

Reanalysis of Hout et al

Kieran Healy*
University of Arizona

November 20, 2004

Hout et al. (2004) present an analysis of county-level voting patterns in Florida for the Presidential election of 2004. The authors argue that President Bush received a greater than expected share of the vote in counties using electronic voting machines, controlling for various demographic characteristics of the counties as well as the proportion of votes cast for the Republican Presidential candidate in 1996 and 2000:

As baseline support for Bush increases in Florida counties, the change in percent voting for Bush from 2000 to 2004 increases, but at a decreasing rate. Electronic voting has a main, positive effect on the dependent variable. Furthermore, there is an interaction effect between baseline support for Bush and electronic voting, and between baseline support for Bush squared and electronic voting. Support for Dole in 1996, county size, median income, and Hispanic population had no significant effect net of the other effects. Essentially, net of other effects, electronic voting had the greatest positive effect on changin percent voting for Bush from 2000 to 2004 in democratic counties. Summing these effects for the fifteen counties with electronic voting yields the total estimated excess votes in favor of Bush associated with Electronic Voting; this figure is 130,733.

Following a discussion by Gelman (2004), Figure 1 plots the percentage swing to Bush in the Florida counties by the percent vote for Bush in 2000.

We can see that Broward and Palm Beach counties (which have very large populations and lean strongly Democratic) swung more toward Bush

*Social Sciences 404, University of Arizona, Tucson AZ 85721. Email at kjhealy@arizona.edu.

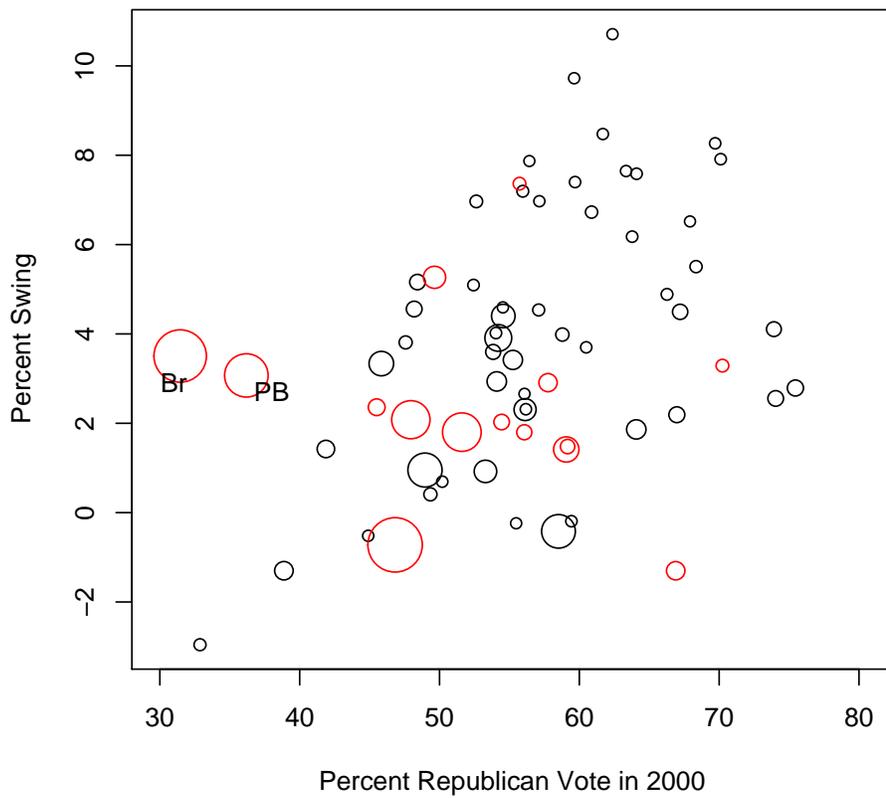


Figure 1: Percentage Swing to Bush in 2004. Counties with Electronic voting are shown in red. Counties are plotted in proportion to the number of votes cast. Palm Beach County is marked “PB” and Broward County “Br”.

than was typical for counties where Republicans won less than 47 or 48 percent of the vote in 2000. It turns out that these two counties are driving the findings in Gelman (2004). Their model is as follows:

```
> out.hout <- lm(b.change ~ b00pc + b00pc.sq + size + etouch +
+   b00pc.e + b00pcsq.e + d96pc + v.change + income + hispanic,
+   data = data.fl)
> summary(out.hout)
```

