

Dear Greg:

Our offices are about 20 yards and one floor apart but I'm replying to you this way since I have received several emails of your election return analysis from locations as distant as Paris and California.

Below I present an alternative analysis, but first I note that you seem to have initiated what we in the Carter Administration (I also worked in the Bush Administration) called advocacy econometrics. I doubt that your results, nor mine to follow, would be acceptable to an academic journal and yet your results appear to be getting widespread credence at least on the Internet. Let me explain why I think your results should be interpreted with caution.

- Choice of dependent variable: A chart has been circulating showing the number of Buchanan votes in Palm County and the visual (and statistical) indication that it is an outlier. If you view the percentage of vote cast for Buchanan (see below), Palm Beach is not an apparent outlier.
- Forecasting an order statistic: Regression analysis minimizes the sum of squared errors and so it emphasizes large or small outcomes in the sample. Removing the highest vote getting county (first order statistic) from the sample has a large impact, whether for Buchanan or Gore.
- Choice of regressors: You regress the number of Bush (or Gore) votes on the number of Buchanan votes. The Bush votes are endogenous (simultaneously determined) and theory predicts a downwardly biased coefficient from your model that would thus under-predict the number of Buchanan votes.
- Robustness of model: How sensitive are your results to the above issues? Consider another model of votes for Buchanan, that they depend on the number of registered voters for different parties. (That variable is "predetermined" in a statistical sense and avoids the expected downward bias of your model.)
 - Analysis of Buchanan votes without assuming Palm Beach is an outlier: Using the voter registration to predict Buchanan votes and incorporating standard linear and quadratic terms provides a good fit (R-Squared of .97) and significantly increases the expected number of Buchanan votes in Palm Beach to 3,297 with a regression standard error 91 placing the Buchanan total well within 2 standard deviations of the prediction. A standard analysis of residuals would not identify Palm Beach as a problem. (see below.)
 - Analysis of Buchanan votes assuming Palm Beach is an outlier: Your approach drops Palm Beach, runs a regression, and predicts Palm Beach voting. The resulting regression predicts the remaining counties almost as well, but when Palm Beach is forecast, the Buchanan vote appears as a statistical outlier; 38 standard deviations away from the predicted value.
 - Analysis of Gore votes assuming Broward County is an outlier: Following your approach to the county with the highest Buchanan results, what happens when the county (Broward) with the highest number of Gore votes is removed, and then forecast? The resulting linear and quadratic equation has an outstanding fit (R-squared of .995) but when the Gore vote in Broward County is forecast, the

Gore vote also appears as a statistical outlier; 23 standard deviations away from the predicted value (see below).

