Replication of Green & Vasudevan

Zenobia Chan, Alicia Cooperman, & Lauren Young

Columbia University

October 2015

Overview

- Theory
- Design
- Replication of main results
- Robustness to other coding of vote buying
- Heterogeneous effects

Theory

- Two-party system: (1) Vote-buying *vs.* (2) Non-vote-buying
- Representative agent model:

$$U(v, l, x) = \underbrace{\alpha \ln(v)}_{\text{reciprocity}} + \underbrace{\beta \ln(l)}_{\text{leisure}} + \underbrace{\gamma \ln(x)}_{\text{public good}}$$

- Public good:

$$x(s, t) = \overbrace{t^h}^{\text{turnout}} \underbrace{s^{h_1}(1-s)^{h_2}}_{\text{vote-shares}} \text{ where } h \text{ is honesty}$$

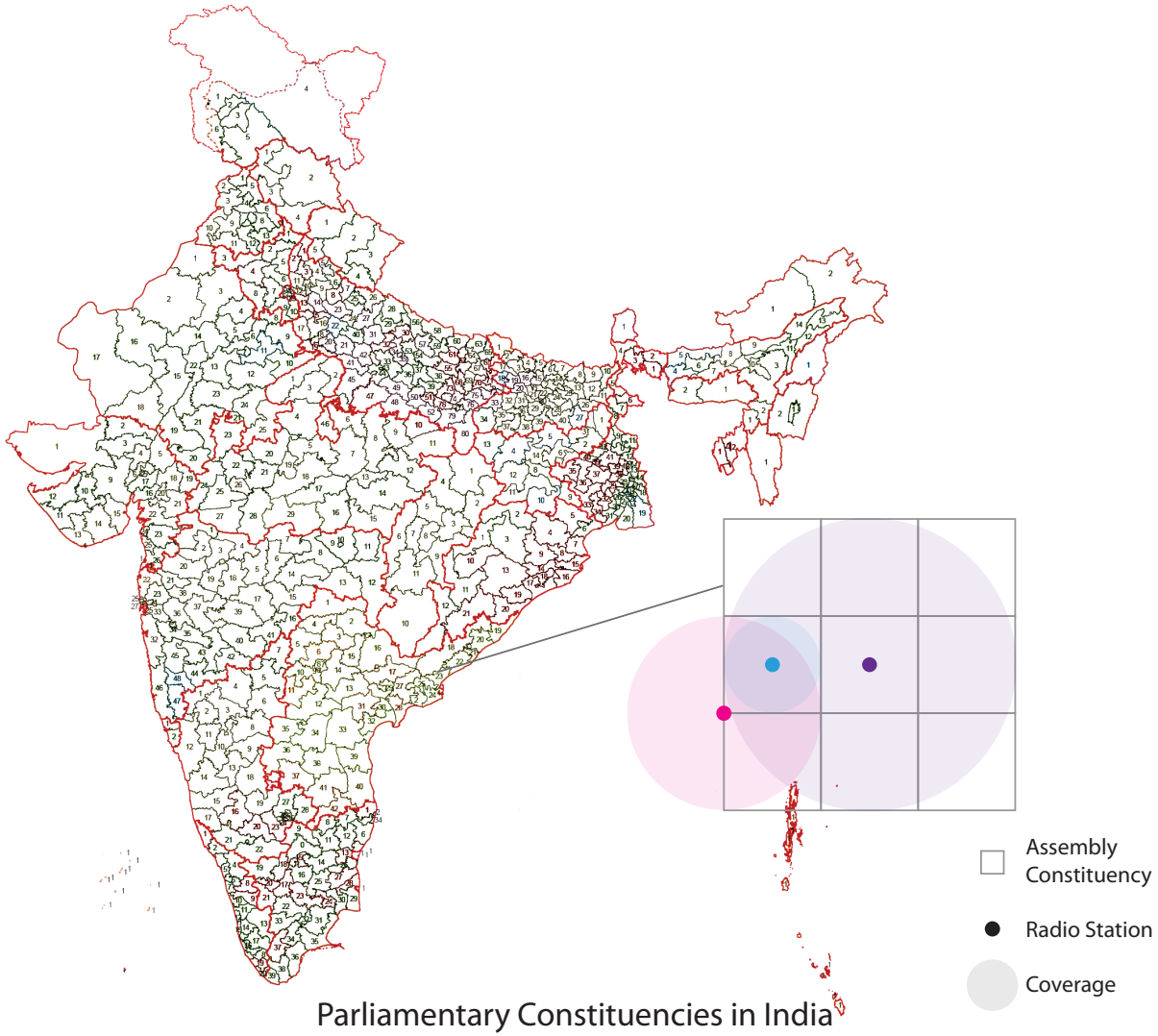
- The message affects voter behavior by:
  1. Decreasing voter reciprocity towards the vote-buying party
  2. Changing voter expectations of honesty of the two parties

Design

- 2014 Indian General Election
- Radio campaigns immediately before polling
  - 60-sec dramatized vignette
  - With information on (1) the nature of vote-buying and (2) its hidden social costs
  - Translated into Hindi and four regional languages