Call:

```
lm(formula = b.change ~ b00pc + b00pc.sq + size + etouch + b00pc.e +
    b00pcsq.e + d96pc + v.change + income + hispanic, data = data.fl)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.05954	-0.01178	0.00082	0.01095	0.04910

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-2.13e-01	9.42e-02	-2.26	0.0277 *
b00pc	1.03e+00	3.22e-01	3.19	0.0023 **
b00pc.sq	-6.64e-01	2.81e-01	-2.36	0.0215 *
size	-3.93e-08	6.69e-08	-0.59	0.5593
etouch	4.17e-01	1.49e-01	2.79	0.0073 **
b00pc.e	-1.28e+00	5.55e-01	-2.31	0.0245 *
b00pcsq.e	9.38e-01	5.19e-01	1.81	0.0759 .
d96pc	-1.52e-01	1.17e-01	-1.30	0.1990
v.change	-2.67e-11	3.00e-07	-8.9e-05	0.9999
income	-8.17e-07	7.56e-07	-1.08	0.2841
hispanic	-5.26e-02	3.07e-02	-1.71	0.0930 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.0214 on 56 degrees of freedom

Multiple R-Squared: 0.537, Adjusted R-squared: 0.455

F-statistic: 6.51 on 10 and 56 DF, p-value: 1.43e-06

The variable of interest is `etouch`, which picks out counties with electronic voting. If we create a dummy variable marking Broward and Palm Beach Counties and re-run the model, we get the following results:

```

> out.pbbr <- update(out.hout, . ~ . + pb.br)
> summary(out.pbbr)

Call:
lm(formula = b.change ~ b00pc + b00pc.sq + size + etouch + b00pc.e +
    b00pcsq.e + d96pc + v.change + income + hispanic + pb.br,
    data = data.fl)

Residuals:
    Min       1Q   Median       3Q      Max
-0.059500 -0.010019  0.000652  0.011719  0.049047

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -2.13e-01  9.49e-02  -2.24  0.0289 *
b00pc        1.03e+00  3.25e-01   3.17  0.0025 **
b00pc.sq    -6.62e-01  2.83e-01  -2.34  0.0230 *
size        -2.88e-08  7.21e-08  -0.40  0.6908
etouch      2.98e-01  3.26e-01   0.92  0.3638
b00pc.e    -8.82e-01  1.13e+00  -0.78  0.4373
b00pcsq.e   6.02e-01  9.71e-01   0.62  0.5377
d96pc      -1.58e-01  1.19e-01  -1.33  0.1881
v.change    -4.41e-08  3.21e-07  -0.14  0.8912
income      -7.89e-07  7.64e-07  -1.03  0.3064
hispanic    -5.21e-02  3.10e-02  -1.68  0.0988 .
pb.br       2.14e-02  5.23e-02   0.41  0.6831
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.0215 on 55 degrees of freedom
Multiple R-Squared: 0.539,    Adjusted R-squared: 0.447
F-statistic: 5.84 on 11 and 55 DF,  p-value: 3.55e-06

```

As we can see, putting in a dummy for Palm Beach and Broward Counties makes the significant effect of `etouch` go away. Now the only variables significant at conventional levels are the ones measuring the percentage voting for Bush in 2000. (Note that there's also a hint of an effect for `hispanic`, as befits their ambiguous role in deciding the election.)

So, all of the e-voting action is explained by two counties. The question is what's happening in those counties. Andrew Gelman again:

One possibility, as suggested by Hout et al., is cheating, possibly set up ahead of time (e.g., by loading extra votes into the machines before the election or by setting it up to switch or not count some votes) — but an obvious alternative explanation is that, for various reasons, 3% more people in those counties preferred Bush in 2004, compared to 2000. As can be seen in the graphs above for 2000, 1996, and 1992, such a swing would be unusual (at least compared to recent history), but that doesn't mean it couldn't happen! It would make sense to look further at Broward and Palm Beach counties, where swings happened which look unexpected compared to the other counties and compared to 2000, 1996, and 1992. But lots of unexpected things happen in elections, so we shouldn't jump to the conclusion that e-voting is related to these particular surprises.

In other words, if there is cheating it's not centralized cheating where all the e-voting machines mess up in the same way. If you believe that the machines were rigged, focus on the ones in Palm Beach and Broward county. But it seems more likely that these results show the Republican Party Machine was really, really well-organized in Palm Beach and Broward, and they were able to mobilize their vote better than the Democrats. The general swing toward Bush in Florida seems consistent with this story.

References

- Gelman, Andrew. 2004. "Vote swings in Florida in counties with and without e-voting." http://www.stat.columbia.edu/~cook/movabletype/archives/2004/11/vote_sw%ings_in.html.
- Hout, Michael, Laura Mangels, Jennifer Carlson, and Rachel Best. 2004. "The Effect of Electronic Voting Machines on Change in Support for Bush in the 2004 Florida Elections." Working Paper, Department of Sociology, University of California at Berkeley.