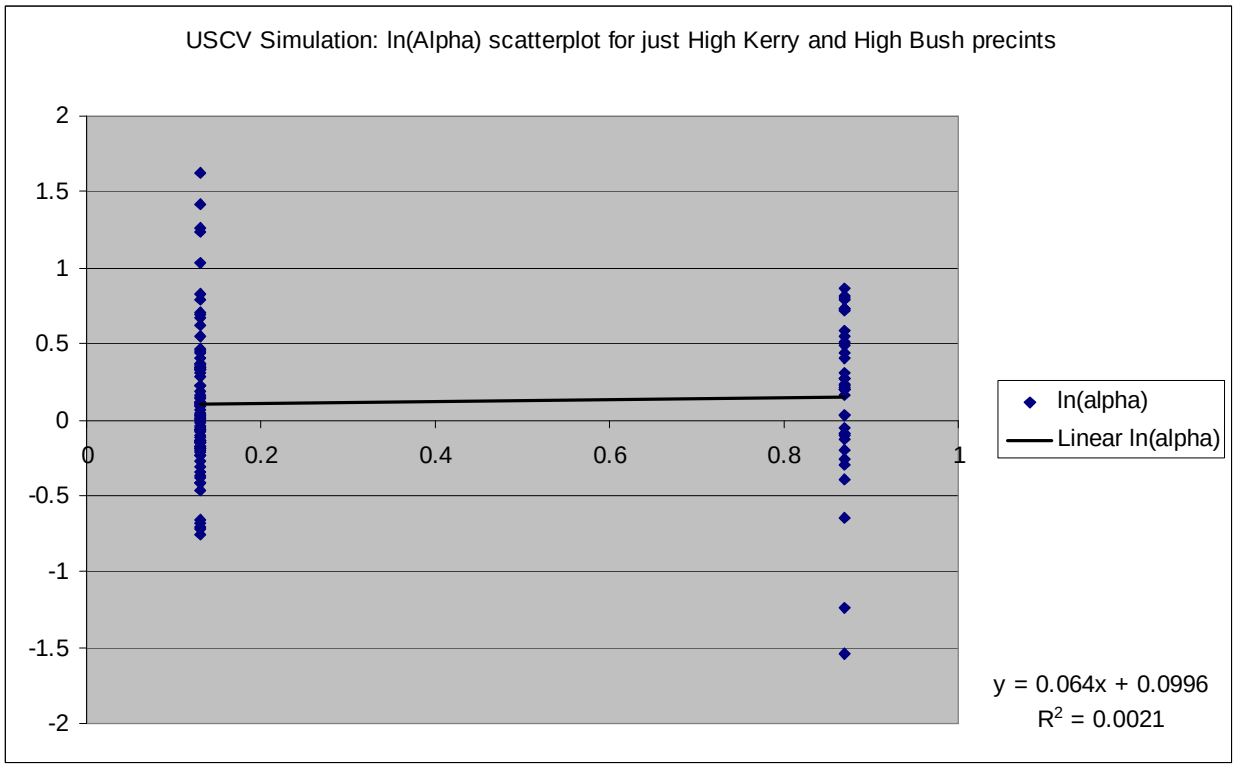


USCV Simulation of 124 High Bush and High Kerry precincts (the original set of 130, minus the extreme three outliers in each category) – note that the WPE regression line now joins the **median** WPE for each category, rather than the mean. For B=87%, WPE = -0.059 (or -5.9%, *extremely close* to E/M's reported *median* WPE of -5.8%).

For this regression, Pearson's coefficient of correlation r = 0.23244, t = -2.6397, and for a two-tailed test of significance, p < .0094 (significant).

When only six extreme outliers are removed – three in each category – the slope of the WPE regression line becomes noticeably shallower. But more importantly, it now intercepts B=87 at -5.9% (the *median* WPE value that E/M reported) rather than at -10.0% (the *mean* WPE value that E/M reported). This illustrates a "sensitive dependence" of aggregate WPE statistics on just a handful of outliers in High Bush and High Kerry territory.

The corresponding scatterplot of ln(alpha) for the same set of 124 High Bush and High Kerry Precincts is shown on the next page.



USCV Simulation of 124 High Bush and High Kerry precincts (the original set of 130, minus the extreme three outliers in each category) – note that the correlation of partisanship and ln(alpha) is no longer statistically significant

For the regression of ln(alpha) against partisanship for the sample of 124 High Bush and High Kerry precincts, Pearson's correlation coefficient r = 0.04554, t = .503526, and for a two-tailed test of significance, p = .6155 (not significant).