Design

- Randomization occurs at radio station-level



Replication process

- Both Stata and Matlab codes were used originally

covbalance_reg.m

```
cd('C:\Users\srini\Dropbox\Green and Vasudevan (2015) Data\4. Analysis\Matlab Data');
clear;
clc;
data = csvread('covbalance_data.csv',1,0);
N= 665;

electors = data(:,1);
turnout = data(:,2);
margin = data(:,3);
num_cand = data(:,4);
num_ind = data(:,5);

num_vb2 = data(:,7);
num_vb3 = data(:,8);

ind_share = data(:,9);

vb2_share = data(:,11);
vb3_share = data(:,12);

lit = data(:,13);
rural = data(:,14);
scst = data(:,15);

weight = sqrt(data(:,17));
treat0 = weight.*data(:,16);
treat = [weight treat0];

covariates = [weight.*electors weight.*turnout weight.*margin weight.*num_cand weight.*
    num_ind weight.*num_vb2 weight.*num_vb3 weight.*ind_share weight.*vb2_share weight.*
    vb3_share weight.*lit weight.*rural weight.*scst];

station1 = data(:,18);
station2 = data(:,19);
station3 = data(:,20);

dep_matrix = zeros(N,N);

for i = 1:N
        for j = 1:N
            if station1(i) == station1(j)
                dep_matrix(i,j) = 1;
            end;
            if station1(i) == station2(j)
                dep_matrix(i,j) = 1;
            end;
            if station1(i) == station3(j)
                dep_matrix(i,j) = 1;
            end;
            if station2(i) == station1(j) && station2(i) ~= -9
                dep_matrix(i,j) = 1;
            end;
            if station2(i) == station2(j) && station2(i) ~= -9
                dep_matrix(i,j) = 1;
            end;
            if station2(i) == station3(j) && station2(i) ~= -9
```

covbalance_data.do

```
/*
Covariate balance data
*/

set more off
cd "C:\Users\Srini\Dropbox\Green and Vasudevan (2015) Data"

*****************************************2009 Variables*********************
*Number votebuyers, number parties, number candidates, votebuyer voteshare, ir
use "Electoral Data/results_10states2009_clean.dta", clear
drop if ac_num<0 //dropping postal ballot

*Candidates
bys state_name ac_num: egen num_cand_2009 = count(ac_num)

*Independents
gen ind = cand_votes if cand_party == "IND"
bys state_name ac_num: egen indvoteshare_2009 = sum(ind)
replace indvoteshare_2009 = 100*indvoteshare_2009/total_ac_votes
bys state_name ac_num: egen num_ind_2009 = count(ind)

*Margin
egen rank = rank(-cand_votes), by(state_name ac_num) unique

egen winner_votes = total(cand_votes/(rank==1)), by(state_name ac_num)
egen runnerup_votes = total(cand_votes/(rank==2)), by(state_name ac_num)
gen margin_2009 = 100*(winner_votes-runnerup_votes)/total_ac_votes

drop rank winner_votes runnerup_votes

bys state_name ac_num: drop if _n >1

merge 1:1 state_name ac_num using "Sample Data/AC_expt_sample.dta", nogen forc
merge 1:1 state_name ac_num using "Temp/voteshare_data.dta", keep(matched) kee
merge 1:1 state_name ac_num using "Temp/turnout_data.dta", keepusing(turnoutra
merge 1:1 state_name ac_num using "Electoral Data/electors_2009.dta", keepusir

*****************************************************************************
gen state_election = (inlist(state_name,"Andhra Pradesh","Orissa"))
egen cluster = group(station_id1 station_id2 station_id3)
tab num_eligible, gen(num_stations)
tab poll_date, gen(poll_date)

replace illit_pc = 100-illit_pc
rename illit_pc lit_pc

save "Temp/covbalance_data.dta", replace
```

Suggestions for replication package

- Code written in Matlab + Stata
  - Randomization - Stata
  - Data Building - Stata
  - Regressions - Matlab
  - Standard Errors - Matlab
  - Randomization Inference Simulations - Stata
  - p-values - Matlab
- Possible to do everything in R
- Include a roadmap (master R file, markdown, etc)

Main results from the paper

Table 6: Average Treatment Effect (ATE) of receiving radio ads on vote-share of vote-buying parties and on the voter turnout rate

	Vote-share of vote-buying parties (%)						Turnout rate (%)	
	Specification 1[5]		Specification 2		Specification 3			
	IPW	FE	IPW	FE	IPW	FE	IPW	FE
ATE[1]	-5.86	-6.04	-7.68	-7.73	-3.68	-3.41	-0.49	-0.61
SE[2]	3.97	4.08	3.92	4.18	1.92	2.04	0.96	0.99
p-value[3]	0.08	0.08	0.00	0.00	0.02	0.03	0.64	0.57
R-squared	0.44	0.43	0.38	0.28	0.51	0.33	0.80	0.76
Mean[4] (Control)	67.23		90.85		91.73		68.45	
N	628		665		665		665	
Control	315		324		324		324	
Treatment	313		341		341		341	

All specifications have the lagged outcome variable as covariate.
[1]IPW are inverse probability weighted and FE are fixed effects regression estimates respectively.
[2]Standard errors are robust to heteroskedasticity and known cross-sectional dependence of the error term.
[3]p-values obtained from randomization inference with 10,000 iterations.
[4]Control Means are inverse probability weighted.
[5]Responses identifying vote-buying parties for 37 ACs are missing.

Main results from the paper

	Spec 1		Spec 2		Spec 3	
	IPW	FE	IPW	FE	IPW	FE
ATE[1]	-5.86	-6.04	-7.68	-7.73	-3.68	-3.41
SE[2]	3.97	4.08	3.92	4.18	1.92	2.04
p-value (Barrios)[3]	0.07	0.07	0.03	0.03	0.03	0.05
p-value (RI)[4]	0.08	0.08	0.00	0.00	0.02	0.03
R^2	0.44	0.43	0.38	0.28	0.51	0.33
Mean (Control)	67.24		90.82		91.68	
N	628		665		665	
control	315		324		324	
treat	313		341		341	

[1] ATEs are estimated using OLS weighted by the inverse propensity of receiving treatment (IPW) or with dummies for the probability of receiving treatment as fixed effects (FE). All specifications include the pre-treatment measure of the outcome as a covariate.

