

Polling Bias or Corrupted Count?

Accepted Improbabilities and Neglected Correlations in 2004 US Presidential Exit Poll Data

Steve Freeman, Center for Organizational Dynamics
University of Pennsylvania

American Statistical Association, Philadelphia Chapter
Fall Meeting, October 14, 2005

Two Questions to Consider Today

1. How much confidence do we have in the official count – and the explanations about the exit poll discrepancy?

2. How much doubt must be raised about the count – and the explanations about the exit poll discrepancy – before we demand answers?

John Roberts on voting rights during his confirmation hearings

The right to vote [and have that vote counted as cast] ... is preservative of all other rights

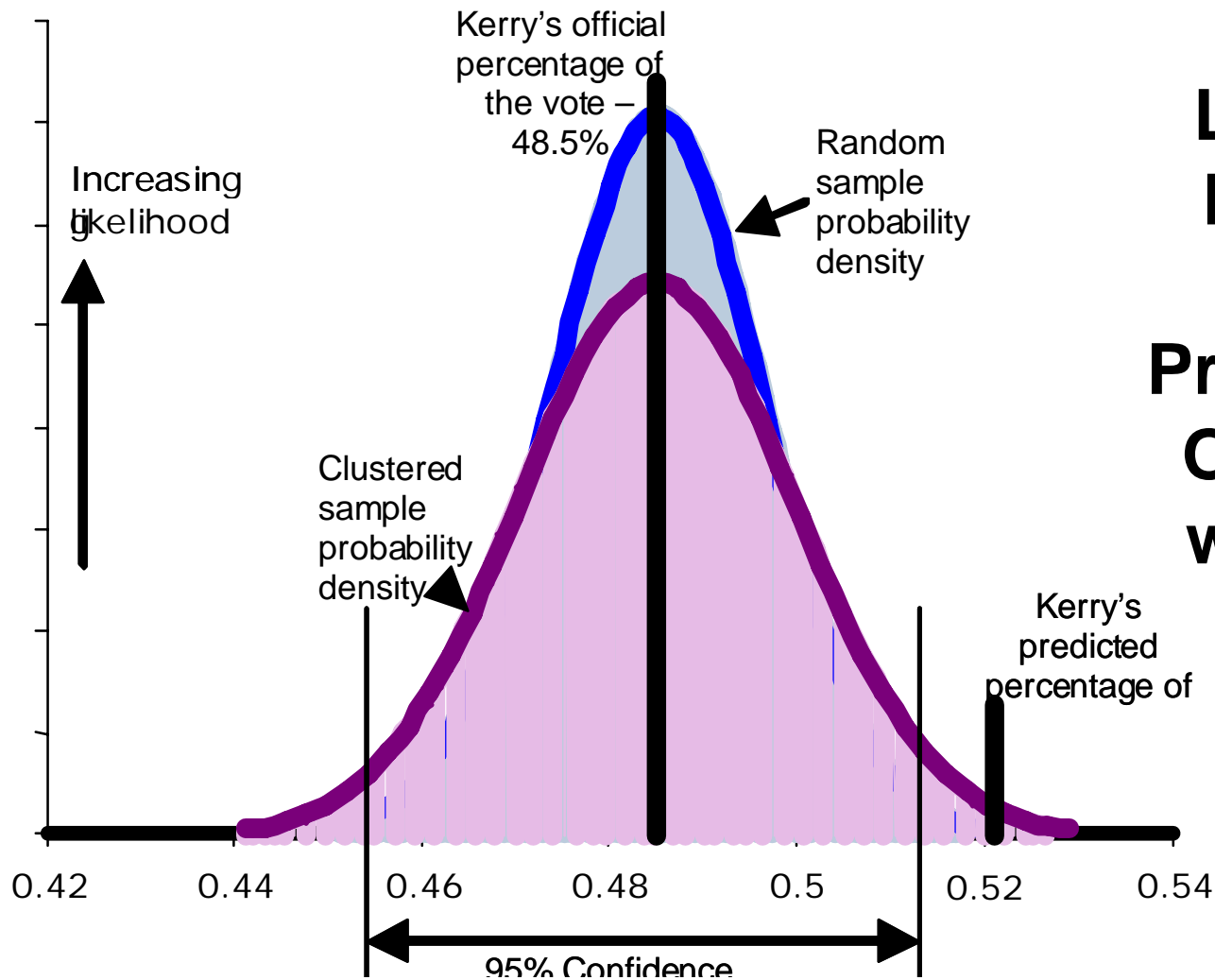
The Voting Rights Act was extremely valuable in securing not just the right to vote but all other rights derivative of that.

Election Night Nov 2, 2004: Exit Poll vs. Official Percentages in Battleground States

	Bush exit poll	Kerry exit poll	Exit Poll differential	Bush official	Kerry official	Official differential	Official vs. Exit Poll
Colorado	49.9%	48.1%	Bush 1.8	52.0%	46.8%	Bush 5.2	Bush 3.4
Florida	49.8%	49.7%	Bush 0.1	52.1%	47.1%	Bush 5.0	Bush 4.9
Iowa	48.4%	49.7%	Kerry 1.3	50.1%	49.2%	Bush 0.9	Bush 2.2
Michigan	46.5%	51.5%	Kerry 5.0	47.8%	51.2%	Kerry 3.4	Bush 1.6
Minnesota	44.5%	53.5%	Kerry 9.0	47.6%	51.1%	Kerry 3.5	Bush 5.5
Nevada	47.9%	49.2%	Kerry 1.3	50.5%	47.9%	Bush 2.6	Bush 3.9
New Hampshire	44.1%	54.9%	Kerry 10.8	49.0%	50.3%	Kerry 1.3	Bush 9.5
New Mexico	47.5%	50.1%	Kerry 2.6	50.0%	48.9%	Bush 1.1	Bush 3.7
Ohio	47.9%	52.1%	Kerry 4.2	51.0%	48.5%	Bush 2.5	Bush 6.7
Pennsylvania	45.4%	54.1%	Kerry 8.7	48.6%	50.8%	Kerry 2.2	Bush 6.5
Wisconsin	48.8%	49.2%	Kerry 0.4	49.4%	49.8%	Kerry 0.4	No dif

Probability that 10 out of 10 states would favor the President: 1 out of 1,024

Election Night Nov 2, 2004: Official Vote Count in Ohio and the Exit Poll Projection



Likelihood of Kerry's Ohio Exit Poll Projection if the Official Count were Correct:
 $p = .0073$

About the Nov 2, 2004 Exit Poll

- Conducted by Mr. Mitofsky for the National Election Pool, a consortium of the major US media corporations and associations: CBS, NBC, ABC, CNN, Fox & AP.
- At a representative precincts across each state, every n th voter exiting the polling place is asked to fill out a confidential questionnaire. The interval is chosen so that approximately 100 interviews will be spread evenly over the course of the day.

Vote Suppression/ Manipulation

- Vote Spoilage: 2-3% of Ballots lost nationally, overwhelmingly concentrated in poor, black voting districts
- Disenfranchised Felons: 800,000 in Florida (7% of VAP) not eligible to vote (1/3 of African American males)
- Faux-Felon lists: 82,389 voters wrongly purged
- Democratic Absentee ballots rejected (signature)
- GOP registration groups discard Democratic forms
- Ohio registrations rejected if not on precise paper-weight
- Long lines – 11 hours at Oberlin U.; 7 hours in black precincts of Columbus

Vote Suppression Success

US ranks #139 out of 172 nations in the world in
Federal election turnout: 48.3% vs. Europe 80%+

- Registration obstacles
- Negative campaigns
- Limited Choice (Republocrats)
- Limited Effect (Congressional Gerrymandering, Electoral Vote)
- No time off of work
- Long Lines

Election Turnout around the World

(Average from 1945-1998)

Country	%vap	Country	%vap
1 Italy	92.5	135 Cent African Rep	50.3
4 Iceland	89.5	136 Antigua&Barbuda	50.2
6 Indonesia	88.3	137 Burma/Myanmar	50.0
8 New Zealand	86.2	138 Switzerland*	49.3
12 Austria	85.1	139 USA**	48.3

* Switzerland: Women not permitted to vote until 1971

** Ranking based on turnout, not ballots counted (percentage of votes counted is lower)

Vote Manipulation by Campaign Managers / Chief Election Officers

Florida 2000 (Katherine Harris)

- Faux-Felon lists
- Differential treatment of absentee ballots
- Impossibly tight recount deadlines
- Disregard law to acknowledge voter Intent
- Disregard law to conduct machine recount (1/4 of the state ballots)
- Miami manual count obstructed

Ohio 2004 (Ken Blackwell)

- Arbitrary guidelines
- Unequal distribution of voting machines
- Diversion of HAVA funding to observers, who were there to challenge voting qualifications.
- Failure to conduct recount
- Obstruction of judicial review
- Attempt to disbar attorneys that challenged the process

Ohio Vote Fraud

- Secret Count in Warren County due to “FBI terrorist alert,” an alert denied by the FBI
- 25% spoilage in black precincts of Republican Montgomery Co.
- Systematic vote switching in Cuyahoga Co. (Cleveland): Ballots with candidate position “rotation” tabulated at wrong precincts – Big net Dem loss
- Appalachian precincts with 90%+ turnout rates and more far votes cast than recorded voters
- Ghosts in the Machines – voters tried to vote for Kerry, Bush’s name came up

Ohio Vote Fraud (cont.): Non-Recount throughout the state

Throughout the state: The requirements for the selection of a random recount sample were clearly laid out in Directive 2004-58. Instead of following these guidelines, staff members were instructed to purposefully select certain precincts that would easily balance during a hand recount.

