Mathematical Proof that Election Sciences Institute's Test to Rule Out Vote Fraud Is Logically Incorrect

Even If Logically Corrected, ESI's Test Would Require More Data and Have Many Pitfalls

Contents

Author: Kathy Dopp
Contributors: Ron Baiman, David Dodge, Jonathan Simon, Victoria Lovegren
Content Reviewers: Ron Baiman, Robert Klauber, David Dodge, David Griscom, Joseph Baugher
Style/Grammar Reviewer: Jill Hacker

Abstract

John F. Kerry won the 2004 presidential election according to exit polls. Yet George W. Bush won the election according to the official election results. Due to numerous questions about the integrity of U.S. vote counting systems which are rarely, if ever, independently audited for accuracy, the validity of the official presidential election results was thrown into doubt. Because Kerry quickly conceded, these questions were left to be resolved by independent and often unpaid analysts and patriots concerned about American democracy. Exit pollster Warren Mitofsky has been particularly active in defending the accuracy of the vote counts. In a January 19, 2005 paper, Mitofsky stated, without any supporting evidence, that the exit poll discrepancy was caused by Kerry voters responding to exit polls more than Bush voters. From January through June 2005, the National Election Data Archive (NEDA) derived new algebraic methods for studying exit poll discrepancy patterns in order to evaluate this hypothesis. NEDA found that this "response bias" hypothesis was contradicted by the available exit poll data.[1] In June, Mitofsky, working with the Election Science Institute (ESI), publicly revealed another hypothesis which they claimed could rule out vote fraud as a cause of the 2004 presidential exit poll discrepancies. On October 14, Mitosky presented an analysis, based on this hypothesis, at a debate in Philadelphia where he proclaimed, "This kills the vote fraud argument".

The ESI hypothesis, proven herein to be logically incorrect, is stated in an ESI paper[2] as follows.

> "If systematic fraud or error in vote counting [favoring Bush] occurred [in precincts] in 2004 but not in 2000, [then] Bush would have done significantly better in those precincts in 2004 [than in

[1] See "History of the Debate Surrounding the 2004 Presidential Election" for references. http://electionarchive.org/ucvAnalysis/US/Presidential-Election-2004.pdf
[2] "**Analysis of the 2004 Ohio Exit Polls and Election Results**" http://electionscience.org/reports/view_reports

> 2000], and we would see larger differences between the reported vote and exit poll in those precincts [than in other 2004 exit-polled precincts]."

This paper presents a proof, using mathematical logic, that Mitofsky's newest hypothesis is logically invalid and therefore any analysis based on it is likewise invalid. Vote fraud *cannot* be ruled out by using this hypothesis presented first in ESI's June 6, 2005 paper: "Ohio Exit Polls: Explaining the Discrepancy"[3]

The text in brackets, "[]," is added above in order to clarify ESI's meaning. ESI's interpretation can be determined by examination of both ESI's analysis in its June paper and the analysis presented by Mitofsky at the October 14 Freeman-Mitofsky debate.

This paper shows that ESI's hypothesis cannot be used as a test to rule out vote fraud unless it is reformulated in a logically correct way and Mitofsky/ESI provides unadjusted exit poll data for 2000 for the same precincts to support it. However, even if ESI had provided this data, the corrected hypothesis could not be used without compensating for other factors, listed below, that were not mentioned or accounted for by ESI in its analysis.

NEDA will be releasing a new study based on the Ohio precinct-level exit-poll data provided in the ESI paper that shows virtually irrefutable evidence for vote miscount in Ohio's 2004 presidential election.

Nonmathematical persons may want to skip now to the section in this paper entitled "The English-Language Explanation: Why is ESI's Analysis Illogical?"

Overview of Math Logic Proof of the Invalidity of ESI's Hypothesis

ESI's hypothesis, and how they use it to "rule out vote fraud"

ESI's hypothesis is an inference of the form $A \rightarrow (B \wedge C)$. That is, if statement A is true, then statement B and statement C must follow as true. If ESI's inference, upon which it bases its analysis on, were logically valid, then its contrapositive $(\neg B \vee \neg C) \rightarrow \neg A$ would logically prove the absence of vote fraud. That is, if ESI's hypothesis were valid, then if statement C is not true or statement B is not true, then it must follow that statement A is not true, and so it would be logically correct for ESI to rule out vote fraud with their analysis.

e.g. An example of using a statement's contrapositive to make a conclusion is "if it rains I will stay indoors; I was outdoors so it didn't rain." An example more like ESI's inference is: "if it rains, I will stay indoors AND do the wash; I didn't do the wash, so it didn't rain;"

ESI's hypothesis, as stated and analyzed by ESI, is not a logically valid hypothesis.

