Edison/Mitofsky Exit Polls 2004: differential non-response or vote-count corruption?

Updated 27th April 2005. The previous version is archived at http://www.geocities.com/lizzielid/WPEpaperARCH.pdf

Abstract

During the 2004 US presidential election, exit polls initially appeared to over-estimate the proportion of votes cast for John Kerry, the Democratic challenger, relative to the proportion cast for the Republican president, George W. Bush. Because the margin was so close, the discrepancy between the early exit poll indicators and the result made the difference between indications that was Kerry ahead, and a result that gave Bush a victory, and helped fuel widespread suspicion that the error had arisen not from errors in the exit poll, but from errors in the counting of the votes. In January, 2005 the polling companies, Edison Media Research and Mitofsky International, issued a report in which they concluded that the errors arose from differential response rates in the two groups of voters, exacerbated by problems with voter sampling methodology. In March 2005, a group of academics working for the US Counts Votes (USCV) National Election Data Archive Project, issued a paper on the internet detailing analyses of data provided in the Edison-Mitofsky report which they claim indicate that the patterns of error reported are more consistent with a hypothesis of vote-corruption than one of differential non-response. One of their claims is rooted in the apparent differentials between "Within-Precinct Errors" (WPEs) found in highly partisan precincts and those found in more evenly divided precincts. However, the WPE as a measure is itself confounded by precinct partisanship. This paper addresses problems presented by the use of the WPE as a dependent measure, particularly when precinct partisanship is a predictor variable. It describes a computation whereby an unconfounded index of sampling bias can be retrieved from the data, models the nature of the confound introduced into analyses by the WPE, discusses implications for the inferences drawn by USCV from the Edison-Mitofsky data, and makes a recommendation for further investigations into the factors that led to discrepancies between exit poll predictions and election results.

Exit poll methodology in US presidential elections normally involves two levels of sampling, both of which are vulnerable to error. Firstly, from all the precincts in each US state, a selection of precincts is sampled, which, it is hoped, will be representative of the totality of precincts in that state. Assessing how representative a sample the selected precincts actually were can be relatively easily ascertained after the election simply by comparing the estimate made by the actual results from the sampled precincts with the actual results from the totality of precincts. However, a second level of sampling occurs within each of the sampled precincts, namely the sampling of the voters themselves. As this is done at the precincts on election day, and involves human interactions, it is potentially more vulnerable to bias.

During the 2004 US presidential election, exit polls were conducted for the National Election Pool (NEP) by two companies, Edison Media Research and Mitofsky International, commonly known as Edison-Mitofsky (E-M.) In an evaluation released in January 2005, E-M conclude that their precinct sampling was good, and closely predicted the counted result, the error being in the direction of a slight over-estimate of George W Bush's share of the vote.[1] However, the sampling of the voters within each precinct was associated with a large and significant 6.5 percentage point over-estimate of John Kerry's share of the vote.[2] This source of error is referred to as the Within-Precinct Error (WPE), and represents the difference between the predicted and the actual percentage margin between the candidates for that precinct. Two hypotheses have been advanced to account for the magnitude and direction of the large mean WPE in 2004. One is that a greater proportion of Kerry voters were sampled than Bush voters.[3] The other hypothesis is that a greater proportion of Kerry votes than Bush votes went uncounted, or alternatively that extra Bush votes were somehow added to the total. This second hypothesis has been at the heart of allegations that the exit polls in 2004 are evidence of election fraud.

In their evaluation, E-M summarize factors that appeared to contribute to the observed WPE. They conclude that:

[1] Edison Media Research and Mitofsky International. 2005, p.28
[2] Edison Media Research and Mitofsky International. 2005, p.31
[3] Edison Media Research and Mitofsky International. 2005, p.31

> While we cannot measure the completion rate by Democratic and Republican voters, hypothetical completion rates of 56% among Kerry voters and 50% among Bush voters overall would account for the entire Within Precinct Error that we observed in 2004.[4]

They note that the overestimate of the Kerry vote was greater wherever factors were present that were likely to compromise random sampling – for example when the interviewers were obliged to be positioned more than 25 feet from the precinct, where precincts were large, and where a smaller proportion of all voters were selected.[5]

However, in March 2005, a group of academics working for the US Counts Votes (USCV) National Election Data Archive Project issued a paper entitled *Analysis of the 2004 Presidential Election Exit Poll Discrepancies* [6] in which they analyse the data given in the E-M evaluation, and conclude that what they term the "reluctant Bush responder" hypothesis is not supported by the E-M data as given in the E-M evaluation. They then consider a plausible alternative hypothesis entitled: "reluctant Bush responder in mixed political company", which states that Bush voters may have been more reluctant to respond in predominantly Kerry-supporting precincts.

In the E-M report, a table[7] gives the mean and median WPEs for precincts with different degrees of partisanship as indicated by the final vote count. Five categories are given: highly Democratic precincts; moderately Democratic precincts; even precincts; moderately Republican precincts; and highly Republican precincts. Their table is reproduced as Table 1 below.

[4] Edison Media Research and Mitofsky International. 2005, p.31.
[5] Edison Media Research and Mitofsky International. 2005, p.47
[6] Mitteldorf et al. 2005
[7] Edison Media Research and Mitofsky International. 2005, p.36.

Precinct partisanship	mean WPE	median WPE	mean Abs (WPE)	N
High Dem (Kerry>=.80)	0.3	-0.4	8.8	90
Mod Dem (0.60<=Kerry<.80)	-5.9	-5.5	13.4	165
Even (0.40<=Kerry<.60)	-8.5	-8.3	15.2	540
Mod Rep (0.20<=Kerry<.40	-6.1	-6.1	13.2	415
High Rep (0.00<=Kerry<.20)	-10.0	-5.8	12.4	40

Table 1

Mean and median WPE, mean absolute WPE and number of precincts (N) by precinct partisanship, from page 36 of the Edison/Mitofsky report (means and medians are graphed in Figure 4.)