- Hocking County's Cheat Sheet – so workers would know what the numbers should add up to. TRIAD presence – to “ the machines in anticipation of the recount.
- Clermont Co.'s Optical-Scan Scam – stickers covering up Kerry ovals (when no stickers were used at the polls)

From a Hocking Co. affidavit:

One week before the recount, a Triad rep came 'to check out [our] tabulator' ... He said that the stored information was gone [but that] he could put a patch on it ... Our computer is 14 years old and always worked in the past....

[He] then asked me which precinct we were going to count. I told him... He went back into the tabulation room. Shortly after that he stated that the computer was ready for the recount ... He said not to turn the computer off until after the recount. He advised [us] on how to post a 'cheat sheet' on the wall so that only [we] would know about it ... so the count would come out perfect and we wouldn't have to do a full hand recount of the county.

Electronic Voting: An Invitation for Mass Scale Electoral Fraud

- No confirmation at all that votes are counted as cast. It's as though you're asking a Man Behind a Curtain to faithfully record your vote.
- Easy to Manipulate (e.g., Easter eggs)
- Qualities of the e-voting companies
 - NOT impartial (Diebold chair a major Bush supporter)
 - Criminal convictions and civil suit settlements
 - conflict-of-interest (e.g., Sen Hagel (R-Neb))
 - No transparency in operations (all proprietary)

Electronic Voting Machines Compared to Las Vegas Slot Machines

	Las Vegas Slot Machines	Electronic Voting Machines
Software	State has access to all software. Illegal to use software that is not on file.	Software is a trade secret.
Spot-checking	Board inspectors show up unannounced at casinos to compare computer chips to those on file. If a discrepancy, the machine is shut down and investigated.	No checks are required. election officials have no chip to compare to the one found in the machine.
Standards	Meticulous standards (e.g., machine must work when subjected to a 20,000-volt shock) constantly updated.	Two-years old standards with gaping security holes
Back-ground Scrutiny	Manufacturers subjected to background checks. Employees are investigated for criminal records.	Citizens have no way of knowing, for example, if e-voting programmers have been convicted of fraud.
Equipment Certification	By a public agency at arms-length from manufacturers. Public questions invited.	By for-profit companies chosen and paid by the manufacturers. No public information on how the testing is done.
In the event of disputes ...	Casino must contact the Gaming Board, which has investigators on call around the clock. Investigators can open up machines to inspect internal mechanisms, and records of recent gambling outcomes.	In most cases a voter's only recourse is to call a board of elections number that may or may not work to lodge a complaint that may or may not be investigated.

Snohomish County's Parallel Systems and Divergent Results

Absentee Ballots cast on paper/optical scan (2/3 of vote)
Election Day on electronic voting machines (1/3 of vote)

Gregoire wins on paper; Rossi wins big on electronic voting machines, bigger yet in precincts with problems and machines with maintenance

Voting Technology, Condition	Republican Candidate	Gubernatorial Candidate Dino Rossi	Democratic Candidate	Gubernatorial Candidate Christine Gregoire	Winner/Margin
Paper / optical scan	95,228	49.5%	97,044	50.5%	Gregoire 1.0%
Electronic voting machines	50,400	54.5% (+5.0%)	42,145	45.5% (-5.0%)	Rossi 9.0%
Polling places with Election Day problems	21,847	56.1% (+6.6%)	17,100	43.9% (-6.6%)	Rossi 13.2%
Precincts with CPU changes	4,237	58.1% (+8.6%)	3,050	41.9% (-8.6%)	Rossi 16.2%
Malfunctioning DREs	155	60.5% (+12%)	101	39.5% (-12%)	Rossi 21.0%

Precinct Level Disparity (PLD)

The difference between how people said they voted in confidential questionnaires as they walked out of the voting booth, and the way those votes were officially recorded.

E/M calls this term, WPE (Within Precinct Error) or more specifically IM WPE, the only measure they report that does not exclude outliers precincts with the larger deviations.

Official vs Exit Poll Survey Results

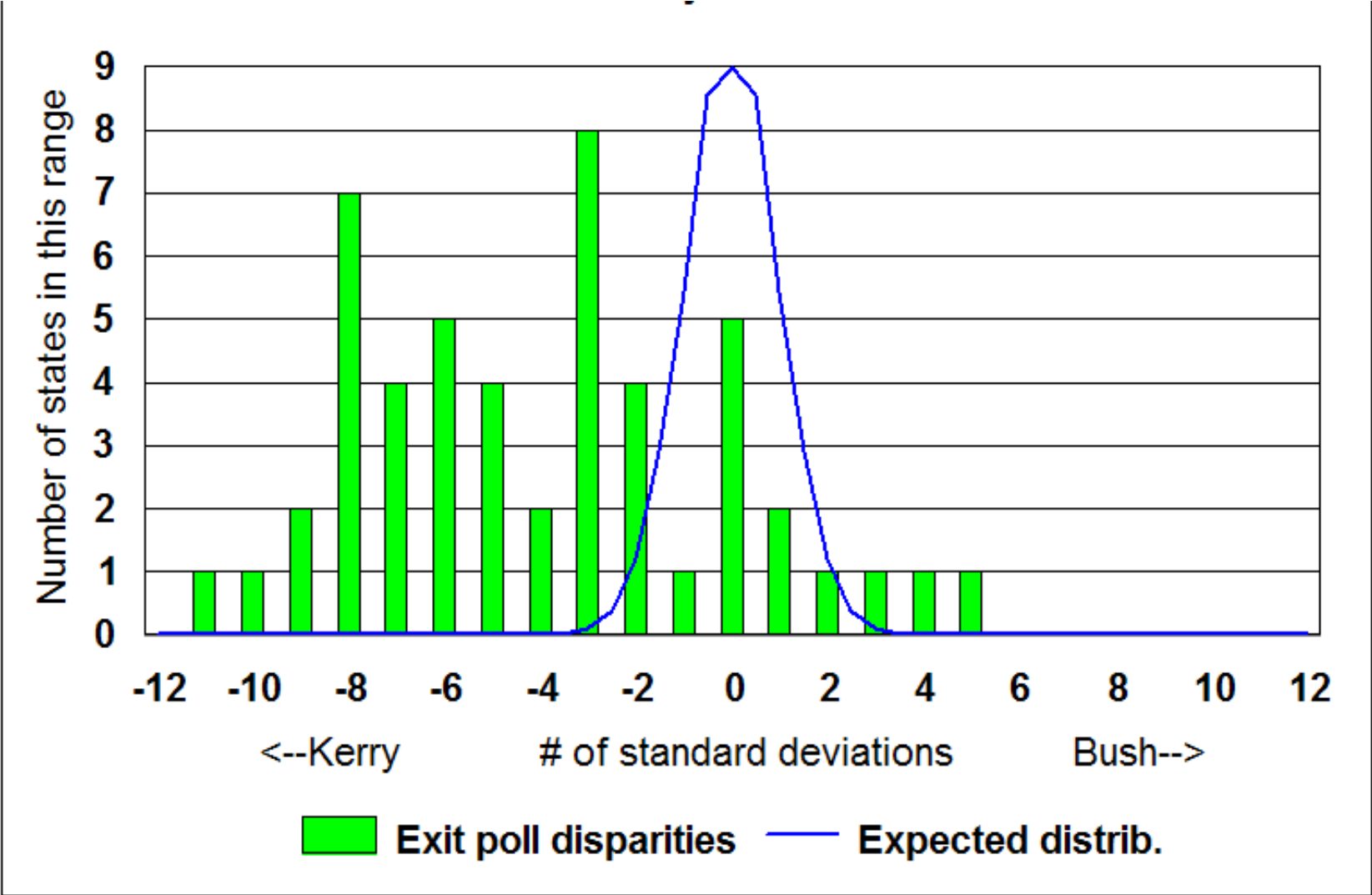
Official Result: Bush defeated Kerry by 3,000,000 votes nationally (2.5%) and a slim majority in the Electoral College.