ESI's hypothesis is an inference of the form $A \rightarrow (B \wedge C)$.[4] That is, if statement A is true, then statement B and statement C must follow as true. However, as we will show, $A \rightarrow (B \wedge C)$ is clearly not a valid

[3] Authors are Susan Kyle, Douglass A. Samuelson, Fritz Scheuren, and Nicole Vicinanza with Scott Dingman and Warren Mitofsky. ESI originally presented this logically incorrect hypothesis in June 2005 to wrongly "rule out vote fraud" in the Ohio exit poll data. Mitofsky used this same incorrect logic to "rule out vote fraud" in the national exit poll data at an October 14, 2005 debate at the University of Pennsylvania at Philadelphia with Steven Freeman. Mark Lindeman and Elizabeth Liddle both posted the same analysis, based on the same invalid hypothesis, on the Internet after the debate, which Mitofsky claimed ruled out vote fraud for the entire national exit poll dataset.

inference under any interpretation, because there are many hypothetical counterexamples when A exists and not B ($A \wedge \neg B$). Therefore $A \rightarrow (B \wedge C)$ is not a logically correct inference and one cannot reach ESI's final conclusion $\neg A$, that vote fraud is ruled out.

e.g. The original statement 'if it rains, I will stay indoors AND do the wash' must be false if "I didn't do the wash but it rained;" and therefore we can't use either the wash or my not being inside to tell us anything about the weather (i.e. votefraud).

Even if corrected, ESI's hypothesis *cannot* be used without additional data.

If one "corrects" ESI's hypothesis to a valid inference $(A \wedge B) \rightarrow C$, that is, if A and B are both assumed to be true, then C must be true, then this hypothesis can be used to rule out vote fraud only if the unadjusted exit poll data for the 2000 election were provided for the same precincts and if the data, in fact, support this hypothesis.

However, the hypothesis $(A \wedge \neg B) \rightarrow C$ is also valid. That is, if A and not B are assumed true, then C must be true. Hence ESI's hypothesis would be more elegantly stated as $A \rightarrow C$, that is if A occurs, then C must occur. However, the inference $A \rightarrow C$ can be used to rule out vote fraud only if the unadjusted exit poll discrepancies for 2000 are also provided, which they were not.

Even if the hypothesis is corrected logically and the data provided, the hypothesis requires consideration of other factors.

Even if Mitofsky provided the unadjusted exit poll discrepancy data for the 2000 election, and one applied a logically correct version of the hypothesis, *any hypothesis comparing the 2000 and 2004 elections must take other factors into consideration if it is to be useful for investigating vote fraud*. We provide a list of other factors that must be considered to use such a hypothesis properly.

---

[4] We know that this is the correct interpretation of ESI's hypothesis because they use it to claim $\neg A$ (not vote fraud) by showing $\neg(B \wedge C)$ because there are not larger discrepancies in precincts where Bush did better in 2004 than in 2000.

Converting ESI's Hypothesis to a Math Logic Statement for Examination

The ESI/Mitofsky/Liddle conclusion that the nationwide exit poll data shows no evidence of vote fraud in the 2004 presidential election[5], is based on the inference that:

> "If systematic fraud or error in vote counting [favoring Bush] occurred [in precincts] in 2004 but not in 2000, [then] Bush would have done significantly better in those precincts in 2004, and we would see larger differences between the reported vote and exit poll in those precincts [than in other 2004 exit-polled precincts]."[6]
>
> We examine the logic of ESI's hypothesis by labeling its clauses:
>
> A = systematic fraud or error in vote counting occurred in 2004 but not in 2000
>
> B = Bush would have done significantly better in those precincts in 2004
>
> C = We would see larger differences between the reported vote count and exit poll in those precincts.

Stated in terms of the inference $A \rightarrow (B \wedge C)$, ESI's hypothesis is:

> Hypothesis 1:
>
> IF systematic fraud or error in vote counting [favoring Bush] occurred [in precincts] in 2004 but not in 2000, THEN Bush would have done significantly better in *those* precincts in 2004 than in 2000), AND we would see larger differences between the reported vote count and exit poll results in *those* precincts.

Or, because English can be ambiguous, ESI could possibly have meant that $(A \wedge B) \rightarrow C$ so ESI's hypothesis could have been:

> Hypothesis 2:
>
> IF (systematic fraud or error in vote counting [favoring Bush] occurred [in precincts] in 2004 but not in 2000, AND Bush had done significantly better in *those* precincts in 2004 than in 2000), THEN we would see larger differences between the reported vote count and exit poll results in *those* precincts in 2004 than in 2000.

By reviewing ESI's report, however, it is clear that ESI intended hypothesis #1 above because ESI's analysis included all of the Ohio exit-polled precincts. Their analysis was *not* limited to just the precincts where Bush did significantly better in 2004 than in 2000; and it did *not* include exit poll discrepancy data for 2000.

We now show that hypothesis #1, above, upon which ESI based its Ohio analysis and Mitofsky based his national analysis, is an invalid inference and that only hypothesis #2 is a valid inference.

---

[5] I was informed by October 24, 2005 Internet postings by Elizabeth Liddle, who has worked frequently for Mitofsky since April, 2005, that " the failure to find a linear relationship between the magnitude of the exit poll discrepancy and the magnitude of Bush's increase in vote share since 2000 is a major problem for the argument that the discrepancy indicates fraud." And "I was not talking about [only] Ohio. I was talking about an equivalent analysis done on the whole [U.S.] dataset, with the same finding, and presented by Mitofsky at the Freeman-Mitofsky debate. Mark Lindeman, with Mitofsky's permission, posted an account here which includes Liddle's graphs as presented by Mitofsky in October: http://inside.bard.edu/~lindeman/slides.html "

[6] Page 3 in its June 6, 2005 paper: "*Ohio Exit Polls: Explaining the Discrepancy*" by Susan Kyle, Douglass A. Samuelson, Fritz Scheuren, and Nicole Vicinanza with Scott Dingman and Warren Mitofsky. See http://electionscience.org/Members/stevenhertzberg/report.2005-07-19.7420722886/report_contents_file/

ESI's Hypothesis, as Stated and Analyzed by ESI, is Logically Invalid

Consider hypothesis #1: $A \rightarrow (B \wedge C)$

> IF (systematic fraud or error in vote counting [favoring Bush] occurred [in precincts] in 2004 but not in 2000), THEN Bush would have done significantly better in *those* precincts in 2004), AND we would see larger differences between the reported vote count and exit poll results in *those* precincts.