In the sign convention used in the E-M report, a negative WPE indicates an overestimate of the proportion of votes for the Democratic candidate, and a positive WPE indicates an over-estimate of the proportion of votes for the Republican candidate. From Table 1 it can therefore be seen that the most negative mean and median WPEs are in the "high Republican" category of precincts. Thus, contrary to the "reluctant Bush responder in mixed political company hypothesis", the greatest overestimates of the Kerry vote appear to have occurred in highly Republican precincts.

From the table, the USCV authors calculate the relative response rates of the two groups of voters, and conclude that in order to satisfy the "reluctant Bush responder in mixed political company" hypothesis, implausible patterns of non-response rates have to be inferred, and that the data are more consistent with a "Bush strongholds have more vote-count corruption" hypothesis.[8]

The WPE as a confounded variable

However, both the E-M WPE analyses given in the E-M report, and the inferences drawn from the WPE data in the E-M report by the USCV authors, are potentially confounded by the nature of the WPE as a variable. To compute the WPE, the actual margins are calculated by subtracting the proportion of votes counted for the Republican candidate from the proportion of votes counted for the Democrat. The predicted margins similarly are calculated by subtracting the predicted proportion of

[8] Mitteldorf et al. 2005

votes that will be counted for the Republicans from the predicted proportion that will be counted for the Democrat. The predicted margin is then subtracted from the actual margin. It can be expressed algebraically as in Equation 1:

$$\textit{Equation 1:}^{9}\ WPE = (D_V - R_V) - (D_P - R_P)$$

where D_V and R_V are the actual proportions of the Democratic and Republican votes in that precinct respectively, and D_P and R_P are the predicted proportions, based on the sampled respondents.

Note that in this equation, the WPE is a function of the actual proportion of votes cast for each candidate. Therefore, any use of the WPE as a dependent variable in analyses in which the proportions of votes cast for each candidate appears as a predictor variable is potentially confounded. Indeed, it will be shown later in this paper that, for a given level of bias, the WPE will tend to be larger in precincts where the support for the candidates is fairly even than in more highly partisan precincts, and also larger in partisan precincts in which the relatively under-polled group of voters are in the majority than in precincts in which they are in the minority. In order to investigate factors affecting polling bias, we therefore need a measure of bias that is independent of precinct partisanship. The following section of the paper will describe a set of algebraic manipulations that produce such a measure.

A bias index

We can express differential non-response, or sampling bias, as the ratio between the achieved sampling rate[10] of the voters who voted for the Democratic candidate and the achieved sampling rate of the voters who voted for the Republican candidate. If there is no bias, the achieved sampling rate will be the same for both groups of voters, giving a 1:1 ratio. If the achieved sampling ratio is not 1:1, we will have sampling bias. In the mathematical discussion that follows, I have expressed the ratio as one in which the numerator is the sampling rate of voters who voted for the Democrat, and

[9] Note that equations have been re-numbered in this update.

[10] By "achieved sampling rate" I mean the number of responses recorded for a candidate and the number of votes counted for that candidate. If voter sampling protocol is accurate, and every Nth voter is approached, the ratio between "achieved sampling rates" for each candidate should be the same as the ratio between the response rates for each candidate. However, where deviations from strict Nth voter sampling protocol exacerbates the differential, these will not be identical. For the purposes of this paper, however, I will assume strict protocol when using the term "response rate."

the denominator is the sampling rate of voters who voted for the Republican. This ratio I will call α. It is defined in Equation 2:

$$\textit{Equation 2: } \alpha = \frac{S_D}{S_R}$$

where S_D is the achieved sampling rate of the voters who voted for the Democrat and S_R is the achieved sampling rate of the voters who voted for the Republican.

We can re-write Equation 2 to express S_D in terms of S_R and alpha:

$$\textit{Equation 3: } S_D = \alpha \times S_R$$

Thus if there are 200 voters in a precinct, 140 of whom voted for the Democrat (70%), and around 50% of voters in the precinct are approached, and all respond, approximately 70 "Democrat" responses should be recorded. Similarly, if around 50% of the 60 voters who voted for the Republican candidate are polled, around 30 "Republican" responses will be recorded, giving a total of around 100 responses. Of these responses, around 70 will be "Democrat", giving an estimate of 70% for the actual vote, and around 30 will be "Republican" giving an estimate of 30%. As this is the same as the proportions in the counted vote, the exit poll will correctly estimate the result, subject to a computable degree of sampling error.

However, a factor that results in a systematically reduced proportion of "Democrat" voters sampled may lower that 50% to, say 40%. The proportion has thus been reduced by a factor of 0.8 (0.8*50% = 40%.) In this case we will record responses from only 40% of the 140 "Democrat" voters, giving only 56 interviews instead of 70. If no such reduction factor applies to the Republicans, we will have still have the 30 "Republican" responses, giving us a total of 56 + 30 responses, i.e. a sample of 86. This will lead us to estimate that 56 out of 86 voters voted for the Democrat i.e. only 65%, and that 30 out of 86 voted for the Republican, i.e. 35%. As the margin between the candidates in the counted vote is 40 percentage points, but our estimated margin is only 30 percentage points, we will have a WPE of 10, and an overestimate of the Republican's vote.

The WPE equation can thus be written:

$$\textit{Equation 4: } WPE = (D_V - R_V) - \frac{\alpha D_V S_R n - R_V S_R n}{\alpha D_V S_R n + R_V S_R n}$$

where α is the bias factor, S_R is achieved sampling rate for the voters who voted for the Republican candidate, and n is the total number of voters in the precinct. As both the achieved sampling rate for voters who voted Republican (S_R), and the number of voters in the precinct (n) cancel out of the expression, this can be re-written:

$$\textit{Equation 5: } WPE = (D_V - R_V) - \frac{\alpha D_V - R_V}{\alpha D_V + R_V}$$

An α of 1 will indicate that the differential is zero, i.e. a Democratic to Republican response rate ratio of 1:1. An α value equal to 2 will indicate that the response rate for Democrats was double that of the Republicans (2:1); while an α of .5 will indicate a response rate for Democrats that was only half of that of the Republicans (1:2.) Defining alpha with Republican voters as the numerator will simply reverse the way the ratio is expressed.