Exit poll data based on 114,559 sampled voters at 1,460 precincts across the nation indicated that Kerry defeated Bush by 7,000,000 votes nationally (4.6%) and a decisive majority in the Electoral College.
(see handout: *The Election Outcome Based on Exit Poll Reported Voting*)

Election outcome changes if the count matched how Voters said they cast they their ballots

STATE	Bush Official Vote	Kerry Official Vote	Official Margin	(5) PLD	Exit Poll Bush	Exit Poll Kerry	Exit Poll Margin	EP EV Bush	Too Close	EP EV Kerry
* Colorado	51.7%	47.0%	4.7	-6.1	48.6%	50.1%	-1.4		*9	
** Florida	52.1%	47.1%	5.0	-7.6	48.3%	50.9%	-2.6			**27
* Iowa	49.9%	49.2%	0.7	-3.0	48.4%	50.7%	-2.3		*7	
Missouri	53.3%	46.1%	7.2	-5.8	50.4%	49.0%	1.4		11	
** Nevada	50.5%	47.9%	2.6	-10.1	45.4%	52.9%	-7.5			** 5
**New Mexico	49.8%	49.0%	0.8	-7.8	45.9%	52.9%	-7.0			** 5
No. Carolina	56.0%	43.6%	12.4	-11.3	50.4%	49.2%	1.1		15	
** Ohio	50.8%	48.7%	2.1	-10.9	45.4%	54.2%	-8.8			**20
Virginia	53.7%	45.5%	8.2	-7.9	49.7%	49.4%	0.3		13	
Total USA	50.7%	48.3%	2.5	-7.1	47.2%	51.8%	4.6	174	55	309

Distribution of Standard Deviations



Only Two Possible Sources of PLD: Non-Response Bias or Count Corruption

Edison/Mitofksy reject count corruption out of hand – despite acknowledging that, “it is difficult to pinpoint precisely the reasons that, in general, Kerry voters were more likely to participate in the exit polls than Bush voters.”
(p.4)

E/M suggests that the discrepancy is due to their interviewers. They report that PLD is higher...

- when interviewers are more than 25 feet away from the polling place
- among with younger interviewers
- among interviewers with advanced degrees
- among interviewers in large precincts

Interviewer Effects? Location

Distance	mean WPE	median WPE	mean Abs(WPE)	N
Inside	-5.3	-4.2	11.8	416
Right outside the entrance	-6.4	-7.5	13.4	207
10 to 25 feet away	-5.6	-4.2	14.0	220
25 to 50 feet away	-7.6	-7.3	14.8	150
50 to 100 feet away	-9.6	-10.3	17.7	97
More than 100 feet away	-12.3	-12.1	16.7	37

Interviewer Effects?

Interviewer Age

Interviewer Age	mean WPE	median WPE	mean Abs(WPE)	N
24 and under	-7.4	-8.6	15.0	430
25-34	-8.2	-7.2	13.4	182
35-44	-4.0	-3.9	13.4	167
45-54	-6.3	-4.7	12.5	191
55-64	-7.0	-5.8	12.6	143
65 and over	-3.7	-5.4	12.9	68

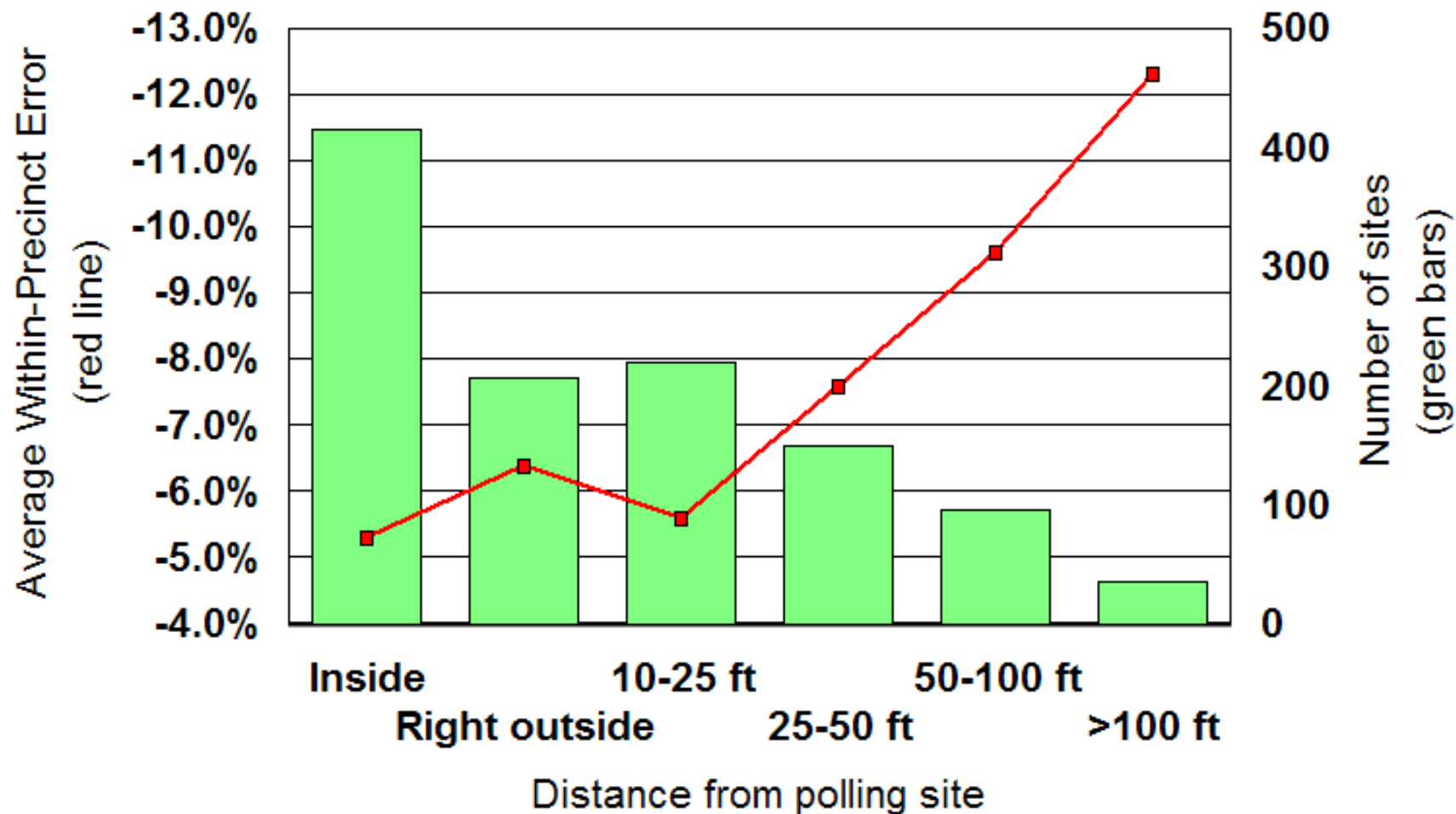
Interviewer Effects? Education

Interviewer Education	mean WPE	median WPE	mean Abs(WPE)	N
High school or less	-3.9	-4.6	14.7	177
One to three years of college	-7.3	-7.0	14.0	526
Four year college degree	-6.3	-6.3	12.8	222
Some graduate credits	-5.4	-5.9	11.9	71
Advanced degree such as MA, MBA or PhD	-7.9	-5.2	13.1	123

The completion rates tend to be slightly higher in precincts with more educated interviewers :

Interviewer Education	Completion Rate	Refusal Rate	Miss Rate
High school or less	0.52	0.36	0.11
One to three years of college	0.53	0.37	0.11
Four year college degree	0.55	0.34	0.11
Some graduate credits	0.57	0.34	0.10
Advanced degree such as MA, MBA or PhD	0.60	0.32	0.08

Error increases with distance between exit pollster and polling place



Interviewer Effects?

Interviewer Age

Interviewer Age	mean WPE	median WPE	mean Abs(WPE)	N
24 and under	-7.4	-8.6	15.0	430
25-34	-8.2	-7.2	13.4	182
35-44	-4.0	-3.9	13.4	167
45-54	-6.3	-4.7	12.5	191
55-64	-7.0	-5.8	12.6	143
65 and over	-3.7	-5.4	12.9	68