[2] Standard errors that are robust to heteroskedasticity and known cross-sectional dependence of the error term calculated using the Barrios et al (2012) method.

[3] p-values obtained from the Barrios et al (2012) estimates of uncertainty.

[4] p-values obtained from randomization inference with 10,000 simulations.

Correcting standard errors

Imagine a scenario of 3 clusters with 2 units each.

Table : Constant error variance

	e_{11}	e_{12}	e_{21}	e_{22}	e_{31}	e_{32}
e_{11}	σ^2	0	0	0	0	0
e_{12}	0	σ^2	0	0	0	0
e_{21}	0	0	σ^2	0	0	0
e_{22}	0	0	0	σ^2	0	0
e_{31}	0	0	0	0	σ^2	0
e_{32}	0	0	0	0	0	σ^2

Table : Not-constant error Σ

	e_{11}	e_{12}	e_{21}	e_{22}	e_{31}	e_{32}
e_{11}	σ^2_{11}	0	0	0	0	0
e_{12}	0	σ^2_{12}	0	0	0	0
e_{21}	0	0	σ^2_{21}	0	0	0
e_{22}	0	0	0	σ^2_{22}	0	0
e_{31}	0	0	0	0	σ^2_{31}	0
e_{32}	0	0	0	0	0	σ^2_{32}

$\text{Var}(\hat{\beta}) = (X'X)^{-1}(X'\Sigma X)(X'X)^{-1}$

Huber-White "Robust" SEs estimate $\hat{\Sigma}$ where σ_i^2 is $\hat{u}_i^2$

But, still assumes no clustered or spatial correlation

Correcting standard errors

Imagine a scenario of 3 clusters with 2 units each.
Cluster-robust "block diagonal"

Table : Cluster robust

	e_{11}	e_{12}	e_{21}	e_{22}	e_{31}	e_{32}
e_{11}	σ_{11}^2	$\sigma_{11}\sigma_{12}$	0	0	0	0
e_{12}	$\sigma_{12}\sigma_{11}$	σ_{12}^2	0	0	0	0
e_{21}	0	0	σ_{21}^2	$\sigma_{21}\sigma_{22}$	0	0
e_{22}	0	0	$\sigma_{22}\sigma_{21}$	σ_{22}^2	0	0
e_{31}	0	0	0	0	σ_{31}^2	$\sigma_{31}\sigma_{32}$
e_{32}	0	0	0	0	$\sigma_{32}\sigma_{31}$	σ_{32}^2

Correcting standard errors

Imagine a scenario of 3 clusters with 2 units each,
but Station 1 covers 11, 12, 21; Station 2 covers cluster 2;
Station 3 covers cluster 3.

Table : Barrios Dependency Matrix

	e_{11}	e_{12}	e_{21}	e_{22}	e_{31}	e_{32}
e_{11}	1	1	1	0	0	0
e_{12}	1	1	1	0	0	0
e_{21}	1	1	1	1	0	0
e_{22}	0	0	1	1	0	0
e_{31}	0	0	0	0	1	1
e_{32}	0	0	0	0	1	1

Multiply this matrix element-by-element with $\hat{u}\hat{u}'$

Correcting standard errors

Imagine a scenario of 3 clusters with 2 units each,
but Station 1 covers 11, 12, 21; Station 2 covers cluster 2;
Station 3 covers cluster 3.

Table : Barrios $\hat{\Sigma}$

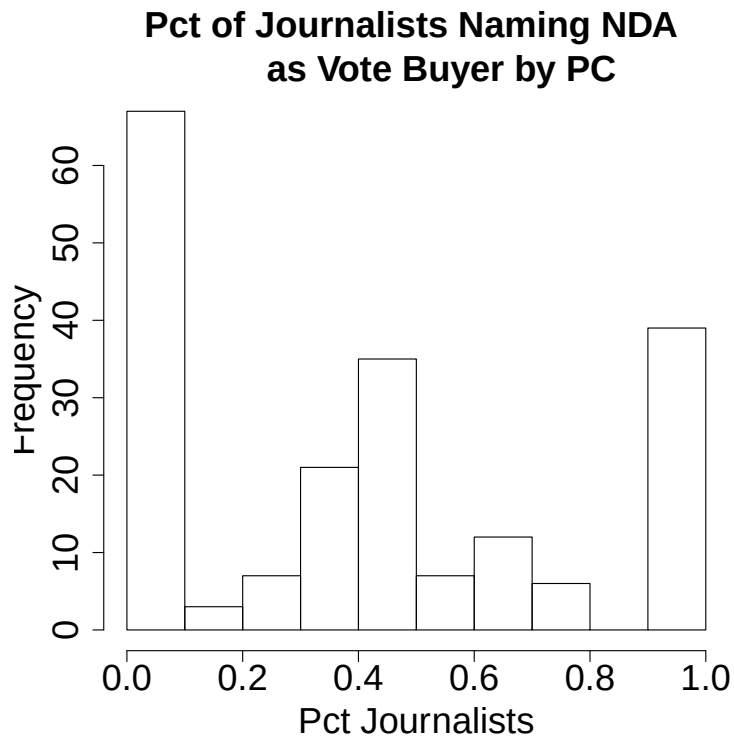
	e_{11}	e_{12}	e_{21}	e_{22}	e_{31}	e_{32}
e_{11}	σ_{11}^2	$\sigma_{11}\sigma_{12}$	$\sigma_{11}\sigma_{21}$	0	0	0
e_{12}	$\sigma_{12}\sigma_{11}$	σ_{12}^2	$\sigma_{12}\sigma_{21}$	0	0	0
e_{21}	$\sigma_{21}\sigma_{11}$	$\sigma_{21}\sigma_{12}$	σ_{21}^2	$\sigma_{21}\sigma_{22}$	0	0
e_{22}	0	0	$\sigma_{22}\sigma_{21}$	σ_{22}^2	0	0
e_{31}	0	0	0	0	σ_{31}^2	$\sigma_{31}\sigma_{32}$
e_{32}	0	0	0	0	$\sigma_{32}\sigma_{31}$	σ_{32}^2