Solving for α, we obtain:

$$\textit{Equation 6: } \alpha = \frac{R_V(1 + D_V - R_V - WPE)}{D_V(1 - D_V + R_V + WPE)}$$

which can also be written:

$$\textit{Equation 7: } \alpha = \frac{R_V(1 + M - WPE)}{D_V(1 - M + WPE)}$$

where M is the final margin, expressed as the proportion of Democratic votes cast minus the proportion of Republican votes cast.

However, because alpha is a ratio, with the sampling rate for one group of voters as the numerator and the sampling rate for the other group of voters as the denominator, α will be non-symmetrical around 1 (=no bias.) To generate a symmetrical "bias index", therefore, we take a log of α.

Equation 8: $\text{Bias index} = \log_e \left[\frac{R_V (1 + M - WPE)}{D_V (1 - M + WPE)} \right]$

A value of 0 for this bias index represents zero bias (Democratic:Republican sampling ratio = 1:1), a value of +0.693 represents a Democratic:Republican sampling of 2:1, and a value of -0.693 represents Democratic:Republican sampling of 1:2. A geometric increase or decrease in the value of α will result in a linear increase or decrease in the value of the bias index. This index thus provides a linear measure of any underlying sampling bias that may have contributed to an over- or under-estimate of the predicted margin between the candidates for that precinct, without being confounded by the partisanship of the precinct itself.

Moreover, if the overall response rate (r) for a precinct is known, the response rates for each group of voters can be derived α from as follows. The response rate for voters who voted for the Republican (r_R) is given by Equation 9:

Equation 9: $r_R = \frac{r}{\alpha D_v + R_V}$

From this, the response rate for voters who voted for the Democrat (r_D) can be derived, using Equation 10:

Equation 10: $r_D = \alpha \times r_R$

Given the confound between WPE and precinct partisanship it is perhaps surprising that it has been widely used as a dependent variable in analyses of the performance of exit polls.[11] It would certainly be of interest to know what the values in Table 1 would be if they represented the means and medians of a bias index that was uncontaminated by partisanship, such as that formulated above, as opposed to the mean and median WPEs given. However, in the absence of these, it may be of use to delineate the nature of the relationship between WPE and partisanship, and thus the nature of the likely confound. It can be examined most clearly if we take Equation 1, and ignore votes for third party candidates. In such a case the proportion of Republican votes in a precinct will equal one minus the proportion of Democratic

[11] Merkle and Edelman 2002; Merkle and Edelman, 2000

votes. We can therefore substitute (1-D_V) for R_v in Equation 1, and express WPE in terms of D_V and α. This gives us:

$$\textit{Equation 11}: WPE = 2 \times \frac{D_V{}^2(1-\alpha) - D_V(1-\alpha)}{D_V(1-\alpha) - 1}$$

It can be seen from this Equation 11 that when $\alpha = 1$ (and the bias is therefore zero) the coefficients of each D_V term becomes 0, and WPE is equal to zero. It can also be seen that that the relationship between WPE and D_V is given by a quadratic function, divided by a linear function. At high and low values of D_V, WPE will therefore tend to be greater than at mid-values of D_V. Moreover, however, as α departs from 1 (representing either under- or over-sampling of the Democratic vote relative to the Republican vote), not only does the coefficient of both the D_V term and the $D_V{}^2$ term become larger, making the quadratic curve more marked, but the denominator changes in value, altering the symmetry of the function.

Implications for WPE analyses

The equations given above not only allow us to calculate an index of bias that is free from the confound of precinct partisanship, but also allow us to determine the nature of the confound when attempting to investigate differential degrees of bias in precincts with differing degrees of partisanship. Figure 1 illustrates the WPEs that would theoretically result from a uniform and extreme Democratic:Republican response rate ratio of 2:1.

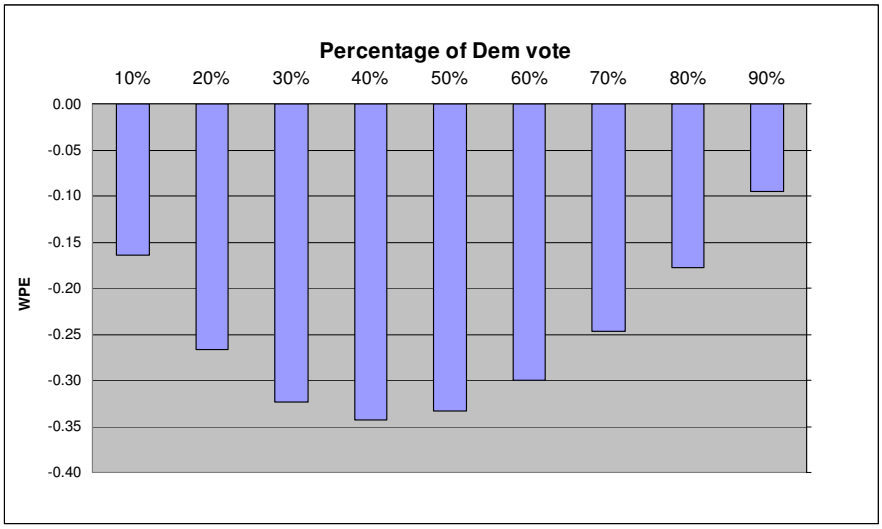


Figure 1

Theoretical values for mean WPEs for precincts in partisanship categories ranging from high Republican on the left (10% Democrat) to low Republican on the right (90% Democrat), for a uniform extreme response rate ratio of 2:1.

Note that the sampling bias in each of these theoretical precincts is identical: the response rate for Democrats was twice that for Republicans. Nonetheless, the resulting WPEs are different for different categories of precinct. For precincts in which the partisanship is fairly even, the WPE is strongly negative. For more partisan precincts it is less negative, although it is more negative for strongly Republican precincts (on the left of the chart) than for strongly Democratic precincts (on the right of the chart.) In other words, relative under-sampling of the minority voter group has less impact on the WPE than relative under-sampling of the majority group.

Figure 2 shows a three-dimensional plot, showing how the curve in Figure 1 changes as the differential response rate changes from an extreme over-sampling of Democrats (3:1 bias) through lessening degrees of Democratic over-sampling and beyond, to an extreme and opposite over-sampling of the Republican voters (1:3 bias.)