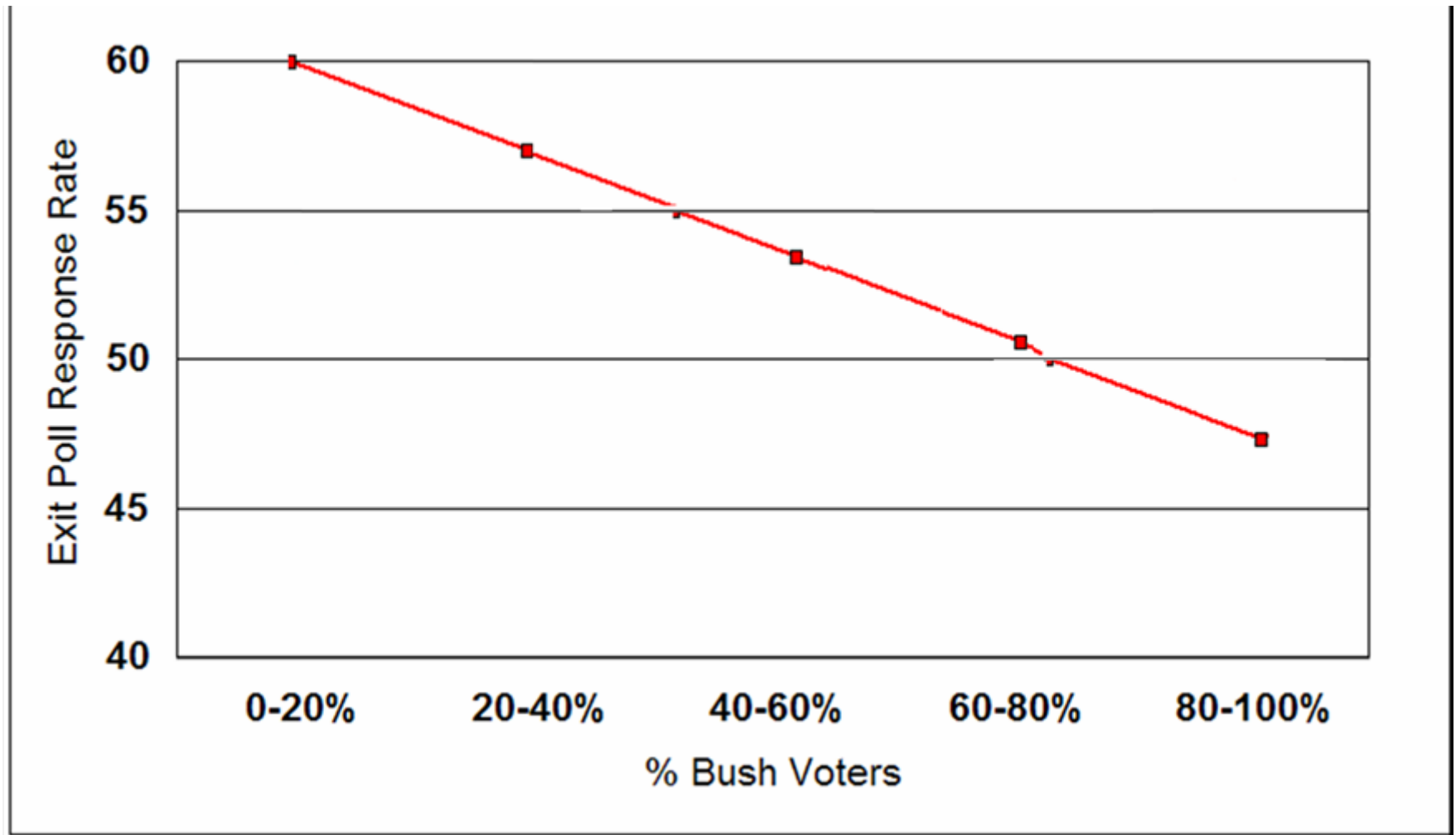
Interviewer Effects?

- Under the best scenario, Interviewer Effects only explain a small part of overall PLD
- Interviewer groups with lowest mean PLD may not be the most accurate
- Alternative explanations (e.g., distance from polling place may be attributable to corruption)
- Non-theory-based partitioning or fishing is guaranteed to produce false findings
- No data presented to permit verification of even the modest potential claims

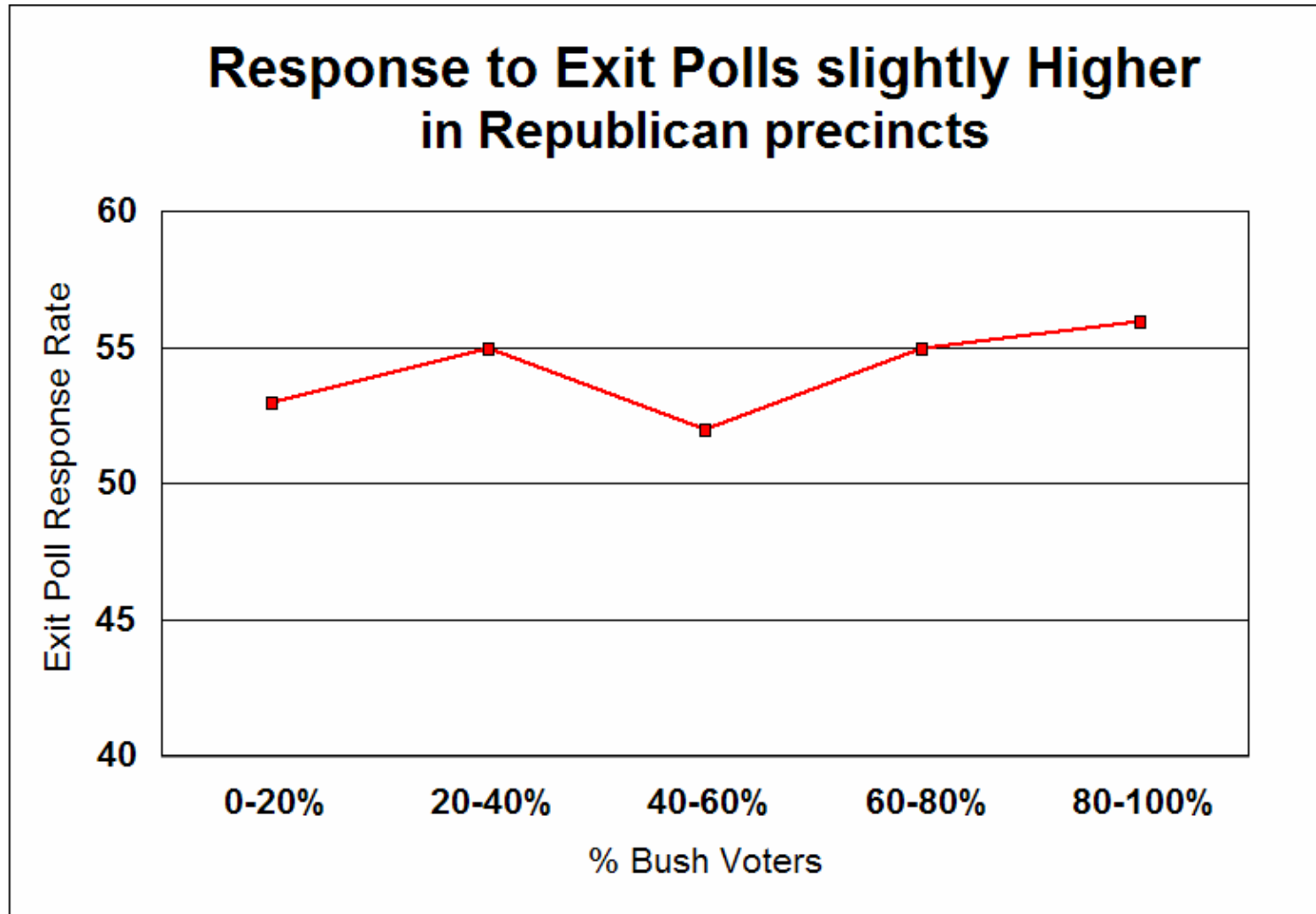
Biased Polls or Corrupted Count?

- ***Precinct level analysis***
- ***State level analysis***

If Bush voters were less likely to participate in the polls, this is the relationship we would expect to see:

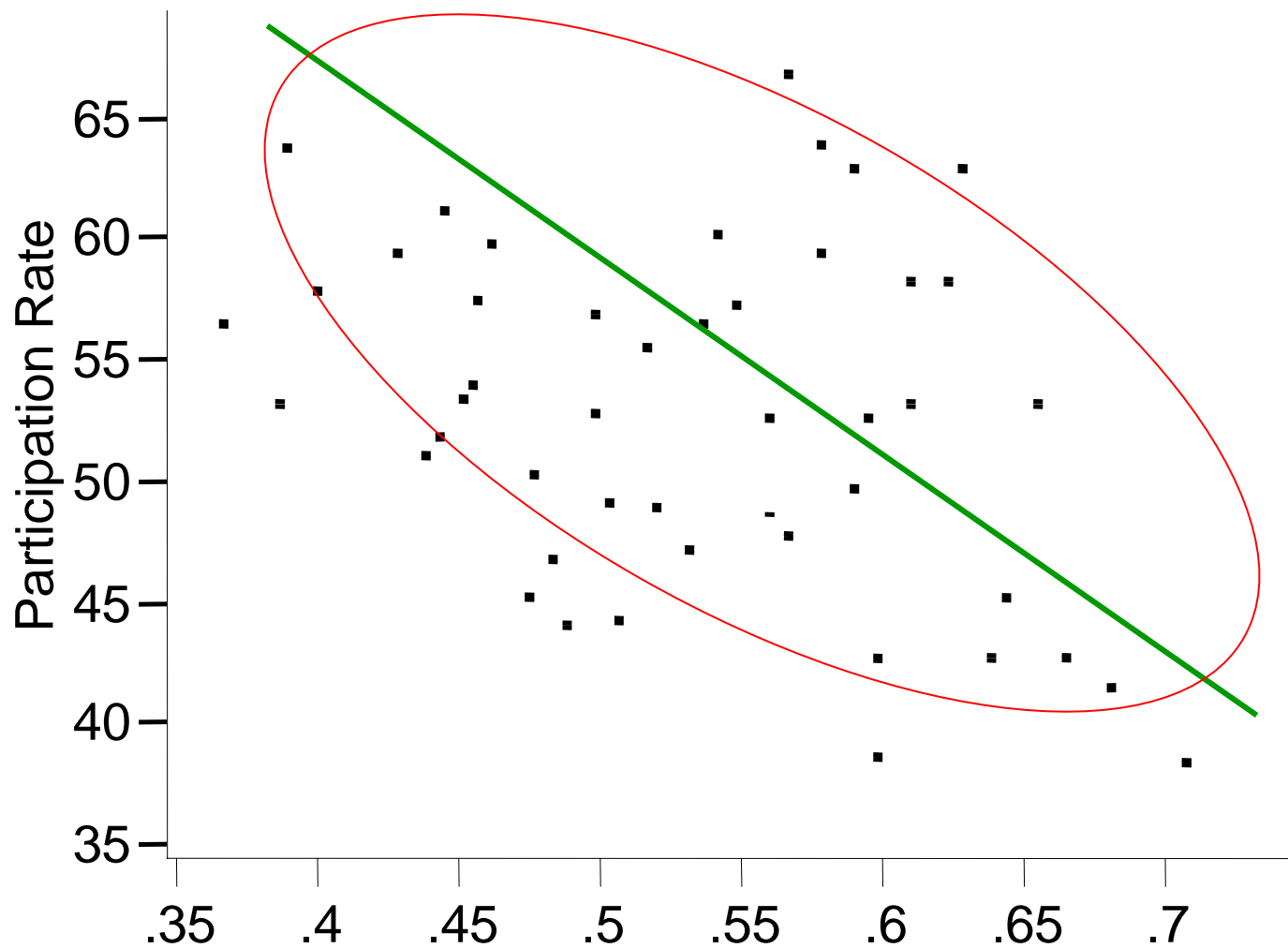


But this is what E/M's figures tell us:



Data drawn from E/M page 37

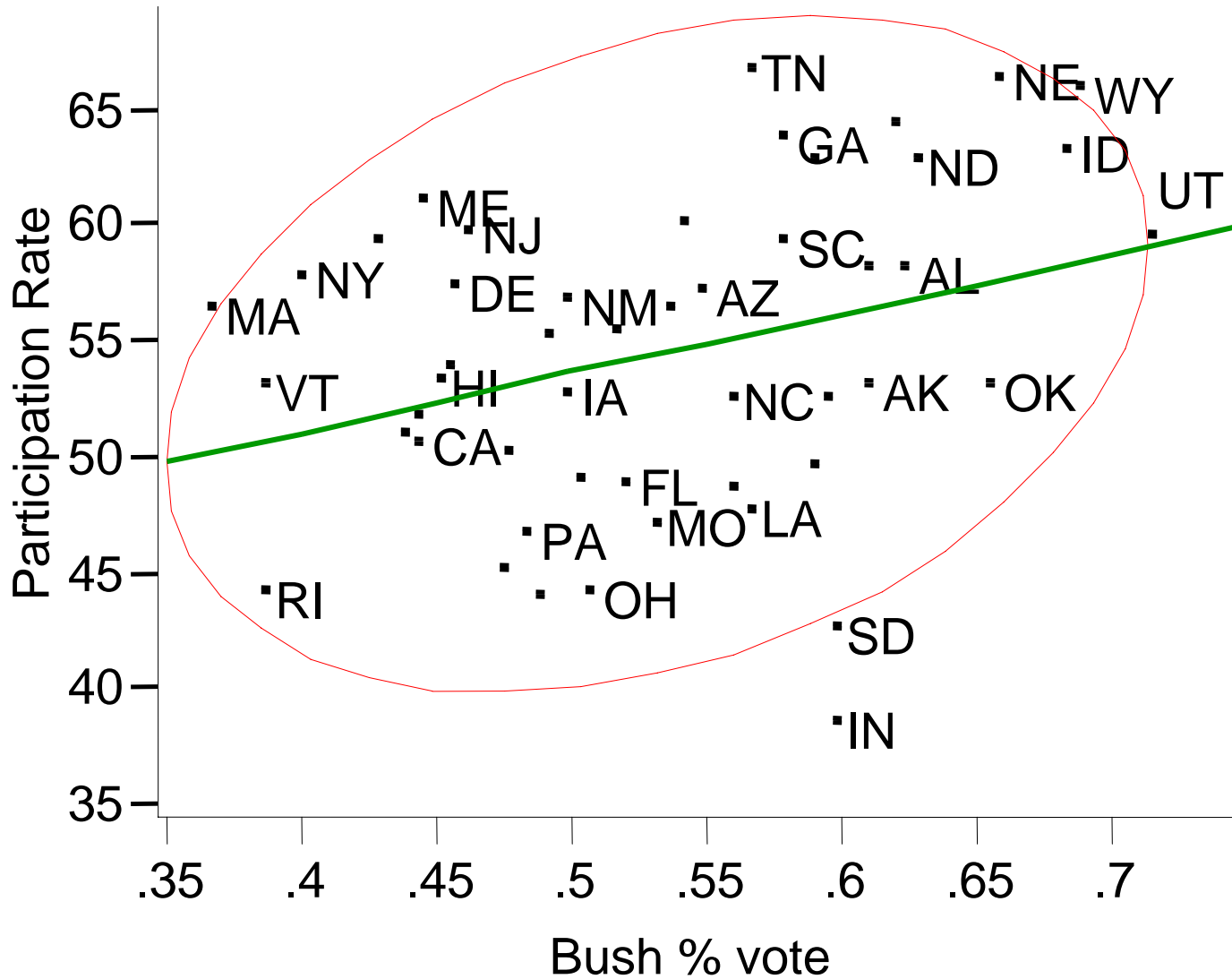
If Bush voters were less likely to participate in the polls, we would expect to see **lower** participation in state with more Bush voters:....



Bivariate Normal Ellipse P=0.900

Bush % vote Correlation = .34; p=0.0245

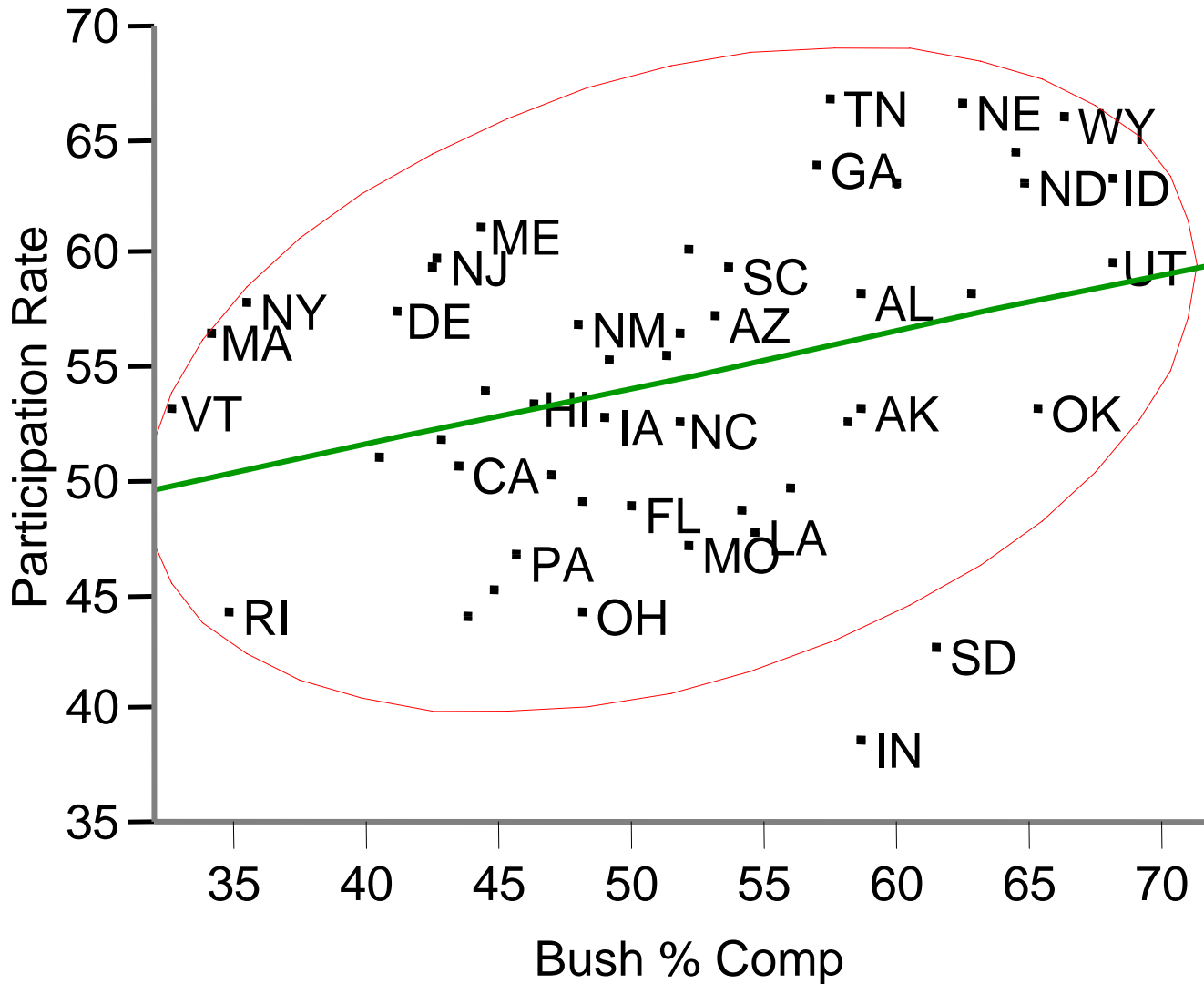
Actual Participation Rate by Bush % of State Vote