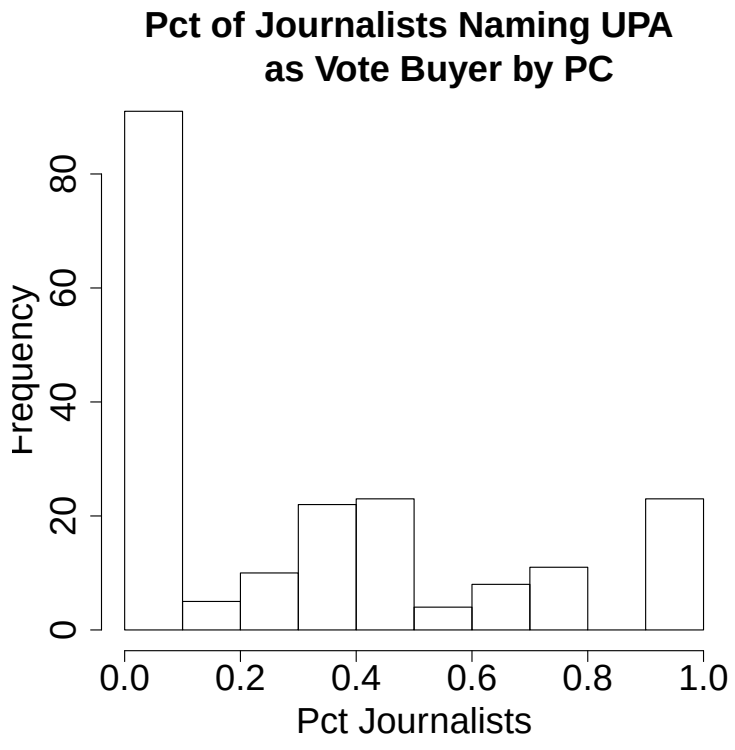
$$\text{Var}(\hat{\beta}) = (X'X)^{-1}(X'\hat{\Sigma}X)(X'X)^{-1}$$

What does it mean to be a vote buying party?

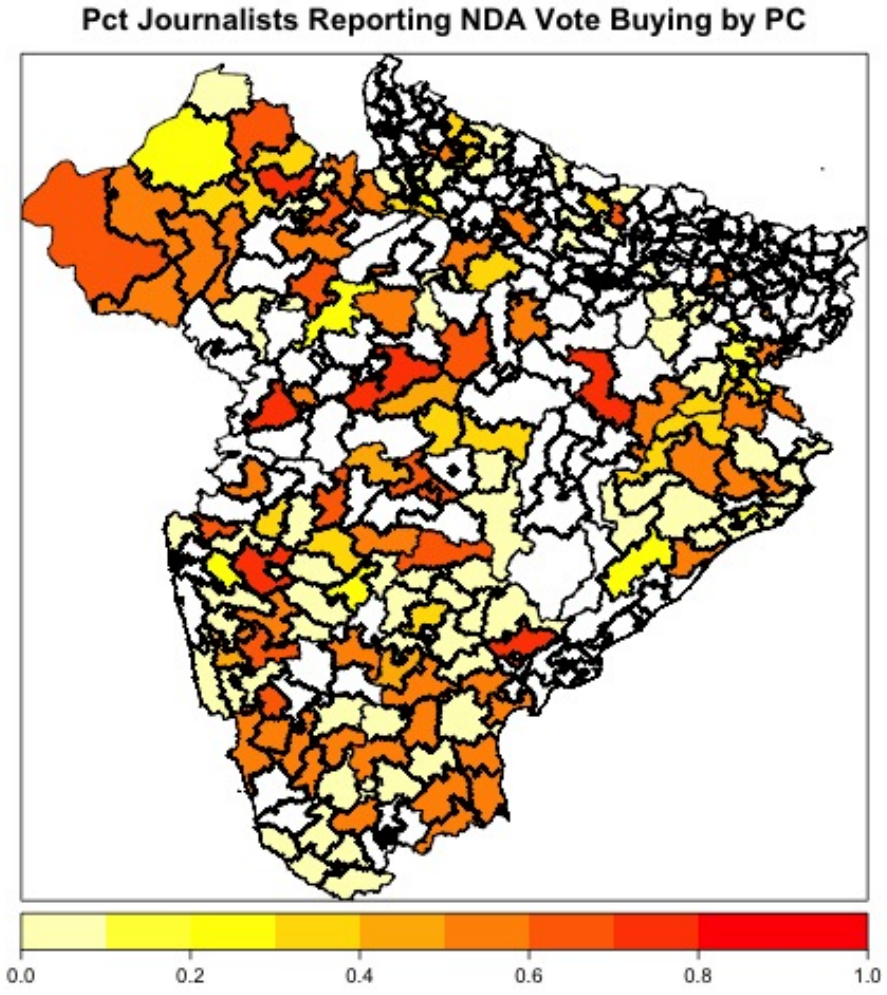
- Very innovative measure of illicit electoral technique
  - Cost-effective
  - Draws on local expertise
  - Covers comprehensive area
- What is the data generating process?
  - Journalistic ethics to tell the truth
  - Ideological biases
  - Pay more attention to major parties
- How to think about uncertainty with journalist data?
  - Under-identification (uninformedness, bias towards parties)
  - Over-identification (bias against parties)
  - Random noise

What does it mean to be a vote buying party?