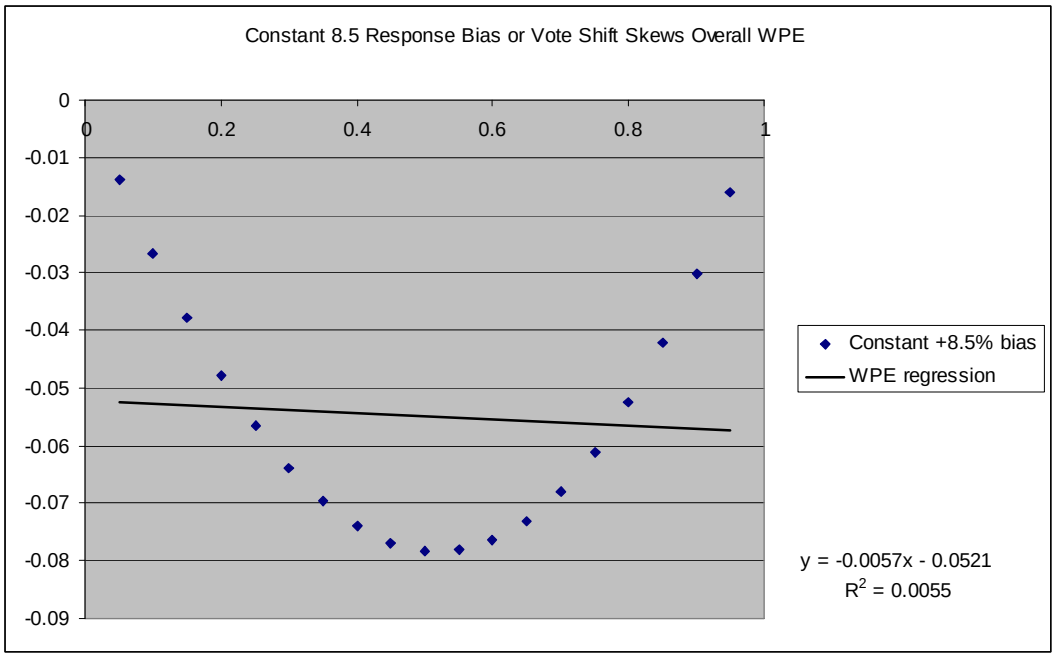
No subset of the precinct level data seems to have any statistically-significant correlation between partisanship and response rate. A few marginally-significant relationships appear on further analysis, but the perceived correlation is sensitively-dependent on just a few outliers, and when those are removed, the significance of the correlation disappears.

While the few, extreme outliers are certainly worthy of further investigation, overall, there is *simply no evidence of any significant correlation of exit poll response rates with partisanship.*

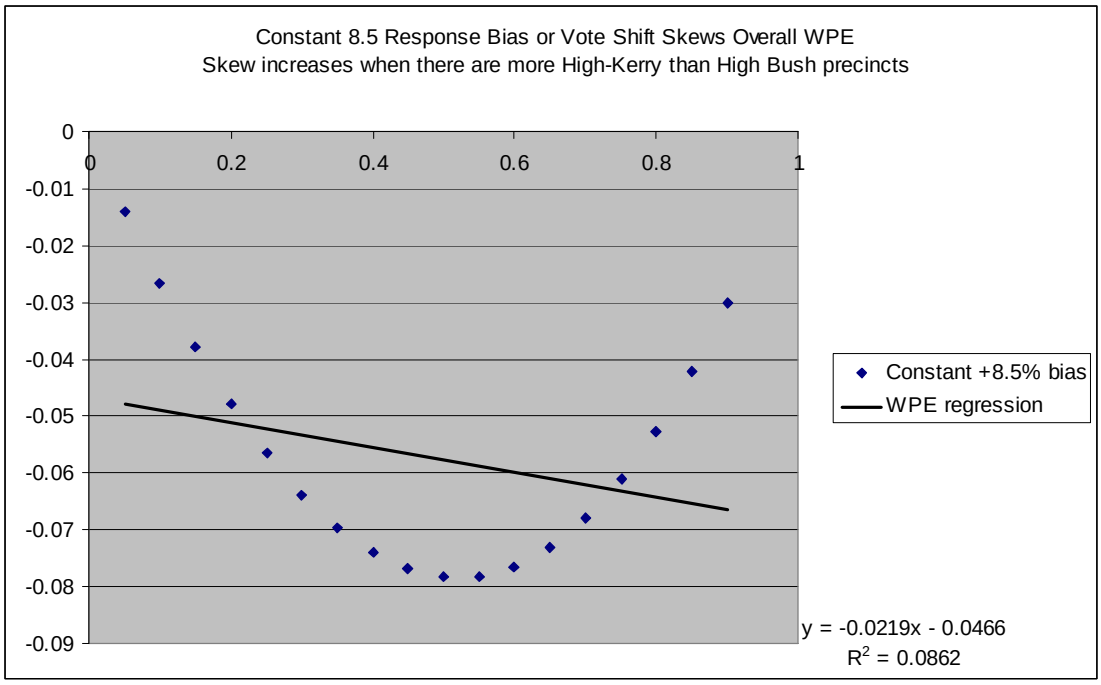
Statistically-significant correlations of partisanship and WPE *were* observed. However, WPE is known to be skewed as a function of partisanship in the presence of uniform mean response rate differences. Alpha and ln(alpha) remove that skew and allow us to correlate response bias against partisanship without distortion. The fact that the USCV O'Dell simulation reproduces highly differential response rates when simulating aggregate WPE – but shows no such effect when simulating precinct level alpha – indicates that aggregate WPE is a highly misleading metric of response bias, and the USCV Working Paper analysis that is based upon it is fundamentally flawed.