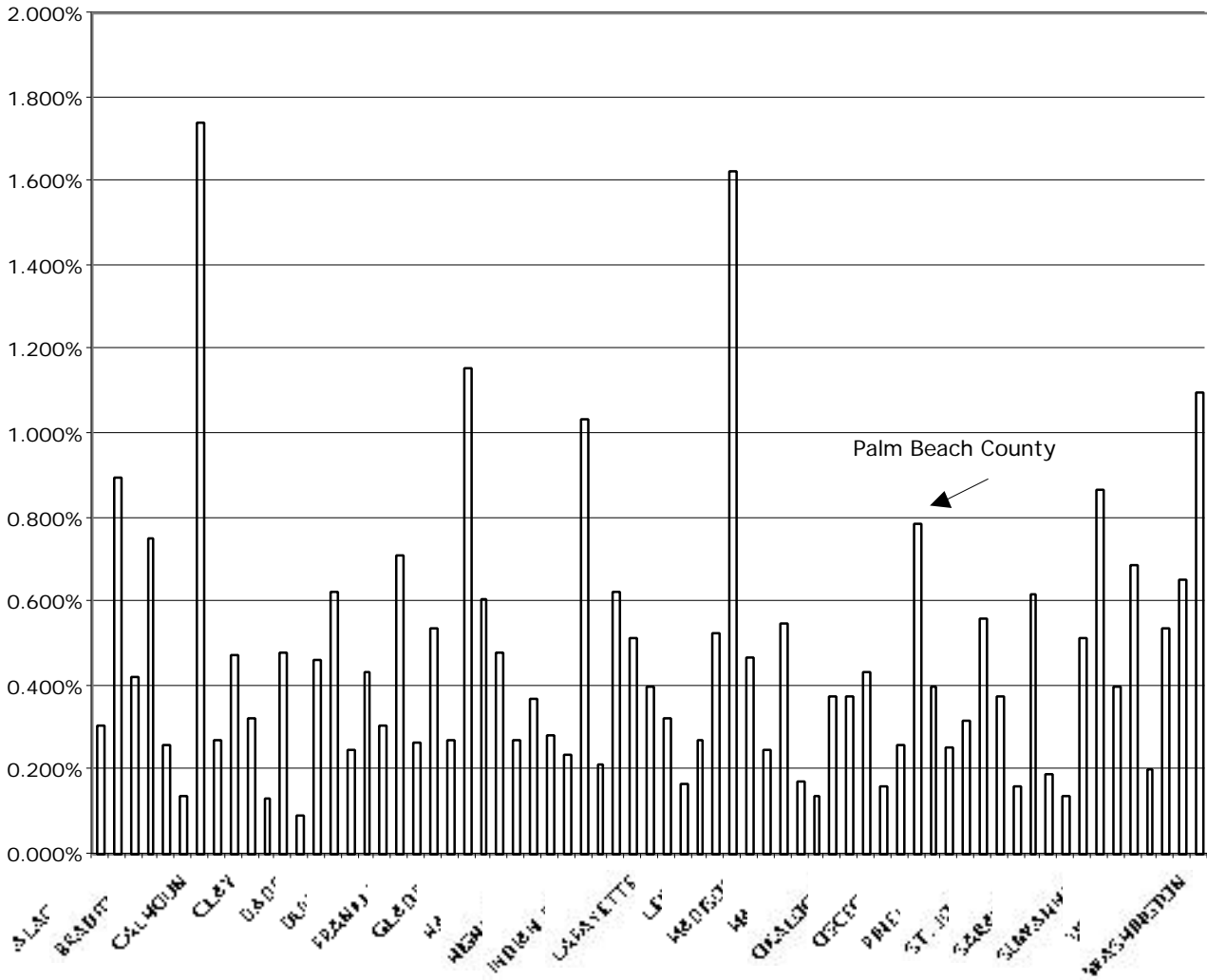
- Numerous other models and statistical issues could be investigated such that I don't think that either of our results meet the rigors of academic publication.

Conclusion: Timeliness is important to policy analysis and you were indeed timely, perhaps too much so. Overconfidence in one's own predictions is a well studied phenomena in your own department. For illustration, I have demonstrated an alternative model that has different implications, namely that the Buchanan vote in Palm Beach County is not an outlier and the number of predicted Buchanan votes is similar to what occurred when the sample is viewed as a whole. When analyses are conducted that drop the county with the highest votes for either Buchanan or Gore, then the vote in the county with the highest number of votes looks like a large statistical outlier compared to the remainder of the sample. I think one should exercise caution in the policy implications of your model.

Scott Farrow
Principal Research Economist
H.J. Heinz III School of Public Policy and Management and
The Department of Engineering and Public Policy
126C Porter Hall
Carnegie Mellon University
Pittsburgh, PA 15213

Sf08@andrew.cmu.edu

Buchanan Percent by County



Percentage Buchanan vote by county

Regression: Y=Buchanan votes; X=linear and squared voter registration; entire sample

SUMMARY OUTPUT

<i>Regression Statistics</i>	
Multiple R	0.98322147
R Square	0.96672445
Adjusted R Square	0.95932989
Standard Error	90.6115112
Observations	67

ANOVA

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	12	12880656.6	1073388.05	130.734439	2.1408E-35
Residual	54	443364.082	8210.44596		
Total	66	13324020.7			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>
Intercept	-43.43079	23.0638624	-1.8830666
Demreg	0.00756504	0.00185744	4.07282794
Repubreg	0.00568627	0.00219256	2.59343666
IndepenRe	-0.0541834	0.02020127	-2.6821787
LibertReg	1.81980468	0.57645481	3.15689043
ReformReg	0.89740838	1.52871845	0.58703313
NoPartyReg	-0.0227108	0.00677838	-3.350479
Demsq	-2.765E-08	6.7217E-09	-4.1139798
Repub2	-4.645E-08	8.8287E-09	-5.2616284
Ind2	1.4424E-05	1.7899E-06	8.05869545
Lib2	-0.0039618	0.00074061	-5.3493961
Reform 2	-0.0046525	0.00608586	-0.764472
NoParty2	3.6109E-07	7.2299E-08	4.99446451

RESIDUAL OUTPUT

	Palm Cty.	Predicted	Residual	Std. Resid.
		3297.24958	109.750417	1.33905339 (see row 5)