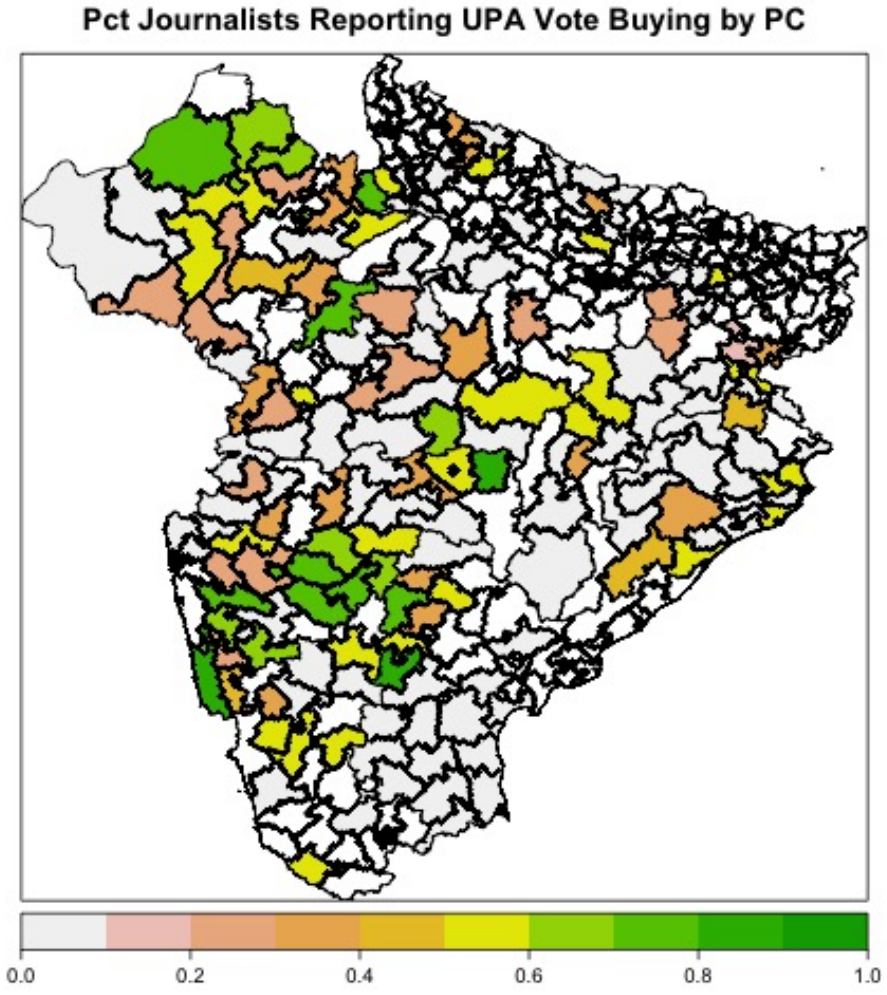




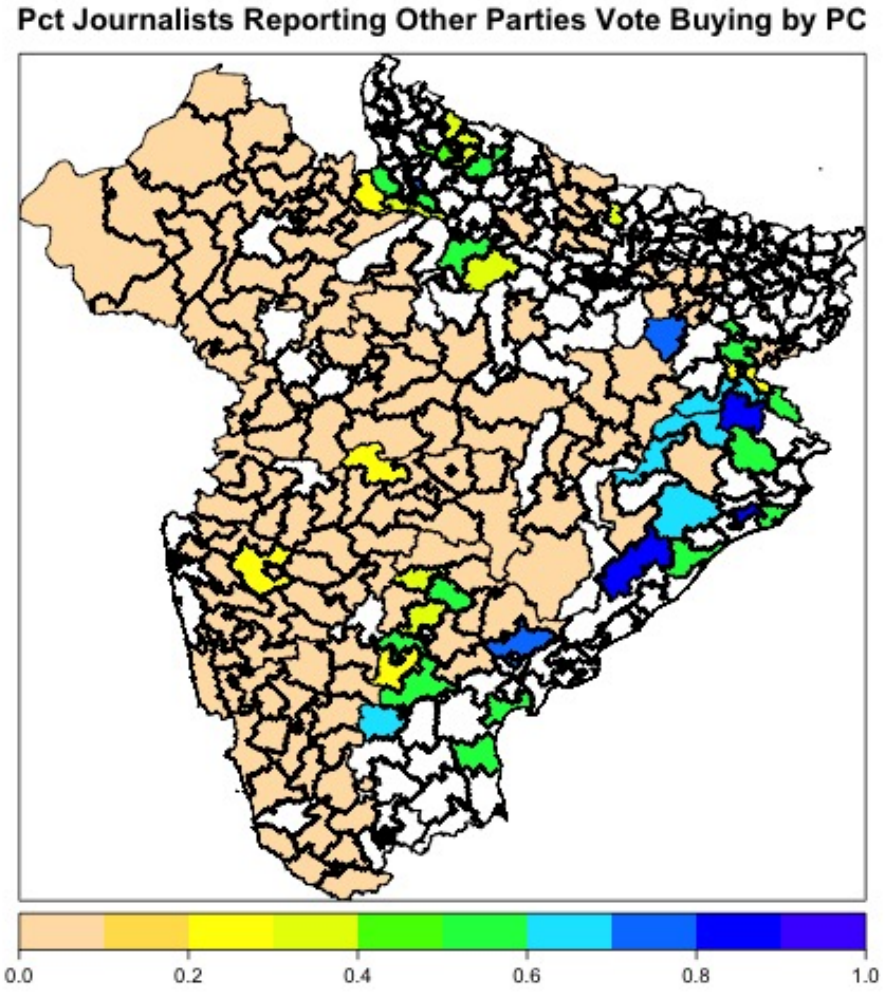
What does it mean to be a vote buying party?