Note that $A \rightarrow B$ must be true for $A \rightarrow (B \wedge C)$ to be true. So we first consider the inference $A \rightarrow B$.

> Hypothesis #1, clause #1:
> "IF systematic fraud or error in vote counting [favoring Bush] occurred [in precincts] in 2004 but not in 2000, THEN Bush would have done significantly better in *those* precincts in 2004 [than in 2000]"

To prove that $A \rightarrow B$ is not a valid inference, we show that A can exist without B existing. That is, we show $A \wedge \neg B$.

In other words, if we give a counterexample that shows how Bush can win with pro-Bush vote count error in 2004, and win in 2000 without vote count error, and *not* do significantly better in those precincts in 2004 than in 2000, then we have shown that $A \rightarrow B$ is not a valid inference. Thus if we find counterexamples, then ESI's hypothesis is logically incorrect and invalid, and therefore so is ESI's analysis and conclusion.

However, because the English language is ambiguous, we consider two possible meanings of "do significantly better in those precincts in 2004". There are two basic interpretations of "do significantly better in those precincts":

1. receive a higher share of the overall votes in 2004 than in 2000, or
2. do better in more precincts in 2004 than he did in 2000 (This second interpretation also covers the case that Bush does significantly better in more precincts; better in significantly more precincts; and significantly better in significantly more precincts - - because if Bush does not have to do better in more precincts at all to win in 2004 with vote fraud, then obviously he does not have to do significantly better in any sense.)

ESI used interpretation #2 above. We know this because ESI/Mitofsky never examine the number of voters in precincts (or precinct size and weight[7]) in their analysis, but rather make their conclusions based on the pattern and number of precincts; and do not consider the 2000 exit poll discrepancy data.

However, as it turns out, the inference (hypothesis #1, clause #1 above) is not logically valid in either of the possible interpretations. We consider both interpretations of ESI's hypothesis #1 to find out whether ESI's hypothesis is logically correct or not.

Counterexample #1:

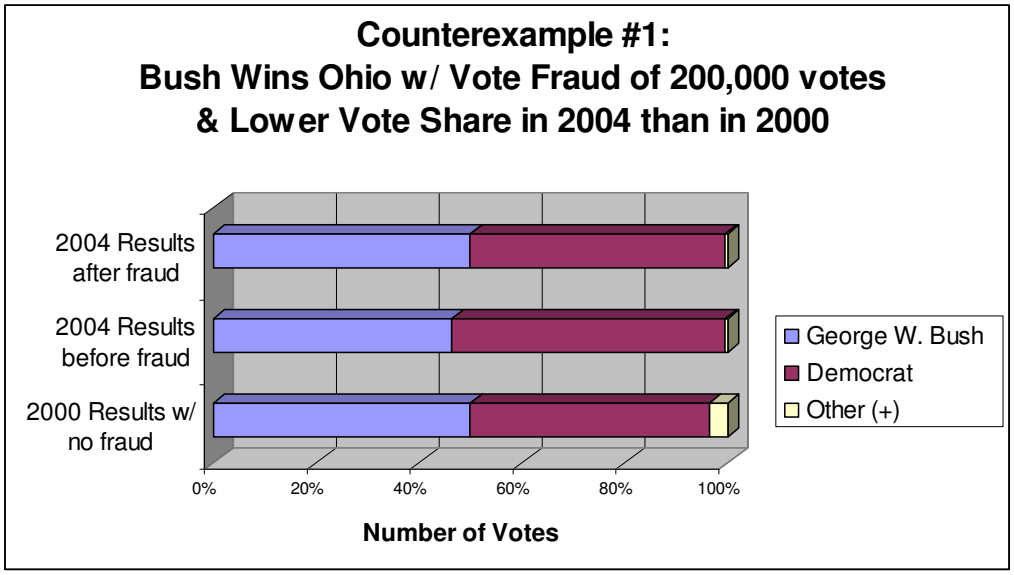
Bush can win in 2004 due to vote miscounts AND have a lower share of overall votes in 2004 than he did in 2000.