The red colours in the figure indicate negative WPEs, in which voters who voted for the Democrat were over-sampled in the poll, resulting in a "red shift",[12] in the counted

[12] Note that the mathematics of this analysis is "fraud-neutral" – under counting of votes for one party would give the same alpha as over-sampling of that party's voters.

vote. The blue colours indicate positive WPEs, in which voters who voted for the Republican were over-sampled, leading to a "blue shift" in the counted vote. Note that the asymmetric "quadratic" function given by Equation 11 results in a series of skewed "quadratic" curves. Both the sign of the quadratic term and the direction of the skew is equal and opposite for opposite directions of bias.

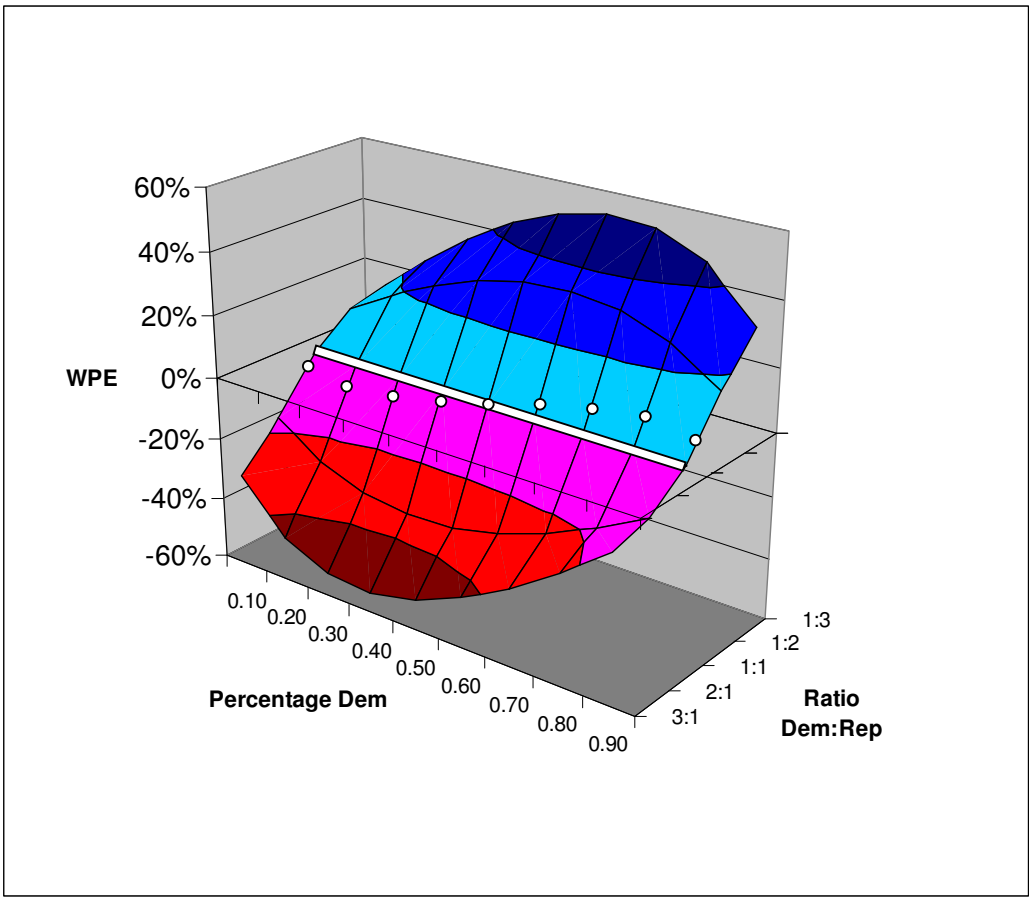


Figure 2

3D plot showing partisanship from left to right, sampling rate ratio from front to back, and WPE on the vertical axis. Negative WPEs are shown in red, positive WPEs in blue. The white dots represent the mean WPEs for each precinct category, and the white line represents the median, when there is no net bias (as much bias in one direction as the other.)

It is unlikely that any real-life sampling bias will be uniform. It is more likely that in some precincts the bias will be greater than in others, and that in some precincts a bias will occur that is in the opposite direction. It is also theoretically possible that biasing factors could have an overall mean of zero, resulting in a mean WPE of zero. This

would not mean that no biased sampling had occurred. It would simply mean that the sampling biases had cancelled each other out across the whole sample of precincts. If, for example, the precincts modelled in Figure 2 were all real precincts, one for every point on the surface of the figure, the average WPE for the entire sample of precincts would be zero, as there are as many precincts in which Republican voters are over-sampled, as precincts in which Democrat voters are over-sampled. However, the mean WPEs for any given category of precinct would not necessarily be zero. The mean WPEs for each precinct category are marked in Figure 2 by the small white dots. The dots on the left are dragged downwards by the deeper "red shift" precincts below, while the dots on the right are pulled upwards by the higher "blue shift" precincts above. Thus the mean WPEs in Republican precincts suggest that the Democrat vote has been over-estimated, and the mean WPEs in Democratic precincts suggest that the Republican has been over-estimated. The median WPEs for each precinct category (represented by the white line), however, remains zero.

I therefore conducted a computer simulation to model the effects of a normal (Gaussian) distribution of biasing factors randomly distributed throughout a population of "precincts" with differing degrees of partisanship. I modelled the resulting mean WPEs, the median WPEs, and the mean "bias index" values as computed by the formula given above. Nine categories of precincts were modelled, ranging from precincts with 10% "Kerry" support, to precincts with 90% "Kerry" support, in a race with only two "candidates": "Kerry" and "Bush." The simulation sampled a Gaussian distribution of response rates for each type of voter, under two conditions. In the first condition, which I will refer to as the "no net bias" condition, the mean of each distribution was 50%, with a standard deviation of 10 percentage points, and a range from 25% to 75%. Thus, the mean differential response rate across all "precincts" was zero. However, in half the precincts, the response rate for "Kerry voters" was greater than for "Bush voters", and vice versa. The response rate differentials thus also had a Gaussian distribution, with a mean of zero. The standard deviation of the differential was 14 percentage points.