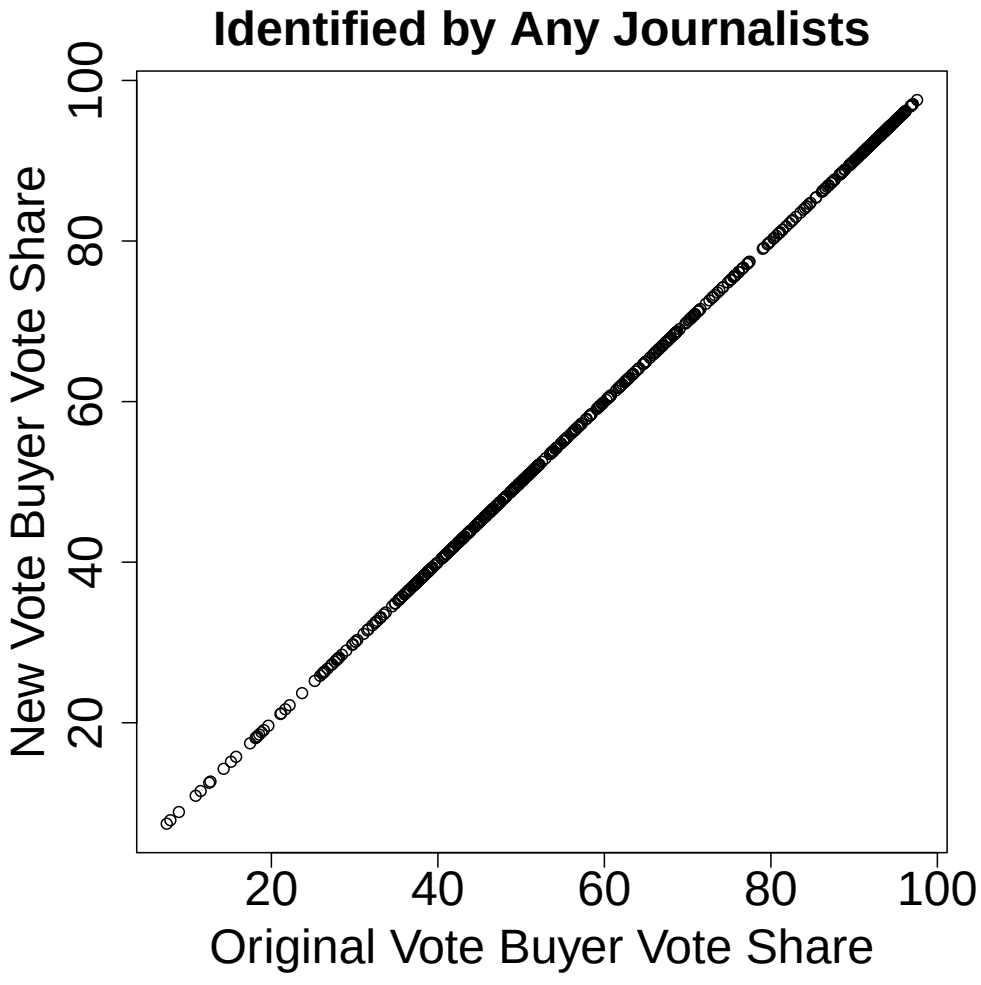
What does it mean to be a vote buying party?



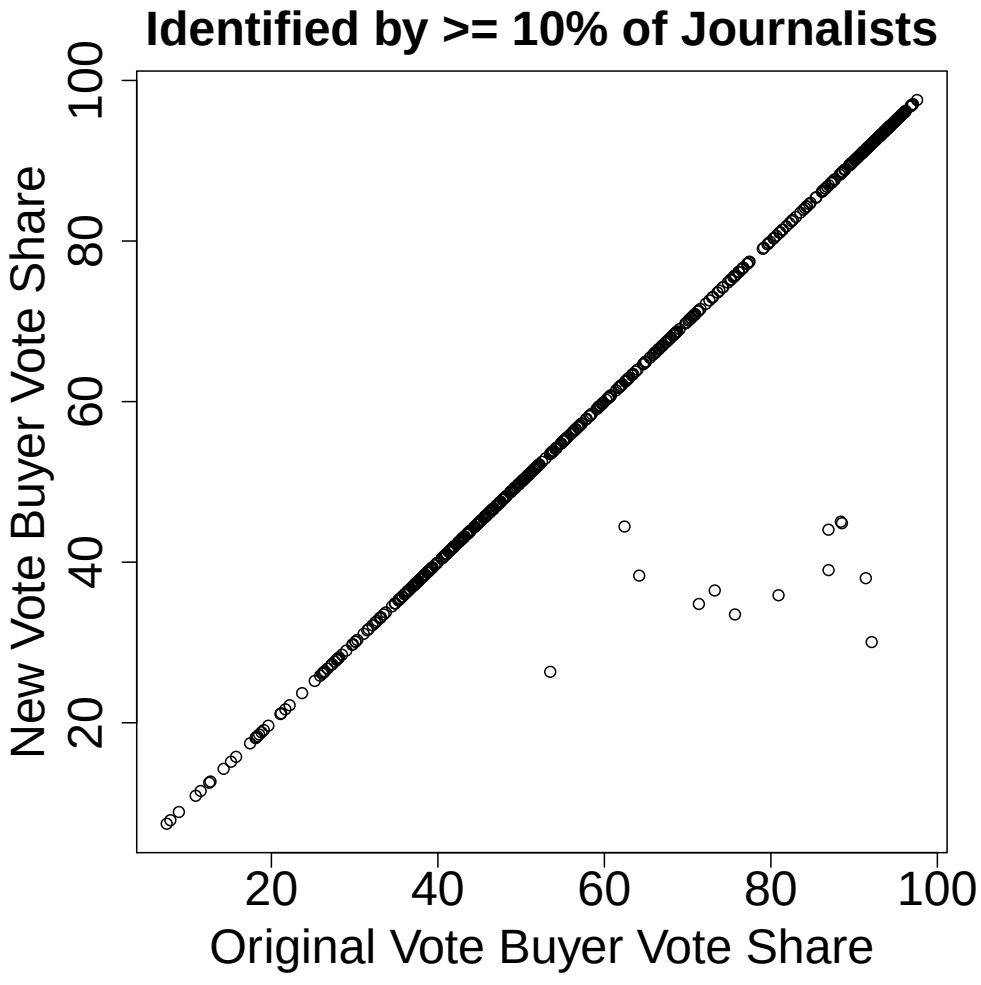
What does it mean to be a vote buying party?