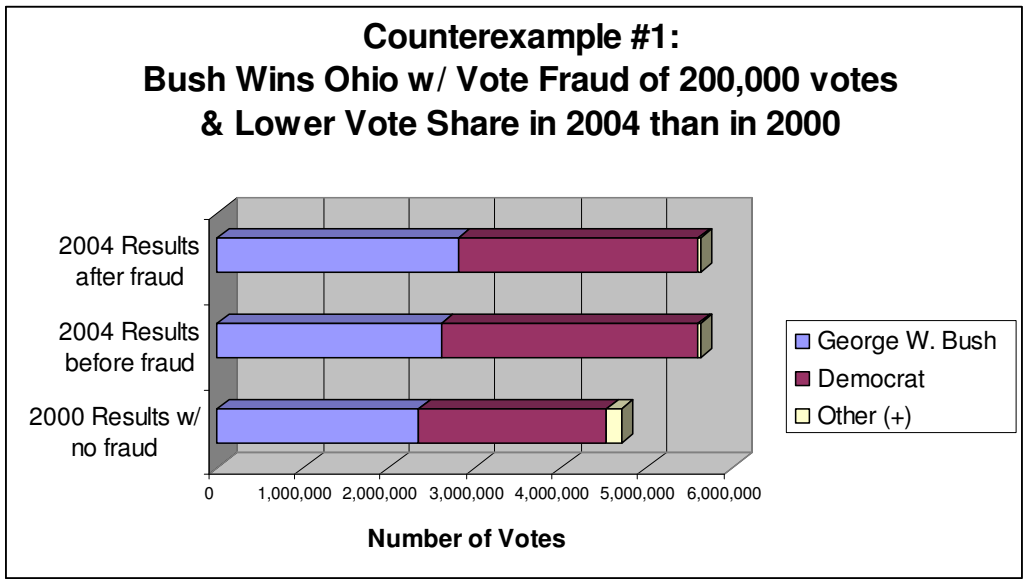
[7] In order to estimate the state vote from exit polled precincts, weights must be assigned to each precinct that reflect the share of overall state votes represented by that precinct.

Hypothetical Counterexample #1, by Percent			
	2000 % with No fraud	**2004 % before fraud**	**2004 % after fraud**
Bush	49.97%	46.37%	49.93%
Gore/Kerry	46.46%	53.15%	49.60%
Other (+)	3.57%	0.48%	0.48%

By number of voters, rather than share, the same example looks like this:

Hypothetical Counterexample #1, by Number of Votes			
	2000 votes with no fraud	**2004 votes before fraud**	**2004 votes after fraud**
Bush	2,351,209	2,609,768	2,809,768
Gore/Kerry	2,186,190	2,991,167	2,791,167
Other (+)	168,058	26,973	26,973





It should be mentioned for clarity that the data presented here is not actual election data but is a hypothetical example to show that such a scenario in which vote fraud favoring Bush takes place in 2004 but not in 2000, but Bush does not do significantly better in 2004 than he did in 2000 is, in fact, possible.

If Bush wins in 2004, he obviously does not have to beat his own 2000 vote share; he merely has to score a higher share of the total vote in comparison to the opposing candidate in 2004 to win, whether or not there are vote miscounts and whether or not he won or lost the last election. Since Bush won the last election, his overall vote share would not have had to increase at all necessarily, as long as his opponent's share did not increase by more than the margin in 2000.

Outcome-changing vote fraud does not have to increase the vote share for a candidate from a prior election because, for example, voters may have tried to vote more for Kerry in 2004, yet not all their votes were counted accurately. Whether there is vote fraud or not does not affect the false nature of this inference because vote fraud can be added/subtracted to any relative vote share of the candidates.

Counterexample #2:

Bush can win in 2004 with vote miscounts and not do better in more precincts than he did in 2000.

According to the official 2004 vote count, Bush won the election, but Bush, in fact, did *not* do better in more exit polled precincts in 2004 than he did in 2000 (Bush outperformed his 2000 vote share in only 31% (15) of the 49 of Ohio exit polled precincts[8]), so obviously it is not necessary to do "significantly better in more precincts" to win an election. Clearly it is also *unnecessary* to do significantly better in "more precincts" to win an election due to vote fraud.

Vote fraud does not have to increase the vote share for a candidate from a prior election because voters may have tried to vote more for the opposing candidate in the 2004 than in 2000, or a strong third party candidate may have been present in 2000 and be absent in 2004 and his vote share may have in reality transferred more to the opposing candidate, yet not all votes were counted accurately. In addition increase in voter turnout can affect the arithmetic. In other words, vote fraud can be added/subtracted to any relative vote share of the candidates. A candidate could do better in fewer precincts than in the last election by committing vote fraud in a small number of more populated urban precincts containing more votes.

The following counterexample uses percentages that are very close to actual percentages in the Ohio 2000 and 2004 presidential elections. Each row represents one precinct. Yellow cells show results in the 2000 election; green the 2004 election. This is only one out of many possible counterexamples that prove that the Mitofsky/ESI's hypothesis[9] is flat wrong. In the hypothetical example in the table below, Bush does better in 2004 than in 2000 in only 35% of precincts (similar to the actual 31% of exit polled Ohio precincts). There is outcome-changing vote fraud favoring Bush. Bush receives a higher vote share and more votes in 2004, but Kerry wins in 2004 according to voters' choice. And there is *not*, contrary to Mitosfky/ESI's claim, a positive correlation between Kerry exit poll overestimates[10] and precincts where Bush's vote share is better. In fact, just the opposite is true, and there is a small *negative* trend between Kerry exit poll overestimates and Bush vote share increase from 2000 to 2004.

[8] These are the number of precincts with above zero values on the vertical axis in Figure 3 of the ESI report, relative to the total number of exit polled precincts in Figure 3.

[9] "In those precincts where fraud took place, you would tend to see an increase in Bush's share from 2000 to 2004 that is correlated with increase in Kerry exit poll overestimates."