Appendix 4: Why alpha corrects the inherent bias in WPE

The following charts illustrate how a constant bias (due to either differential poll response or vote shifting) applied to all partisanship categories "skews" WPE.



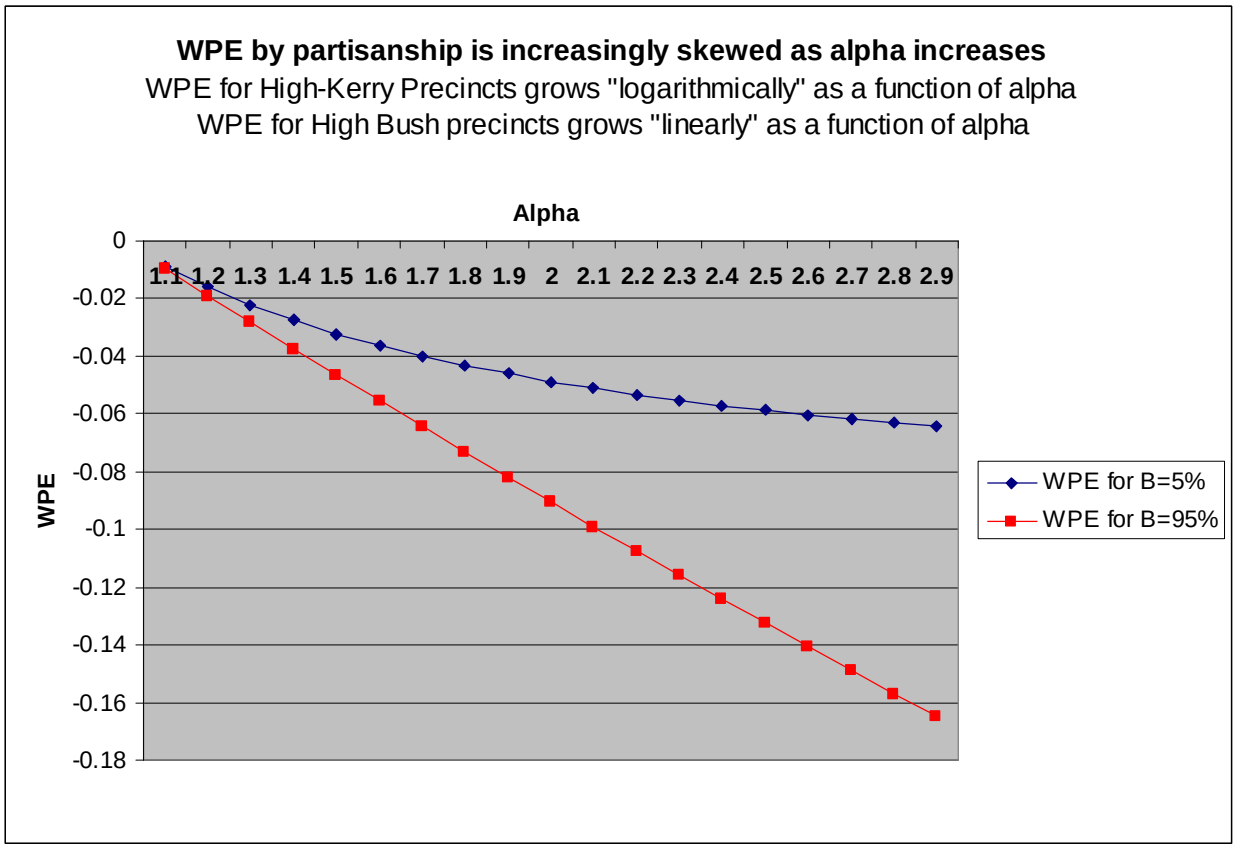
In the E/M precinct samples, there were fewer High Bush than High Kerry precincts. This would tend to introduce additional skew, as illustrated below:



The previous two charts were simple illustrations to show that a skew exists. The magnitude of the skew effect is unmistakable, but smaller in either case than the 7.5% skew Mitofsky reported for the E/M full scatterplot, and that I reproduced in the USCV O'Dell simulation. So, isn't it possible that the authors of the Working Paper are correct and that WPE is still a rather good measure of response bias?

No. The two charts above were idealized simplifications designed to illustrate the paradox of constant bias creating skewed WPE. They do not reflect the extremely high variance found in the E/M data, with WPE values ranging from -60% to +50%.

Large values of WPE introduce even more apparent skew relative to alpha, as shown in the chart below.



Recall that alpha is the ratio of Kerry poll participation to Bush exit poll participation, or [(Kerry poll participants/Kerry actual vote)] / [(Bush poll participants/Bush actual vote)]

The fact that *the same value of alpha* – or relative poll participation – produces WPE in High-Bush precincts that is very disproportionate to the WPE it produces in High-Kerry precincts is I think the definite argument for using alpha, and for discarding WPE, as a measure of exit poll bias.

Alpha detects both differential response bias and vote shifting, but we can't determine which of those two causes may have occurred by exit polls alone.