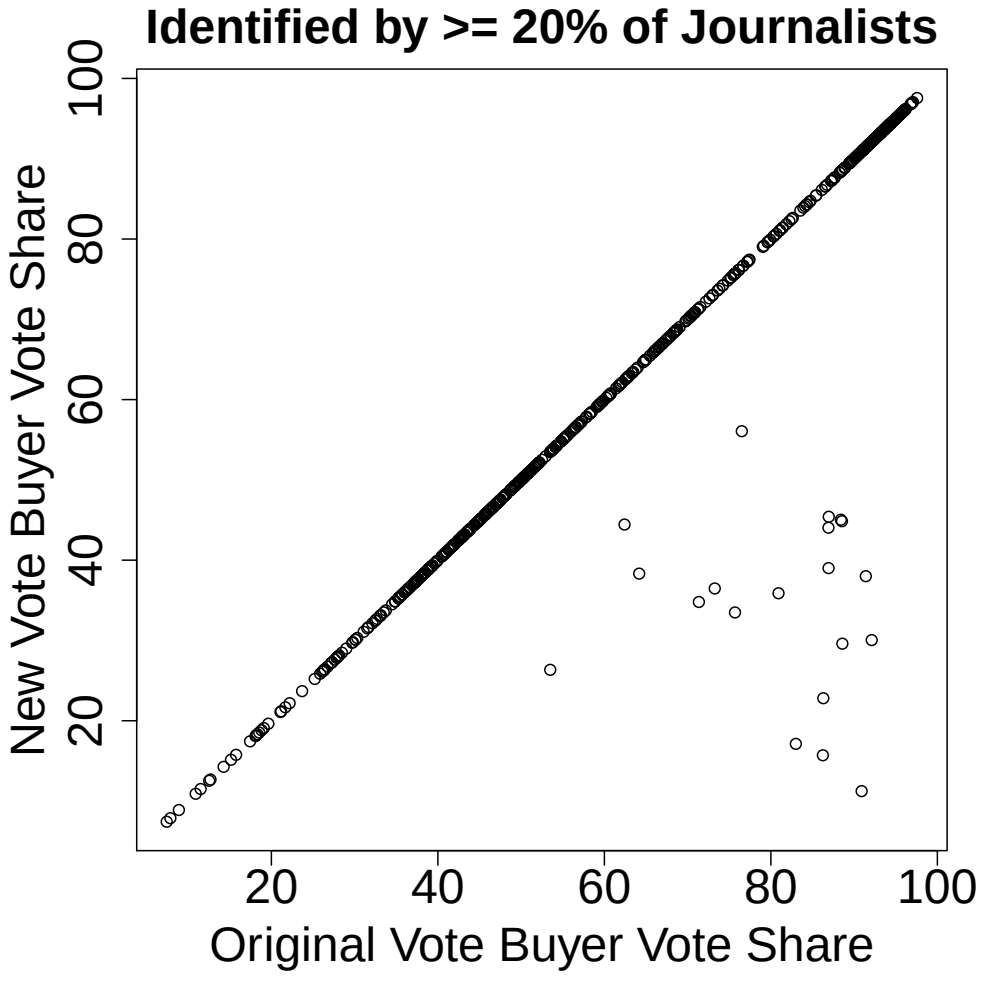
What does it mean to be a vote buying party?