[10] The term “Kerry exit poll overestimate” here means an instance of Kerry’s exit poll numbers exceeding his official vote count for a precinct. The term is value neutral, not meant to imply that either the exit poll data or the official vote count is wrong.

Hypothetical Counterexample #2 to show that $\mathbf{A} \rightarrow (\mathbf{B} \wedge \mathbf{C})$ is faulty logic.

		2000 Number of Votes - No Fraud		2000 % Votes			2004 Number of Votes - No Fraud		Voters Choice 2004 %		WPD Vote Shift to Bush	Diff 2000 to 2004	2004 Number of Votes After Fraud		After Fraud %		Indicator: "2004 vote margin exceeded 2000 voter margin"
	precinct size 2000	Bush #Votes 2000	Gore #Votes 2000	Bush % 2000	Gore % 2000	precinct size 2004	Bush #Votes	Kerry #Votes	Bush % 2004	Kerry % 2004			Bush #Votes 2004	Kerry #Votes 2004	Bush 2004	Kerry 2004	
1	229	69	156	30	68	275	129	129	10	90	-15	-5	69	206	25	75	0
2	117	40	75	34	64	140	31	47	80	19	-4	50	118	22	84	16	1
3	208	79	125	38	60	250	75	158	21	79	-4	-13	62	187	25	75	0
4	216	85	127	39	59	259	28	247	50	50	-3	14	137	121	53	47	1
5	259	104	150	40	58	311	87	176	69	31	-2	31	221	89	71	29	1
6	138	57	79	41	57	166	228	79	28	72	-7	-6	58	108	35	65	0
7	230	99	127	43	55	276	215	95	18	82	-10	-15	77	198	28	72	0
8	220	97	119	44	54	264	112	27	33	67	-1	-10	90	174	34	66	0
9	65	29	34	45	53	78	149	137	39	60	-3	-3	33	45	42	58	0
10	174	82	89	47	51	209	95	131	30	70	-12	-5	88	121	42	58	0
11	195	96	96	49	49	234	149	138	32	68	-5	-12	87	147	37	63	0
12	240	125	110	52	46	288	65	57	52	48	-2	2	155	132	54	46	1
13	188	100	85	53	45	226	261	159	42	58	-1	-10	97	128	43	57	0
14	165	93	69	56	42	198	92	104	44	56	-4	-8	95	103	48	52	0
15	102	58	42	57	41	122	159	141	53	47	-6	2	72	50	59	41	1
16	257	160	93	62	36	308	49	226	74	26	2	10	222	86	72	28	1
17	351	218	126	62	36	421	63	146	62	38	-8	8	295	125	70	30	1
18	163	101	59	62	36	196	52	197	47	53	-15	0	121	74	62	38	0
19	251	166	80	66	32	301	46	119	53	47	0	-13	159	141	53	47	0
20	238	174	60	73	25	286	87	110	52	48	-11	-10	180	105	63	37	0
		2035	**1899**				**2172**	**2623**					**2436**	**2361**			"7 out of 20"
		49.94%	**46.49%**				**45.06%**	**54.46%**					**50.54%**	**48.98%**			**35%**

Bush precincts in 2000 = 9

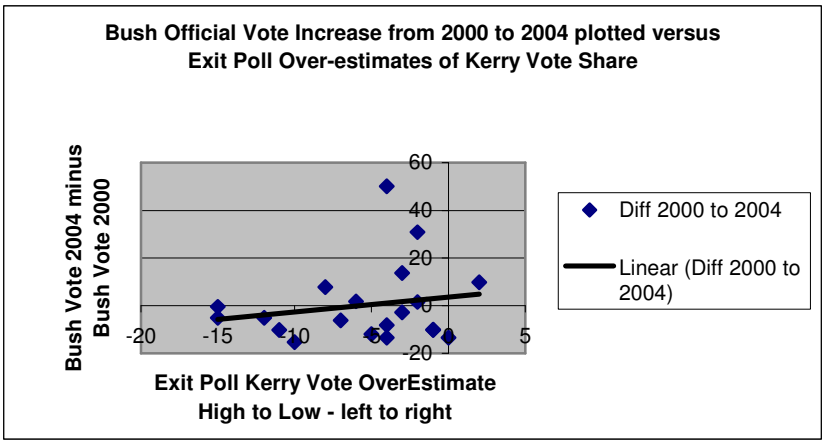
Bush precincts in 2004 = 8

Bush precincts in 2004 = 8

See Appendix A for a full math logic proof of the invalidity of ESI's inference, based on this example.

Clearly, Bush can win *due to* vote fraud in 2004 and win with no vote fraud in 2000 and yet *not* have an increase in his vote share from 2000 to 2004 in those same precincts, and, particularly, not have a Bush vote share that has a positive linear relationship with exit poll overestimates of the Kerry official vote share.

The chart below plots the Kerry exit poll overestimate versus the Bush official vote share increase from 2000 to 2004 from the table above. Contrary to Mitofsky/ESI's claim, there is a small positive correlation for Bush official vote share increase with Kerry exit poll overestimates *decreasing*, yet this is clearly an example of how vote fraud favoring Bush could change the outcome of the 2004 election.



ESI's Hypothesis Cannot Rule Out Vote Fraud in Any Exit Poll Data

Clearly, by giving counterexamples, we have shown that A does not imply $B \wedge C$, so that ESI's hypothesis that $A \rightarrow (B \wedge C)$ does not pass a simple logic analysis and thus *cannot* be validly used, as ESI, Mitofsky, and others claim it can, to rule out vote fraud.