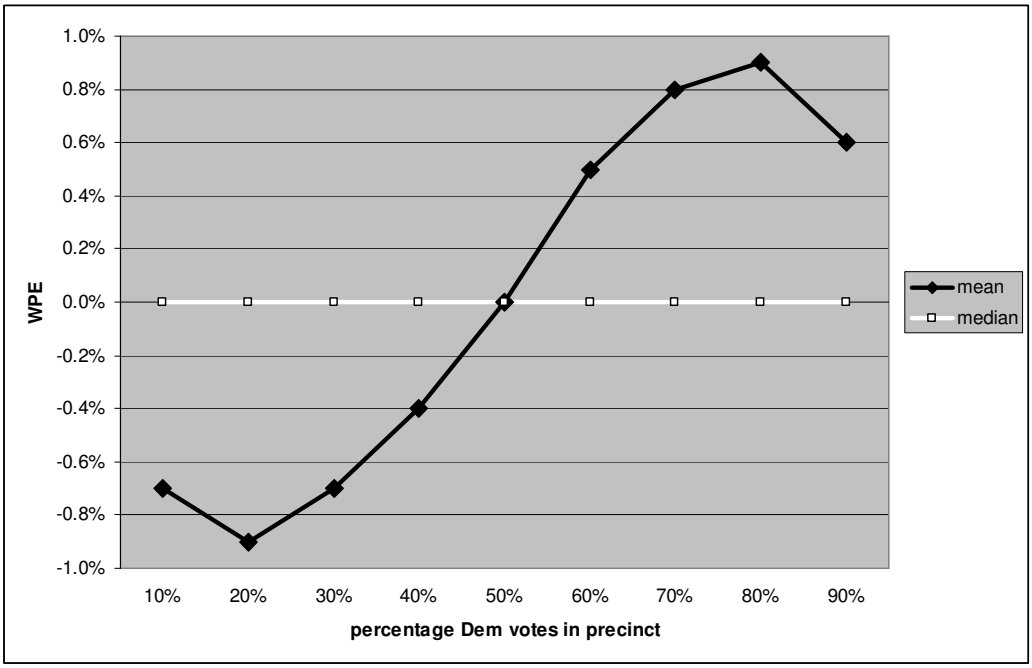
In a second set of "net bias" conditions, I adjusted the mean response rate of one type of voter. For example, I raised the mean response rate of Kerry voters to 56%,

keeping the standard deviation at 10%[13], while keeping the mean response rate for Bush voters at 50%, also with a standard deviation of 10%. The mean differential was thus 6 percentage points, and the standard deviation of the differential was 14. I will refer to this as a "net Kerry bias" condition. I also ran similar simulations, reducing the mean response rate for "Bush voters" to, say, 44%, and keeping the response rate for "Kerry voters" at 50%, again giving a "net Kerry bias." I then repeated these conditions in reverse, giving me a set of "Bush net bias" conditions. I performed 1,000,000 iterations for each "precinct" category for each condition, thus representing 1,000,000 notional "precincts" in each category each time.

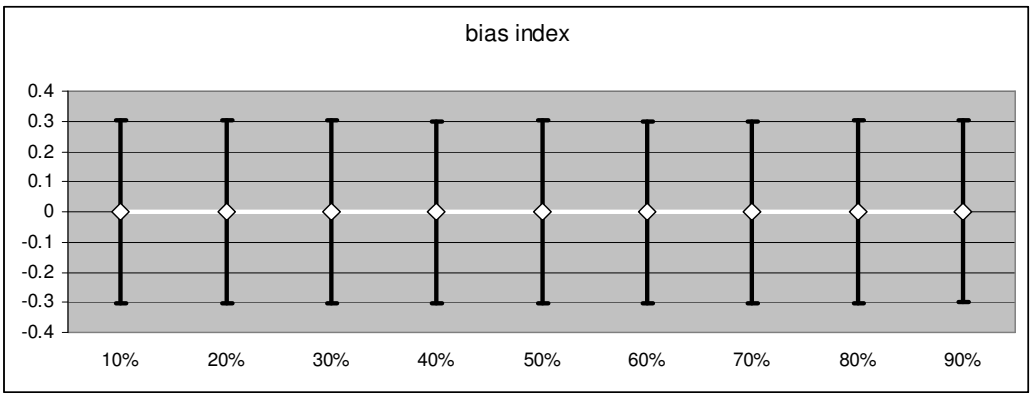
The model output for both the "no net bias" condition is given in Figure 3. In Figure 3a, the X axis shows the percentage of "Kerry" voters in that category of "precinct." On the Y axis, the black line represents the mean WPE for that category of precinct, and the white line represents the median for that category of precinct. In Figure 3b, the "mean bias" index is shown, with error bars representing standard deviation.

It can be seen from Figure 3a, as in Figure 2, that even where the net bias is zero (and therefore the overall mean WPE is zero) the mean WPE values are negative (representing an over-estimate of the Kerry vote) where Kerry's vote is a small proportion of the total, i.e. in highly Republican precincts, but positive (representing an over-estimate of the "Bush" vote) where Kerry's share of the vote is a large proportion of the total, i.e. in highly Democratic precincts. As in Figure 2, the plot of the mean WPEs has a characteristic "S" shape. However, again as in Figure 2, the median WPE values are approximately zero for all precinct categories, and for precincts in which the vote is evenly divided, both the median and the mean WPE are zero. Thus for all precinct categories except the evenly divided precincts, the WPE distribution is skewed, negatively in more Republican precincts, and positively in more Democratic precincts. Figure 3b indicates that mean of the bias index is equal to zero for all categories of precinct, and that variance is also constant across categories. In contrast, the mean WPE variance (not shown) is maximal for even precincts.