<i>Observation</i>	<i>Predicted BUCHANAN</i>	<i>Residuals</i>	<i>Standard Residuals</i>
1	282.119992	-20.119992	-0.2454819
2	38.8464358	34.1535642	0.41670407
3	294.849236	-46.849236	-0.5716026
4	33.0012132	31.9987868	0.39041386
5	312.653967	257.346033	3.1398521
6	826.696336	-37.696336	-0.4599291
7	3.52914902	86.470851	1.05502183
8	180.383564	1.6164358	0.01972196
9	188.851678	81.1483219	0.99008221

10	183.928813	2.0711869	0.02527034
11	183.744496	-61.744496	-0.7533382
12	155.38974	-66.38974	-0.8100143
13	581.498288	-20.498288	-0.2500975
14	19.3654893	16.6345107	0.20295593
15	19.2886403	9.71135966	0.11848729
16	776.475599	-126.4756	-1.5431156
17	499.311605	4.68839476	0.05720262
18	54.2962935	28.7037065	0.35021093
19	0.9879328	32.0120672	0.39057589
20	109.326255	-70.326255	-0.8580433
21	25.1698615	3.83013847	0.04673112
22	-9.6069758	18.6069758	0.22702177
23	25.520623	45.479377	0.55488913
24	1.24808437	22.7519156	0.27759375
25	14.6295742	15.3704258	0.18753296
26	43.8410364	-21.841036	-0.2664802
27	168.31606	73.68394	0.89901007
28	134.548191	-35.548191	-0.4337198
29	546.032639	289.967361	3.53786153
30	26.78597	49.21403	0.60045525
31	181.070836	-76.070836	-0.9281323
32	102.015972	-0.0159717	-0.0001949
33	19.2056154	9.79438459	0.11950026
34	-17.098112	27.098112	0.33062124
35	313.805122	-24.805122	-0.3026447
36	393.70402	-88.70402	-1.0822685
37	476.326326	-194.32633	-2.3709552
38	72.6022923	-5.6022923	-0.068353
39	-12.697682	51.6976821	0.63075803
40	30.0436045	-1.0436045	-0.0127329
41	311.684641	-39.684641	-0.4841882
42	511.969991	51.0300092	0.62261182
43	145.028433	-37.028433	-0.4517801
44	153.979757	-106.97976	-1.3052489
45	97.6672839	-7.6672839	-0.0935477
46	287.038815	-20.038815	-0.2444915
47	57.5754866	-14.575487	-0.177834
48	369.247585	76.7524146	0.9364482
49	232.486621	-87.486621	-1.0674151
50	3297.24958	109.750417	1.33905339 Palm Beach County
51	441.565119	128.434881	1.56702059
52	1191.46392	-181.46392	-2.2140223
53	500.640899	37.3591008	0.45581449
54	151.790495	-4.7904946	-0.0584483
55	240.629327	-11.629327	-0.1418882
56	265.531315	-141.53131	-1.7268088
57	222.638699	88.3613014	1.07808702
58	274.553447	30.4465526	0.37147521
59	301.530975	-107.53097	-1.3119742
60	79.6504378	34.3495622	0.41909543
61	62.6438478	45.3561522	0.55338568
62	34.1703522	-7.1703522	-0.0874847
63	-1.2759632	27.2759632	0.33279118

64	578.35805	-182.35805	-2.2249315
65	45.7120111	0.28798893	0.00351372
66	122.821471	-2.821471	-0.0344245
67	60.6396259	27.3603741	0.33382107

Gore votes: Assuming Broward county is an outlier

<i>Regression Statistics</i>	
Multiple R	0.9976213
R Square	0.99524826
Adjusted R Square	0.9941724
Standard Error	4734.76701
Observations	66

ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	12	2.4886E+11	2.0738E+10	925.067899	4.9437E-57
Residual	53	1188154986	22418018.6		
Total	65	2.5005E+11			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>
Intercept	-2587.7351	1215.14795	-2.1295638
Demreg	0.52650248	0.10015678	5.25678341
Repubreg	-0.1799212	0.15353765	-1.1718378
IndepenRe	3.64933224	1.05743418	3.45112
LibertReg	102.246331	30.1675385	3.38928318
ReformReg	-130.54848	83.3690233	-1.5659112
NoPartyReg	0.57226774	0.37711711	1.5174802
Demsg	-1.451E-06	3.992E-07	-3.6358339
Repub2	6.6191E-07	7.4193E-07	0.89214118
Ind2	-8.494E-05	9.7347E-05	-0.8725788
Lib2	-0.1644114	0.03913006	-4.2016664
Reform 2	0.53531013	0.33221746	1.61132451
NoParty2	9.164E-06	3.9843E-06	2.30003356