The entire June exit poll analysis by ESI of the Ohio exit polls and the analysis of the national exit poll data that Mitofsky presented to the American Statistical Association's October 14th meeting are *not* logically valid, and *vote fraud cannot be ruled out* as they purport.

Logically Corrected Versions of ESI's Hypothesis

Corrected - hypothesis #2: $(A \wedge B) \rightarrow C$

> "If systematic fraud or error in vote counting [favoring Bush] occurred [in precincts] in 2004 but not in 2000, [AND] Bush did significantly better in those precincts in 2004" [THEN] we would see larger exit poll disparities [in 2004 than in 2000] in those precincts [where vote fraud occurred and Bush did better in 2004]."

While these particular precincts in 2004 do display high pro-Kerry discrepancies between exit poll and election results (for instance, there are over 2.5 times more precincts with pro-Kerry discrepancy than with pro-Bush discrepancy in this group of precincts which gave Bush his victory in Ohio), ESI/Mitofsky does not provide the unadjusted exit poll discrepancies data for these same precincts for 2000, so no test can be performed of this hypothesis.

ESI's hypothesis could be logically corrected in another way.

Corrected - hypothesis #3: $(A \wedge \neg B) \rightarrow C$

> "If systematic fraud or error in vote counting [favoring Bush] occurred in 2004 but not in 2000, [AND] Bush did significantly worse in those precincts in 2004" and [then] we would see larger differences between the reported vote and exit poll [in 2004 than in 2000] in those precincts [where vote fraud occurred and Bush did worse in 2004]."

While ESI's 2004 exit poll data clearly shows that there are over 3.5 times as many pro-Kerry as pro-Bush discrepancy precincts in precincts Bush where did worse in 2004 than in 2000, ESI/Mitofsky does not provide any data on the 2000 unadjusted exit poll discrepancies which could be used for this test either.

Because we showed that both $(A \wedge \neg B) \rightarrow C$ and $(A \wedge B) \rightarrow C$ are logically correct hypotheses, the simplest logically correct version of ESI's hypothesis is simply $A \rightarrow C$.
Corrected - hypothesis #4: $A \rightarrow C$

> "If systematic fraud or error in vote counting occurred in 2004 but not in 2000, [then] we would see larger differences between the reported vote and exit poll [in 2004 than in 2000] in those precincts."

Although, in 2004, there are almost four times as many precincts (23 precincts) with over 5% pro-Kerry discrepancy as there are precincts with over 5% pro-Bush discrepancy (6 precincts), we are missing the unadjusted exit poll discrepancies for 2000, to test this hypothesis.

A Corrected Hypothesis Has Other Pitfalls

Even if we did have the unadjusted exit poll discrepancy data for the 2000 election, and we formulate a corrected hypothesis to compare the 2000 with the 2004 elections*, there would be many other factors that would have to be addressed for a valid analysis, including* whether:

1. the 2000 election results are accurate;
2. other (not Democratic and Republican) party influence was the same in the two elections (i.e. that Nader voters did not shift their votes to a particular party)
3. voting patterns and demographics did not change between 2000 and 2004;
4. increased voter turnout would not favor either candidate (there was higher turnout in 2004);
5. the *2004* exit-polled precincts adequately represent the *2000* vote;
6. the *2004* exit-polled precincts are the same *geographic precincts* that they were in 2000;
7. the weights that are used to calculate the state vote share from precinct shares in 2004 are the same as those used in 2000; and
8. given all this, that zero is the correct comparison point for the difference of Bush's vote share in two elections.

Thus ESI's own hypothesis and its Figure 3 data is a fundamentally inappropriate method (at least without much more analysis based on data that has not been provided despite repeated requests[11]) for analyzing the issue of vote fraud in 2004.

[11] See footnote at beginning of report, op. cit. Even if the data were provided, pending the results of a comprehensive analysis, a comparison with 2000 officially reported results for the 2004 exit-polled precincts could well prove to be of little inferential value.

The English-Language Explanation: Why is ESI's Analysis Illogical?

ESI stated that:

> IF systematic fraud or error in vote counting [favoring Bush] occurred [in precincts] in 2004 but not in 2000, THEN Bush would have done significantly better in *those* precincts in 2004 [than in 2000], and AND we would see larger differences between the reported vote count and exit poll results in *those* precincts [than in other 2004 exit-polled precincts].

The problem with ESI's statement is that it incorrectly states that:

> IF systematic fraud or error in vote counting [favoring Bush] occurred [in precincts] in 2004 but not in 2000, THEN Bush would have done significantly better in *those* precincts in 2004 [then in 2000]

ESI is saying that there should be a correlation between the size of exit poll disparity and Bush improvement in precincts. In other words, ESI's idea is that in precincts where there was fraud we would expect better Bush performance than in 2000. However, what if the Democrats won the 2004 turnout battle big-time so that the effect of the fraud in those precincts was to rescue Bush from a worse performance than in 2000 and bring him up to even?

Just because a candidate beat his prior vote share in fewer precincts than in a prior election, does *not* show that there was no vote fraud. If so, then the case for "no vote fraud" favoring Bush would already be made for Ohio because Bush improved his vote share from the 2000 election in only 31% of Ohio's precincts.