[13] As distributions of percentage values will only be symmetrical around a mean of 50%, this would have given the distribution of "Kerry" response rates a slightly asymmetrical distribution. However, the departure from symmetry is very slight for the ranges used. When the model was repeated with percentages of 53% and 47% for the two voter groups, thus giving the two voter groups equal and opposite degrees of asymmetry, similar results were obtained.



a



b

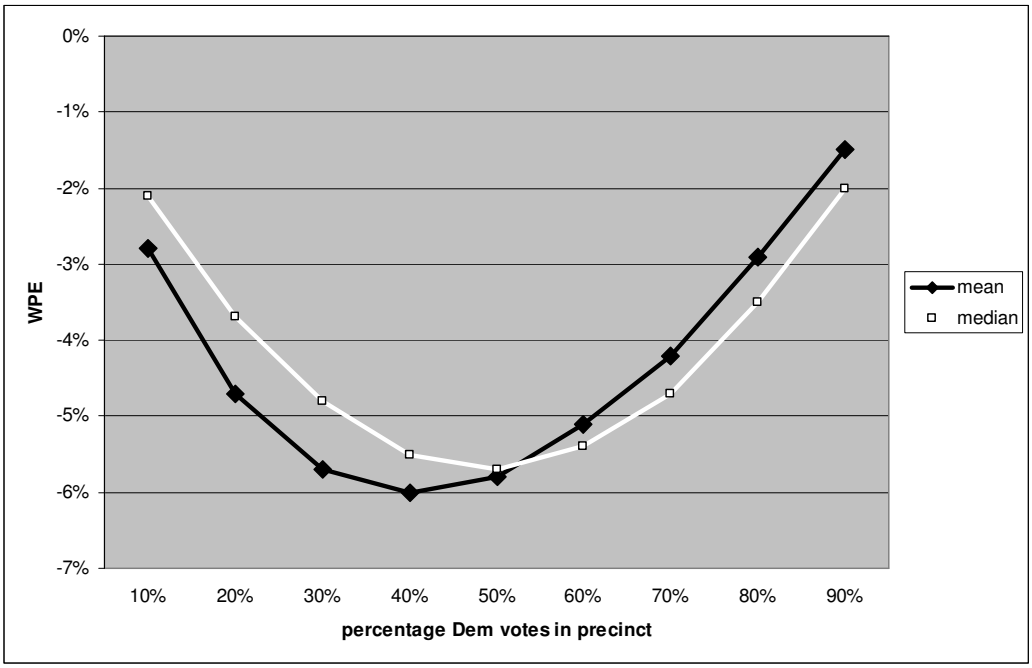
Figure 3

Model output for "no net bias" conditions. Kerry's share of the vote is plotted on the horizontal axis, with "high Republican" precincts on the left, and "high Democrat" precincts on the right.) In Figure 3a, the vertical axis represents the magnitude of the bias as determined by mean WPE (black line) and median WPE (white line.) A negative WPE represents an over-estimate of Kerry's share of the vote, while a positive WPE represents an over-estimate of Bush's share of the vote. In Figure 3b, the vertical axis represents the value of the bias index for each precinct category. A positive value indicates an over-estimate of Kerry's share of the vote. In this scenario the mean bias is zero. The error bars represent the standard deviation. Note that mean WPE varies with precinct category but the median does not. Neither mean nor variance in the bias index varies with precinct category.

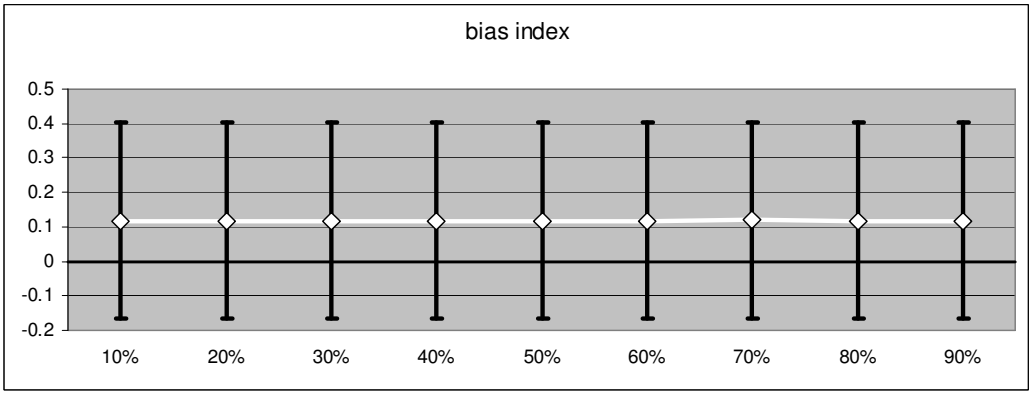
The output for a "net Kerry bias condition" is given in Figure 4. Mean response rates for "Kerry" voter were set at 56%, and mean response rates for "Bush" voters at 50%. In Figure 4a, mean and median WPs are plotted against precinct partisanship. The shape of the plots indicates that both the mean and the median WPE values for each precinct vary with precinct partisanship, and distributions are only normal where partisanship is even. WPEs are greatest (most negative) for even precincts, but mean WPEs are greater for high Republican precincts than for high Democrat precincts. The distribution of WPEs is negatively skewed for more Republican precincts, and slightly positively skewed for more Democratic precincts.

Figure 4b represents the mean value of the bias index for each precinct. It has a constant value greater than zero (=0.118) indicating a net Kerry over-sampling, and again its variance does not change with precinct partisanship. If we take the anti-log of the index we retrieve a value of 1.12 alpha, which is equal to 56% divided by 50% (see Equation 2.)

Reversing the "bias" (i.e. modelling the same degree of bias, but one that tended to overestimate the "Bush's" share of the vote) simply resulted, predictably, in a reversing of the plot, and a mean bias index that was equal and opposite (=-0.118.)



a



b

Figure 4

Model output for "Net Kerry bias" condition. The axes and legends are as for Figure 2. Model assumes a 56% response rate for Kerry voters and a 50% response rate for Bush voters. In Figure 4a, both mean and median WPEs vary with precinct partisanship, and distributions are only normal where partisanship is even. WPEs are greatest (most negative) for even precincts, but means WPEs are greater for high Republican precincts than for high Democrat precincts. Figure 4b shows that the mean and variance of the bias index again remains constant across the range of precincts, although its mean value is now positive, indicating an over-estimate of the Kerry vote.