RESIDUAL OUTPUT

<i>Observation</i>	<i>Predicted GORE</i>	<i>Residuals</i>	<i>Standard Residuals</i>	Coefficient	Broward County	Vote Imp
				-2587.7351	1	-2587.7351
1	48620.2621	-1320.2621	-0.3088021	0.52650248	456789	240500.5
2	2560.57245	-168.57245	-0.0394282	-0.1799212	266829	-48008.2
3	24582.0182	-5732.0182	-1.3406879	3.64933224	125	456.166
4	3177.90804	-105.90804	-0.0247713	102.246331	700	71572.43
5	82264.2287	15053.7713	3.5209953	-130.54848	332	-43342.0
6	710.514403	1444.4856	0.33785733	0.57226774	162699	93107.38
7	28322.8193	1318.18075	0.30831531	-1.451E-06	2.0866E+11	-302847
8	25133.4103	367.589675	0.08597723	6.6191E-07	7.1198E+10	47126.26
9	18249.1685	-3619.1685	-0.8465038	-8.494E-05	15625	-1.3272
10	29972.3417	-67.341669	-0.0157509	-0.1644114	490000	-80561
11	9487.84234	-2440.8423	-0.5708998	0.53531013	110224	59004.02
12	329863.397	-1161.397	-0.2716444	9.164E-06	2.6471E+10	242581.1
13	3935.28809	-613.28809	-0.1434448		Predicted	276999.1
14	2076.41281	-251.41281	-0.0588041		Actual	3865
15	112654.609	-4974.6088	-1.163534		Std. Dev. From predicted	23.13078

16	52574.5506	-11616.551	-2.7170481
17	6536.14036	7354.85964	1.7202617
18	955.409034	1086.59097	0.25414772
19	8339.42707	1225.57293	0.28665485
20	1223.56044	686.439562	0.16055448
21	519.944695	900.055305	0.21051804
22	1958.89609	430.103906	0.10059897
23	447.675438	1270.32456	0.29712201
24	1423.80246	917.197535	0.21452752
25	3275.37386	-36.373858	-0.0085076
26	25335.383	7308.61701	1.7094458
27	13078.8699	1073.13008	0.25099929
28	167580.073	-999.07283	-0.2336777
29	2248.82139	-94.821395	-0.0221782
30	14231.6304	5537.36956	1.29516065
31	6974.28261	-106.28261	-0.0248589
32	1921.33198	1116.66802	0.26118258
33	-701.96617	1489.96617	0.34849499
34	34932.9051	1622.0949	0.37939918
35	83521.495	-9961.495	-2.3299395
36	54061.1389	7363.86114	1.7223671
37	4831.70362	571.296375	0.13362312
38	-418.88886	1429.88886	0.33444323
39	2467.61829	543.381714	0.12709403
40	46631.858	2537.14199	0.59342372
41	48273.2857	-3625.2857	-0.8479346
42	20057.2898	6561.71022	1.53474836
43	18317.9119	-1834.9119	-0.4291759
44	8793.53465	-1841.5346	-0.4307249
45	24137.7877	-7213.7877	-1.6872658
46	1744.8042	2843.1958	0.66500805
47	135479.927	4635.07319	1.08411843
48	37340.3397	-9163.3397	-2.1432554
49	264905.388	4039.61244	0.9448434
50	70249.7215	-699.72151	-0.1636611
51	203018.19	-3358.1898	-0.7854624
52	74402.8681	574.131896	0.13428633
53	13748.9022	-1657.9022	-0.3877743
54	25084.7401	-5602.7401	-1.3104505
55	44767.7284	-3208.7284	-0.7505041
56	16936.6081	-4141.6081	-0.9686996
57	65289.1395	7564.86047	1.76937975
58	62058.447	-3170.447	-0.7415503
59	7465.48178	2168.51822	0.50720463
60	7528.53903	-3444.539	-0.805659
61	3019.84621	-372.84621	-0.0872067
62	449.358895	949.641105	0.22211589
63	93212.4061	3850.59392	0.90063299
64	3206.30162	628.698384	0.14704914
65	8159.73595	-2522.7359	-0.5900542
66	4132.88881	-1336.8888	-0.312691