Bivariate Normal Ellipse $R=0.900$

Correlation = .54, $p=0.0245$

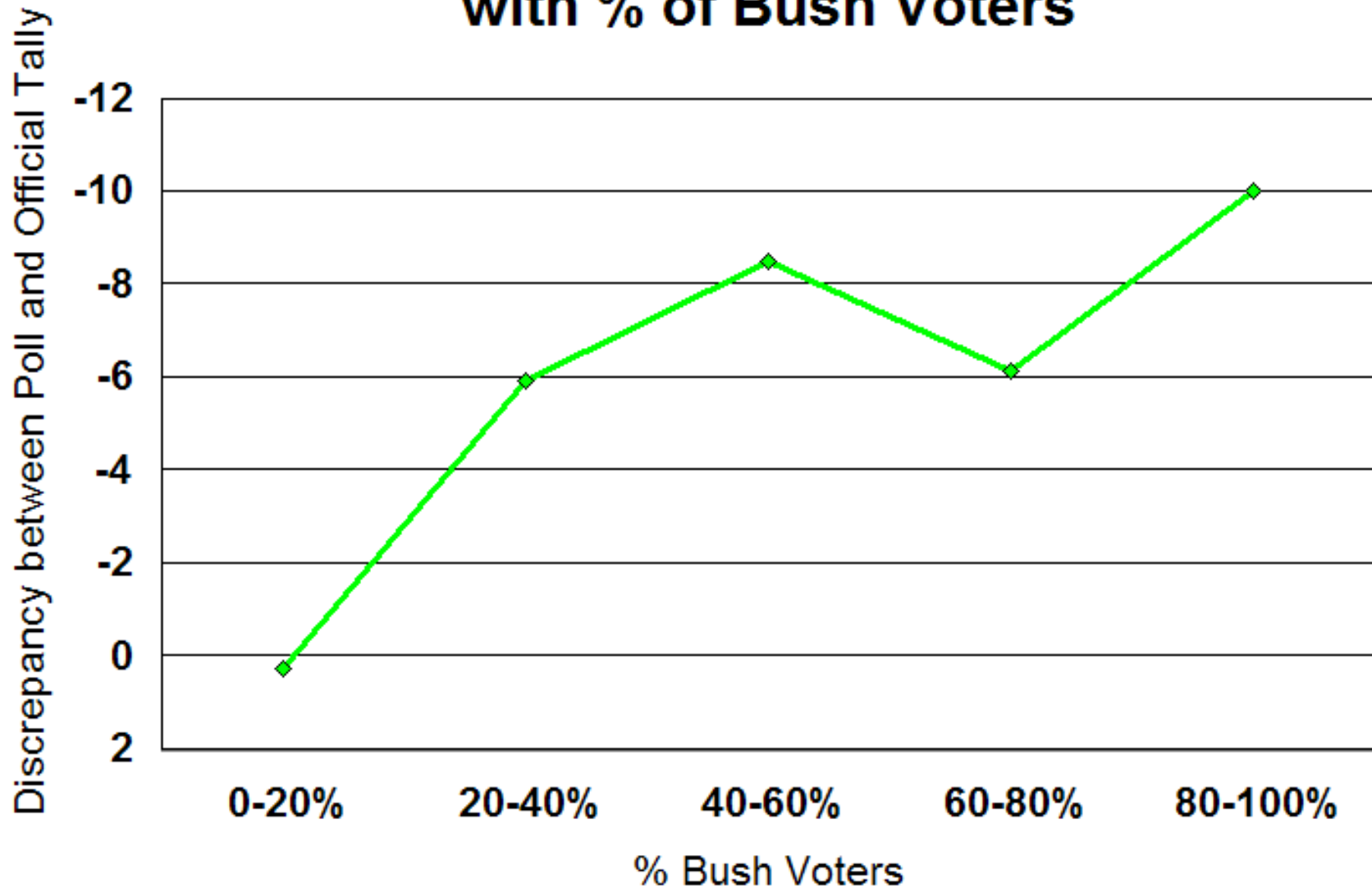
Participation Rate by Bush % Exit Poll



Bivariate Normal Ellipse P=0.900

Correlation = .34; p=0.0167

Exit Poll Discrepancy Rises with % of Bush Voters



Is PLD different for Swing States? (Edison/Mitofsky Report)

12. Swing states:

The WPE was greater in the more competitive “swing” states. For this analysis, the following were considered swing states: Colorado, Florida, Iowa, Michigan, Minnesota, Nevada, New Hampshire, New Mexico, Ohio, Pennsylvania, and Wisconsin.

Swing state	mean WPE	median WPE	mean Abs(WPE)	N
Precinct not in a swing state	-6.1	-5.1	13.1	776
Precinct in a swing state	-7.9	-8.6	14.8	474

Swing state	Completion Rate	Refusal Rate	Miss Rate
Precinct not in a swing state	0.56	0.34	0.10
Precinct in a swing state	0.50	0.38	0.11

This indicates that voters in the swing states (who were exposed to more paid advertising and media coverage than voters in non-swing states) were less likely to respond to the exit poll: but among those who did, more likely to be Kerry voters.

Is PLD different for Swing States?

State-level PLD Analysis

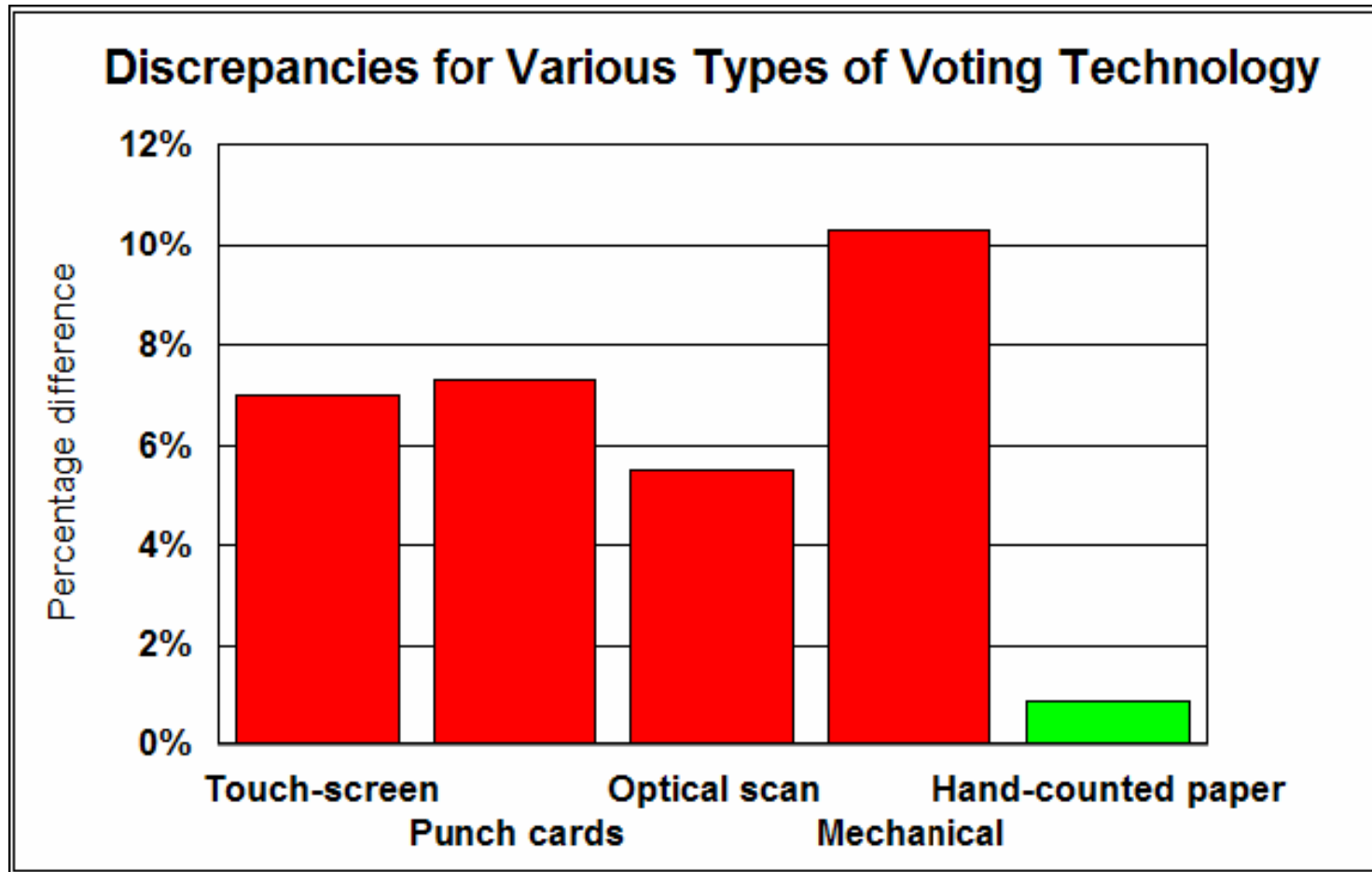
	#	Mean PLD	Median PLD
Non- Swing States	39	5.4	4.7
Swing States	11	8.0	7.8

t-test: p=.04

Among the Swing States:

	#	Mean PLD	Median PLD
Non- Critical Swing States	8	7.6	7.6
Critical Battleground States (Florida, Ohio, Pennsylvania)	3	9.1	8.8

Machine vs. Paper



E/M Dismissal of Paper Correlation

Still lower than machines

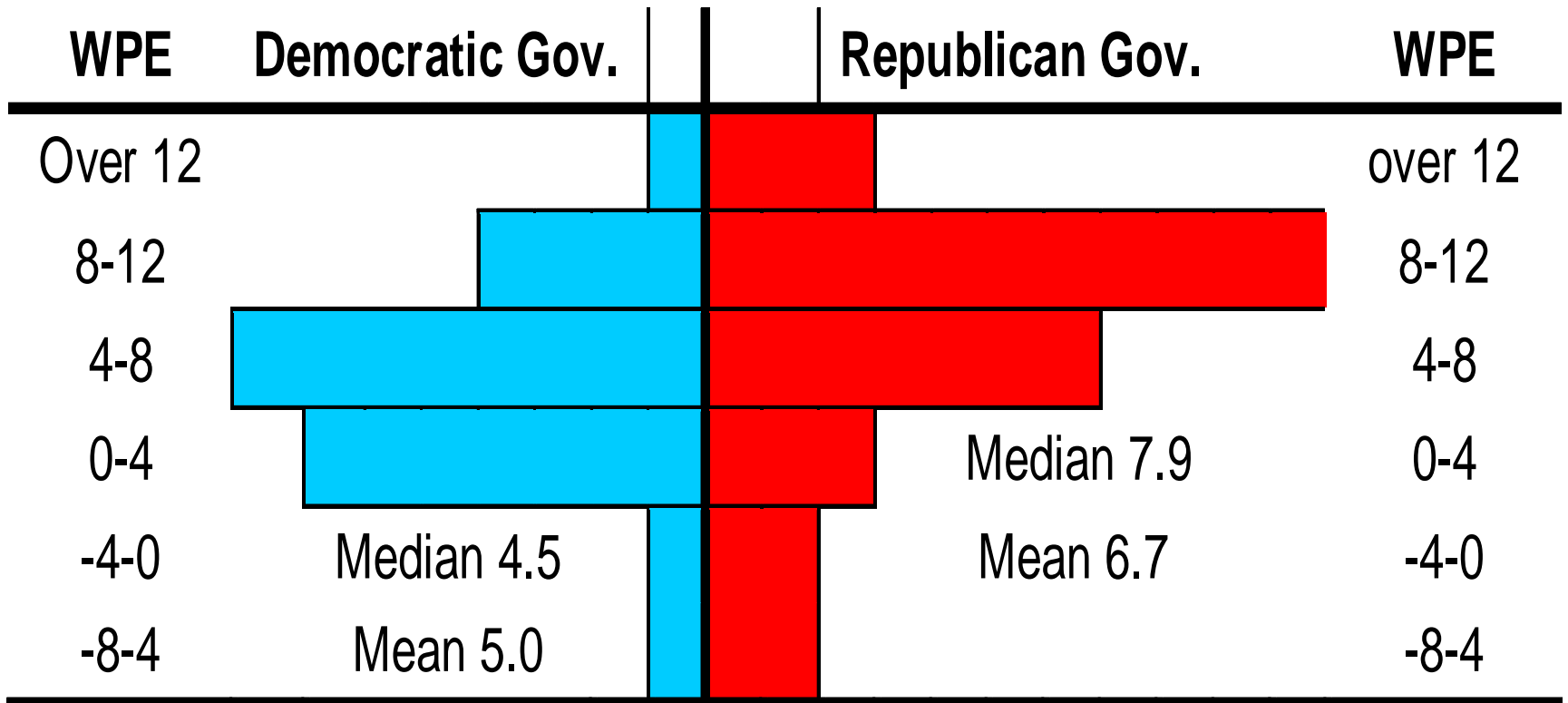
An absurd data partition

Size Of Place	Type of equipment used at polling place	mean WPE	median WPE	mean Abs(WPE)	N
Urban Areas (> 50,000)	Paper Ballot	-6.0	-11.5	15.7	5
	Mechanical Voting Machine	-12.7	-12.5	16.8	92
	Touch Screen	-7.5	-7.6	14.8	272
	Punch Cards	-9.3	-10.0	15.2	108
	Optical Scan	-7.2	-5.8	12.3	350
Rural/Small Town Areas (< 50,000)	Paper Ballot	-1.6	-0.6	10.5	35
	Mechanical Voting Machine	-3.2	-5.4	14.7	26
	Touch Screen	-6.0	-4.8	14.8	88
	Punch Cards	-0.8	-1.7	12.0	50
	Optical Scan	-4.4	-5.0	13.2	223

Rural Area Comparison

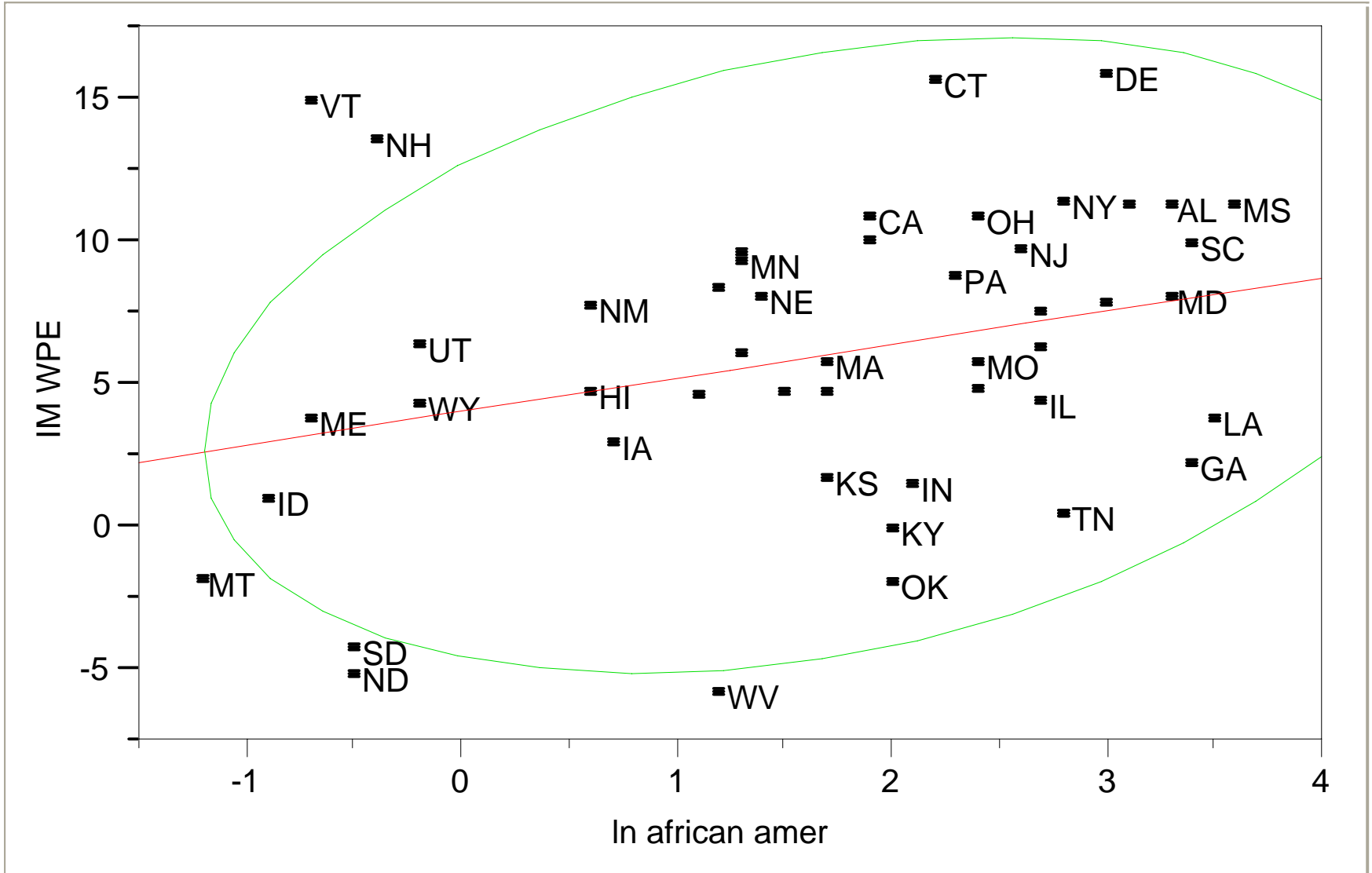
Voting Technology	Mean PLD	N
Paper Ballot	-1.6	35
Machine average	-4.4	1117