This is a simplified model for a number of reasons. Firstly, I deliberately excluded sampling error from the model, as I wanted to demonstrate the effect of randomised differential non-response alone. Secondly, I regarded mean response rates as an output variable for a given pair of differential response rates. Clearly, for a given pair of differential response rates, mean response rates will be a function of precinct partisanship. For example, if "Bush" voters had lower response rates, the model predicts, as also predicted by the USCV authors[14], that mean response rates should be lower in "high Republican" precincts than in "high Dem precincts." This does not appear to be the case for the data in the E-M report, which reports that there was no significant difference in response rates across precinct categories. Thirdly, as this model was designed to investigate principles rather than to simulate actual data, I used arbitrary values for the variance in response rates.

I therefore constructed a second model in which I introduced sampling error, thus allowing me to estimate whether the variance in WPE produced by the model, including that due to differential non-response and as well as to that due to sampling error, was comparable with the variance in the E-M data. I assumed that the number of "voters approached" was 150, and that response rates had a mean of 56% for Democrats and 50% for Republicans, giving mean sample sizes of 79. The inputs thus matched those given in the E-M report.[15] I modelled 5 precinct categories, with partisanship values taken from the USCV authors' assumptions based on the normal curve.[16] I specified the number of precincts in each category to equal the numbers in the E-M sample.[17] I then iterated the model 1,000 times, recording the mean and variance of the overall response rates for each precinct category for each iteration, and took the mean values of the mean absolute WPEs over the 1,000 iterations.

The mean absolute WPEs values produced by the model are given in Table 2.

[14] Mitteldorf et al. 2005, p. 25, p.9.
[15] Edison Media Research and Mitofsky International 2005, p. 31.
[16] Mitteldorf et al. 2005, p. 25, Table 3.
[17] Edison Media Research and Mitofsky International 2005, p. 36.

Category (assumed percentage Kerry)	Absolute WPE (Model)	Absolute WPE (E-M data)
High Dem (87%	5.1	8.8
Mod Dem (67%)	10.3	13.4
Even (50%)	12.0	15.2
Mod Rep (33%)	11.1	13.2
High Rep (13%	6.1	12.4

Table 2

Absolute WPE values for each precinct category produced by the model as compared with those given in the E-M report. Values for assumed percentages for Kerry are those interpolated from the normal curve in the USCV report, Table 3.[18]

In the first column are the absolute values produced by the model, and in the second column are the absolute values given in the E-M report. Although the model reproduces the pattern found in the E-M data of lower absolute values for more partisan precincts, in all cases, the absolute WPE values in the model are lower than those in the E-M data, indicating that the variance in WPE in the model are under- rather than over-estimates of those in the actual data, and that my variance assumptions have probably been conservative.

The mean response rates, as predicted, were lower for the "high Rep" precincts than for the "high Dem" precincts – as this is given by the equations we know it is a "real" effect. However, effect sizes computed for ANOVAs carried out on the response rates for each precinct category were small, and indicated that if the differential response variance in the actual data is greater than in the model, which the absolute WPE values suggest it is, that the response rate differences between precinct categories may be too small to be reliably detected with the statistical power available in the E-M data, even using the one-tailed test required by the hypothesis. Indeed the E-M evaluation states that the differences were not statistically significant[19], although this statement may refer to the results of a two-tailed test.

Conclusions

So what are the implications of these computations?

[18] Mitteldorf et al. 2005, p. 25, Table 3.
[19] Edison Media Research and Mitofsky International 2005, p. 37.

Firstly, they indicate that the WPE is a confounded dependent measure, at best introducing noise into any analysis, but at worst creating artefacts that suggest that bias is concentrated in precincts with particular degrees of partisanship where no such concentration may exist. It is also possible that other more subtle confounds occur where a predictor variable of interest such as precinct size, may be correlated with partisanship. In order to establish the proportion of variance in differential non-response that can be accounted for by known independent variables, the use of an unconfounded index of bias as the dependent variable, such as the one described, would seem to be essential. In the case of the 2004 exit polls, it would therefore be of interest to know what proportion of total variance in genuine within-precinct "bias" could be accounted for by the factors postulated in the E-M report, and whether, after thus accounting for known methodological factors, any precincts/states proved to be statistical outliers that might indicate the possible contribution of vote-count corruption to the exit-poll error. It would also be of interest, where the mean response rates for each precinct are known, to compute not only the bias, but the actual differential response rates required to produce that bias, using Equations 9 and 10. An algebraically equivalent formula is given in the USCV report[20]. Examination of these might answer questions as to whether differential non-response rates arose from lower than average response rates from one group of voters, or higher than average response rates from the other, as well as indicating where factors may have contributed to higher or lower response rates from both groups.

Secondly, it may be worth noting that the results of these computations indicate that even where the net signed WPE is zero, it cannot be inferred that no bias occurred, as randomly distributed precinct-specific biases may nonetheless favour one candidate as frequently and as greatly as the other. However, such randomly distributed precinct specific biases will tend to result in greater over-estimates of the vote for a candidate in precincts where votes for that candidate are in the minority, and greater under-estimates of the vote for a candidate in precincts where votes for that candidate are in the majority, even though biases in any one direction are not concentrated in particular categories of precinct partisanship. The presence of such biases, however, may be able to be inferred, even where the mean WPE is zero, by patterns of

[20] Mitteldorf et al. 2005, p. 26

distribution of mean and median WPEs by precinct partisanship that resemble those modelled in Figure 3a. Sampling error alone will give no such pattern.

Finally, a third implication of the computations reported in this paper is that the pattern of mean and median WPEs by precinct partisanship reported by E-M, may not be, as USCV claim, an indication of bias concentrated in Bush strongholds, but rather the pattern that might be expected if differential response rates were large and widespread, with the mean bias reflecting relative over-polling of Kerry voters (or under-polling of Bush voters.)