But that would obviously be a false conclusion because:

- a candidate could do better in fewer precincts than in the last election due to vote fraud in a few more populated urban precincts with more votes, or
- vote fraud does not have to increase the vote share for a candidate from a prior election at all because voters may have voted more for the opposing candidate in the second election, yet not all their votes were counted accurately. In other words, rather than increase Bush's vote share above what it was in 2000, vote fraud may have masked the increase in Kerry's vote share over Albert Gore's in 2000.

There are numerous examples of ways in which a candidate can win an election due to vote fraud, and yet not increase either his overall vote share or the number of precincts where his vote share increased from a prior election. The section in this paper "ESI's Hypothesis, as Stated and Analyzed by ESI, is Logically Invalid" gives a counterexample that uses percentages similar to those in Ohio's 2000 and 2004 elections.

ESI's and Mitofsky's exit poll analyses claiming to "kill the vote fraud argument" are invalid because their analysis is logically invalid. If they restated the hypothesis on which they base their analysis in a logically correct way, it would require a comparison of unadjusted exit poll discrepancies in 2004 with those of 2000 and would require that many other factors be taken into consideration that they have neglected thus far to mention in their analyses.

Conclusion

Now that we have debunked as illogical the very basis for ESI's and Mitofsky's exit poll analysis, it is important that readers examine a valid analysis of the Ohio precinct-level exit poll analysis, **"**The Gun is Smoking: Ohio Exit Poll Data Provides Clear Evidence of Vote Miscount or Unexplained and Implausible Exit Poll Error**"**[12] published by the National Election Data Archive. It concludes that

> *The limited precinct level Ohio exit poll data that has been released, show impossible and nearly impossible, precinct level exit poll results, and highly irregular patterns of exit poll discrepancies that ... are consistent with a hypothesis of pro-Bush vote miscounts.*

Readers should refer to the October 25, 2005 "*History of the Debate Surrounding the 2004 Presidential Election*"[13] published by the National Election Data Archive, to note the amount of research, analysis, and publishing that has been required to counter the incorrect hypotheses that Mitofsky, Liddle, and the Election Science Institute have put forth which misled the press and the public into accepting without adequate consideration that the 2004 presidential exit poll disparities are not evidence of vote fraud. We ask that *no more* illogical and incorrect hypotheses and analyses be forthcoming, and that the parties involved logically and analytically check and mathematically verify their hypotheses and explanations of the 2004 exit poll discrepancies *before* publicly releasing them. American democracy demands such a level of responsibility from its mathematicians and scientists for analysis of its elections.

Time and resources could, without question, be much better be spent on implementing a national election data archive system to ensure that detailed exit poll and election data showing evidence of vote miscounts be immediately made public after polls close and thus enable Americans to ensure that only validly elected candidates are sworn into office.

Until routine independent audits of vote counts are performed, a national election data archive may be the only tool that could ensure that correctly elected candidates are sworn into office. If steps to implement it are begun today, it could be operable by November 2006. Not one state in America yet collects the needed detailed election data at the state level, so it must be collected from over 3,300 separate county and township offices with the help of volunteers and election officials. Most county and township election offices do not currently release election results data in the detail that is needed to mathematically detect precincts where vote miscounts have probably taken place. Such detailed analysis has successfully pinpointed problems with vote counts in New Mexico and Washington State.[14] Voters have a legal right to the data under every state's freedom of information act, that we could use to ensure that correctly elected candidates are sworn into office following elections. Until routine independent audits of vote count accuracy are performed in every state of every election and race, by hand-counting voter verified paper ballots in a small percentage of precincts[15], mathematical analysis of election data will be the *only* way to detect and correct vote count errors in time to challenge the results if necessary.

---

[12] by Ron Baiman, David Dodge, and Kathy Dopp at http://uscountvotes.org/ucvAnalysis/US/exit-polls/USCV_exit_poll_analysis.pdf

[13] See http://electionarchive.org/ucvAnalysis/US/Presidential-Election-2004.pdf

[14] See this history of the academic debate surrounding the 2004 presidential election: http://electionarchive.org/ucvAnalysis/US/Presidential-Election-2004.pdf

[15] See this scientific paper on how to independently audit vote count accuracy: http://electionarchive.org/ucvAnalysis/US/paper-audits/Paper_Audits.pdf

Appendix A: Symbolic Math Logic Proof - Based on Counterexample #2

Let P = set of Ohio precincts for which exit polls were conducted. From the table, $P = \{1, 2, 3...19, 20\}$ The number of precincts in P, denoted by $[P]$, is 20. ($[P] = 20$)

Let $V_{p,2004,B}$ = number of votes in precinct p in 2004 for Bush

Let $V_{p,2004}$ = number of votes in precinct p in 2004 for all presidential candidates

Let $V_{p,2000,B}$ = number of votes in precinct p in 2000 for Bush

Let $V_{p,2004}$ = number of votes in precinct p in 2000 for all presidential candidates

The above data set can be used as a counter example to the claim that **A** $\rightarrow$**(B** $\wedge$ **C).** Using two different precinct-count-based definitions of **B** ("Bush did better in more precincts"), we can show that **A** $\wedge\neg$ **B,** and therefore that **A** $\rightarrow$**(B** $\wedge$ **C)** is false.