Histogram WPE and Gubernatorial Control



t-test: p=.04




African American Percentage of State Population by PLD



Election Administration Problems by State (per million voters)

State Name	MSNBC election day calls /million	State Name	MSNBC election day calls /million	State Name	MSNBC election day calls /million
Alabama	143	Maryland	248	South Carolina	606
Alaska	82	Massachusetts	125	South Dakota	33
Arizona	131	Michigan	267	Tennessee	128
Arkansas	59	Minnesota	68	Texas	127
California	137	Mississippi	52	Utah	53
Colorado	121	Missouri	108	Vermont	73
Connecticut	82	Montana	72	Virginia	155
Delaware	270	Nebraska	58	Washington	117
Washington DC	465	Nevada	215	West Virginia	90
Florida	537	New Hampshire	82	Wisconsin	152
Georgia	146	New Jersey	281	Wyoming	22
Hawaii	45	New Mexico	106		
Idaho	51	New York	302	average	150
Illinois	88	North Carolina	106		
Indiana	96	North Dakota	27	Correlation with PLD	**0.28
Iowa	116	Ohio	195		**p < .01
Kansas	51	Oklahoma	92		
Kentucky	63	Oregon	56		


Presidential vote in 2004 based on 2000 Presidential vote Election Night Data

PRESIDENTIAL VOTE IN 2000		BUSH		KERRY		NADER
	TOTAL	2004	2000	2004	2004	2004
Did Not Vote (17%)	41%	n/a	57%	2%		
Gore (39%)	8%	n/a	91%	1%		
Bush (41%)	90%	n/a	10%	0%		
Other (4%)	17%	n/a	64%	14%		

Source: CNN Screen Shot November 2, 2004 (13,047 Respondents)

Presidential Vote in 2004 based on 2000 Presidential Vote November 3 “Corrected” Data

<http://www.cnn.com/ELECTION/2004/pages/results/states/US/P/00/epolls.0.html>

PRESIDENTIAL VOTE IN 2000		BUSH		KERRY		NADER
TOTAL	2004	2000	2004	2004	2004	2004
Did Not Vote (17%)	45%	n/a	54%	1%		
Gore (37%)	10%	n/a	90%	0%		
Bush (43%)	91%	n/a	9%	0%		
Other (3%)	21%	n/a	71%	3%		

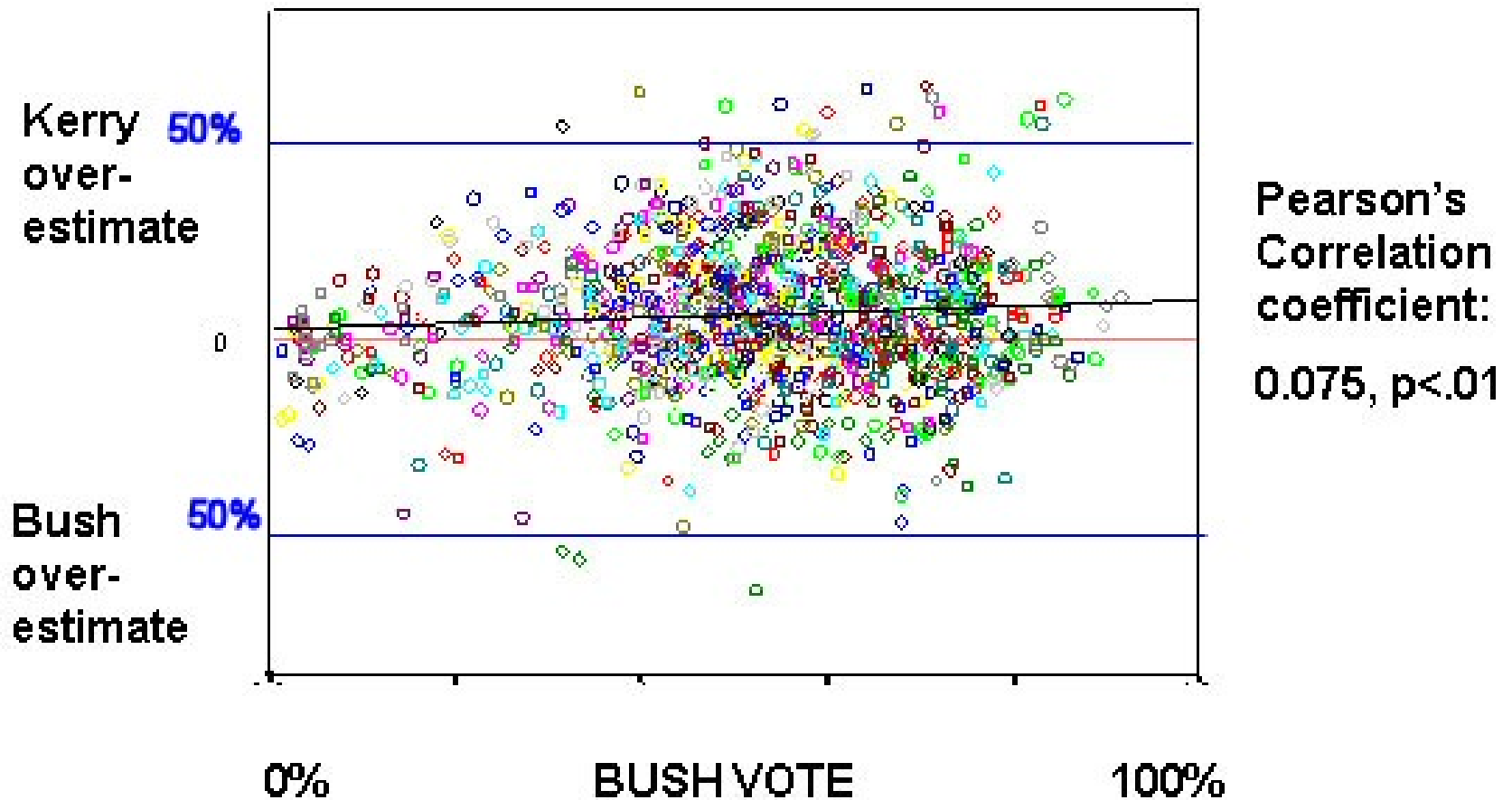
Source: CNN Screen Shot Election night
November 3, 2004 (13,660 Respondents)

Expected Presidential Votes Based on Changes from the 2000 Election

	Bush	Gore/Kerry	Other	Total
2000 official results	50,500,000	48% 51,000,000	48% 4,000,000	4% 105,500,000
2004 official results	62,000,000	51% 59,000,000	48% 1,000,000	1% 122,500,000
New Voters (17% of 2004 electorate did not vote in 2000)				21,000,000
				Repeat Voters: 101,500,000
(0) 96% of 2000 electorate	48,500,000	49,000,000	4,000,000	2%
(1) Bush 2000 Redistributed	43,500,000	90% 5,000,000	10%	48,500,000
(2) Gore Redistributed	4,000,000	8% 45,500,000	91%	500,000 1% 49,000,000
(3) 3rd party redistributed	500,000	17% 2,500,000	64%	500,000 19% 3,500,000
(4) New Voters distributed	8,500,000	41% 12,000,000	57%	500,000 2% 21,000,000
Expected Total	57,000,000	63,500,000	1,500,000	122,100,000
Discrepancy	(5,000,000)	4,500,000	500,000	0

* All vote counts and projections are rounded to the nearest 500,000. Rows do not always total exactly due to rounding.

WPE by Bush Percentage in Precinct



Two Questions to Consider Today

1. How much confidence do we have in the official count – and the explanations about the exit poll discrepancy?

2. How much doubt must be raised about the count – and the explanations about the exit poll discrepancy – before we demand answers?

“Time to Set the Record Straight”

The 2004 exit polls were generally believed to have indicated a victory for John Kerry. Nothing could have been further from the truth. This was compounded by the conspiracy theorists after the election who mistakenly claimed the exit polls validated their claim. There was no evidence in the exit polls to substantiate these claims. On election day the misinformation about the exit polls was spread by inexperienced people trying to make sense of complex statistical data. After the election the academics who yelled election fraud were no better. It is time to set the record straight.

– Warren Mitofsky
abstract for his presentation today