In their updated "Appendix B" dated April 11th, the USCV authors report the results of their analyses when they compute the "required differential response rates" for the mean WPEs and overall response rates given in the E-M report, after adjusting their algebraic formulae to avoid the confound of WPE by partisanship described in this paper. This was done in response to my drawing of their attention to the confound (which they kindly acknowledge.) To ascertain whether applying such a strategy to computations involving the mean WPEs would abolish any artefactual differences between bias rates in precincts of differing partisanship, I re-ran the computational model described above, and computed the "bias index" (using my own transform) that that would result from being applied to the mean WPEs for each category of precinct, rather than to the individual WPEs. This procedure resulted in a series of values for the bias index that still strongly overestimated the bias in partisan precincts. When applied to the medians this artefact was abolished; unfortunately it was highly sensitive to the exact partisanship value. When it was applied to broad bands of precinct category, such as the "greater than 80%" category, rather than to precise partisanship values, the artefact crept back in. For small samples of precincts (there were only 40 precincts in this band in the E-M report), it tended to give values close to the values generated when the formula was applied to the means. These findings underline the importance of deriving an unconfounded measure of bias at the level of individual precinct data.

One way, therefore, of resolving the question as to whether "reluctant Bush responders" were indeed more prevalent in one category of precinct rather than another would be to compute a pure "bias index" for each precinct, rather than applying the formula to either the means or medians given, and then to regress the

"bias index" values on categories or levels of precinct partisanship. However, in lieu of this, it may be informative to observe that the pattern of means and medians in Table 1 is at least consistent with the pattern that would result if differential non-response was not concentrated in precincts of one category, but randomly distributed amongst all precinct categories. To illustrate this, Figure 4 shows the means and medians from the five precinct categories in Table 1.

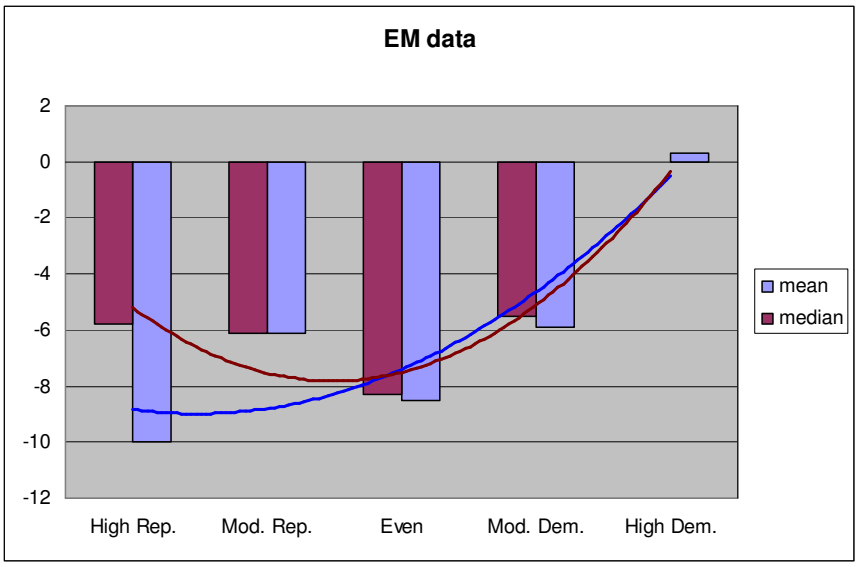


Figure 5

E-M median and mean WPE values for five categories of precinct from Table 1. Trend-lines are best quadratic fit.

Similarities between this plot and the plot in Figure 4 may be noted, particularly the way the median diverges from the mean in the high Republican category. To the extent that the pattern in Figure 4 shares a family likeness with the pattern of the modelled data in Figure 3, the conclusion drawn in the USCV report, that the pattern observed requires "implausible" patterns of non-response and thus leaves the "Bush strongholds have more vote-count corruption" hypothesis as "more consistent with the data", would seem to be unjustified. The pattern instead is consistent with the E-M hypothesis of "reluctant Bush responders", provided we postulate a large degree of variance in the degree and direction of bias across precinct types. Mathematically, the observed pattern could arise from widespread fraud as well as from widespread response bias; differential vote spoilage rates for Kerry votes, or "ballot stuffing" of

Bush votes, would produce results indistinguishable from "reluctant Bush responders." However, this is not the inference currently drawn from the data by USCV in their report.

Elizabeth Liddle, April 27th, 2005

Acknowledgments are due to Mark Blumenthal (http://www.mysterypollster.com), Rick Brady (http://www.stonescryout.org//bios.html#rick), "DemFromCT" at Daily Kos, (http://www.dailykos.com/), and Peter Liddle, without whom this analysis would never have achieved fruition, and to Keith Barratt of for helpful suggestions and for hosting the paper at http://neweuropeantimes.mywowbb.com. Thanks are also due to Campbell Read, from the USCV team, who first drew my attention to the problems in using parametric statistics with proportional data, as well as to Bruce O'Dell, David Dodge and others from the USCV team who have been generous with their time and their ideas in responding to the challenges presented by this paper.

References

Edison Media Research and Mitofsky International. 2005. *Evaluation of Edison/Mitofsky Election System 2004* [PDF on website]. 19th January 2005 [accessed 16th April 2005]. Available from http://www.exit-poll.net/election-night/EvaluationJan192005.pdf.

Merkle, DM, and Edelman, M. 2002. Nonresponse in Exit Polls: A Comprehensive Analysis. In *Survey Nonresponse*, edited by R. Little: John Wiley and Sons, Inc.

Merkle, DM. and Edelman, M. 2000. A Review of the 1996 Voter New Service Exit Polls from a Total Survey Error Perspective. In *Election Polls, the News Media and Democracy*, edited by P.J. Lavrakas, M.W. Traugott. Chatam House

Mitteldorf, J, K Dopp, SF Freeman, B Joiner, F Stenger, RG Sheehan, PF Velleman, V Lovegren, CB Read, J Simon, Ron Baiman, and B O'Dell. 2005. *Analysis of the 2004 Presidential Election Exit Poll Discrepancies* [PDF file on website]. US Counts Votes, 12th April 2005 [accessed 16th April 2005]. Available from http://electionarchive.org/ucvAnalysis/US/Exit_Polls_2004_Edison-Mitofsky.pdf.