Definition of B (#1): "Bush did better in more precincts" ==

> ***The # precincts for which Bush's vote share increased from 2000 to 2004*** *([P_B])* ***is greater than the # precincts for which Bush's vote share <u>did not</u> increase from 2000 to 2004*** *([P_{B^-}]).*
> *So* **B** : $\left[P_B\right] > \left[P_{B^-}\right]$

Definition of B (#2): "Bush did better in more precincts" ==

> ***The # precincts in which Bush's vote share exceeded Kerry's in 2004*** ($\left[P_{B,2004}\right]$) **is** ***greater than the # precincts in which Bush's vote share exceeded Gore's in 2000*** ($\left[P_{B,2000}\right]$)**.** ***So*** **B** : $\left[P_{B,2004}\right] > \left[P_{B,2000}\right]$.

Definition of B (#1): "Bush did better in more precincts" ==

> ***The # precincts for which Bush's vote share increased from 2000 to 2004*** *([P_B])* ***is greater than the # precincts for which Bush's vote share <u>did not</u> increase from 2000 to 2004*** *([P_{B^-}]).*
> *So* **B** : $\left[P_B\right] > \left[P_{B^-}\right]$

Let P_B = set of precincts in P for which Bush's precinct vote share increased from 2000 to 2004. (Clearly $P_B \subset P$)

Let $\frac{V_{p,2004,B}}{V_{p,2004}}$ be the 2004 vote share for Bush in precinct p

and $\frac{V_{p,2000,B}}{V_{p,2000}}$ be the 2000 vote share for Bush in precinct p.

Then $P_B =$ set of precincts p for which $\frac{V_{p,2004,B}}{V_{p,2004}} > \frac{V_{p,2000,B}}{V_{p,2000}}$

Let $P_{B^-} =$ set of precincts in P for which Bush's precinct percentage did not increase from 2000 to 2004. So P_{B^-} is the complement of P_B. (Thus we have $P_{B^-} \subset P$, $P_B \bigcap P_{B^-} = \varnothing$ and $P_B \bigcup P_{B^-} = P$)

Given the set of data found in Table 1, in 2004, "Fraud" has been introduced in each of the 20 precincts by approx 7% up/down for Bush/Kerry, respectively. So we have the condition **A.**

Observe the indicator column of the table. The rows in which the indicator is "1" are those in which the vote share for Bush increased between 2000 and 2004. The 7 precincts in which this occurs are those in the set $P_B = \{2,4,5,12,15,16,17\}$. In particular $\left[P_B\right] = 7$

Now, since P_{B^-} is the complement of P_B, then $P_{B^-} = \{1,3,6,7,8,9,10,11,13,14,18,19,20\}$ and $\left[P_{B^-}\right] = 13$

Since 7 is less than 13, $\left[P_B\right] < \left[P_{B^-}\right]$, or $\neg$ **B.**

Thus, using "precinct-count-based" Definition #1 (for **B**), we have: **A** $\wedge\neg$ **B.**

Definition of B (#2): "Bush did better in more precincts" ==

> ***The # precincts in which Bush's vote share exceeded Kerry's in 2004*** ($\left[P_{B,2004}\right]$) **is *greater than the # precincts in which Bush's vote share exceeded Gore's in 2000*** ($\left[P_{B,2000}\right]$)**. *So* B**: $\left[P_{B,2004}\right] > \left[P_{B,2000}\right]$.

Let $P_{B,2004} =$ set of precincts in P in 2004 for which Bush's vote share was greater than that for Kerry.

Let $P_{B,2000} =$ set of precincts in P in 2000 for which Bush's vote share was greater than that for Gore.

Given the set of data found in Table 1, in 2004, "Fraud" has been introduced in each of the 20 precincts by approx 7% up/down for Bush/Kerry, respectively. So we have the condition of **A.**

Observe that

$P_{B,2004} = \{2,4,5,12,15,16,17,18,19,20\}$ and $\lceil P_{B,2004} \rceil = 8$

$P_{B,2000} = \{12,13,14,15,16,17,18,19,20\}$ and $\lceil P_{B,2000} \rceil = 9$

Since 8 is less than 9, $\lceil P_{B,2004} \rceil < \lceil P_{B,2000} \rceil$, or $\neg$ **B.**

Thus, using "precinct-count-based" Definition #2 (for **B**) we have: **A** $\wedge\neg$ **B.**

Regardless of the definitions of **B** based on precinct counts (#1 or #2), we have a set of data in which **A** $\wedge\neg$ **B.**

Given A $\wedge\neg$ **B,** we can show that it <u>cannot</u> be true that **A** $\rightarrow$**(B** $\wedge$ **C).**

Proof:

Suppose **A** $\rightarrow$**(B** $\wedge$ **C).**

Then, clearly **A** $\rightarrow$ **B.**

Then, by contrapositive**,** $\neg$**B** $\rightarrow$ $\neg$ **A.**

Since there is an example (see above) where **A** $\wedge\neg$ **B ,** then, since $\neg$**B** $\rightarrow$ $\neg$**A,** we have **A** $\wedge\neg$ **A.** *Contradiction***!**

Therefore, it must <u>NOT</u> be the case that **A** $\rightarrow$ **B.**

Since **A** $\rightarrow$ **B** is false, then clearly **A** $\rightarrow$**(B** $\wedge$ **C)** is false.