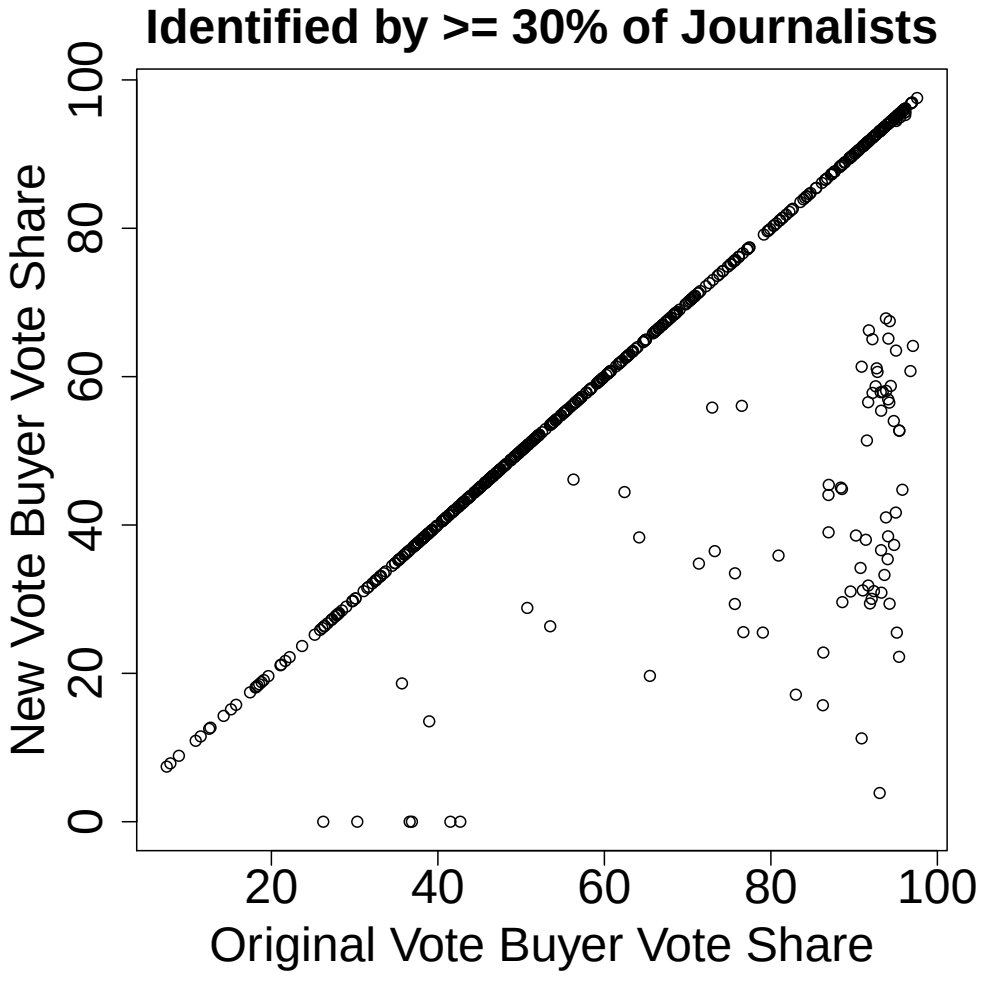
What does it mean to be a vote buying party?



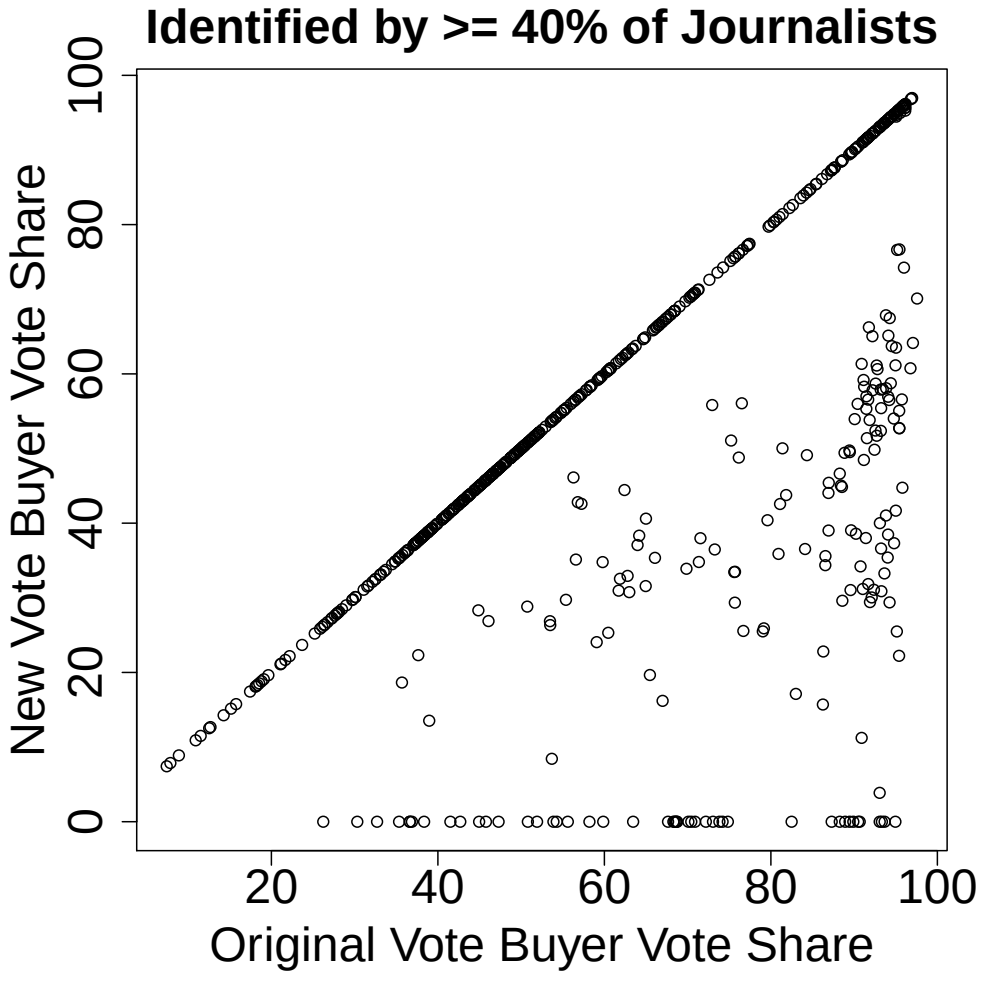
What does it mean to be a vote buying party?



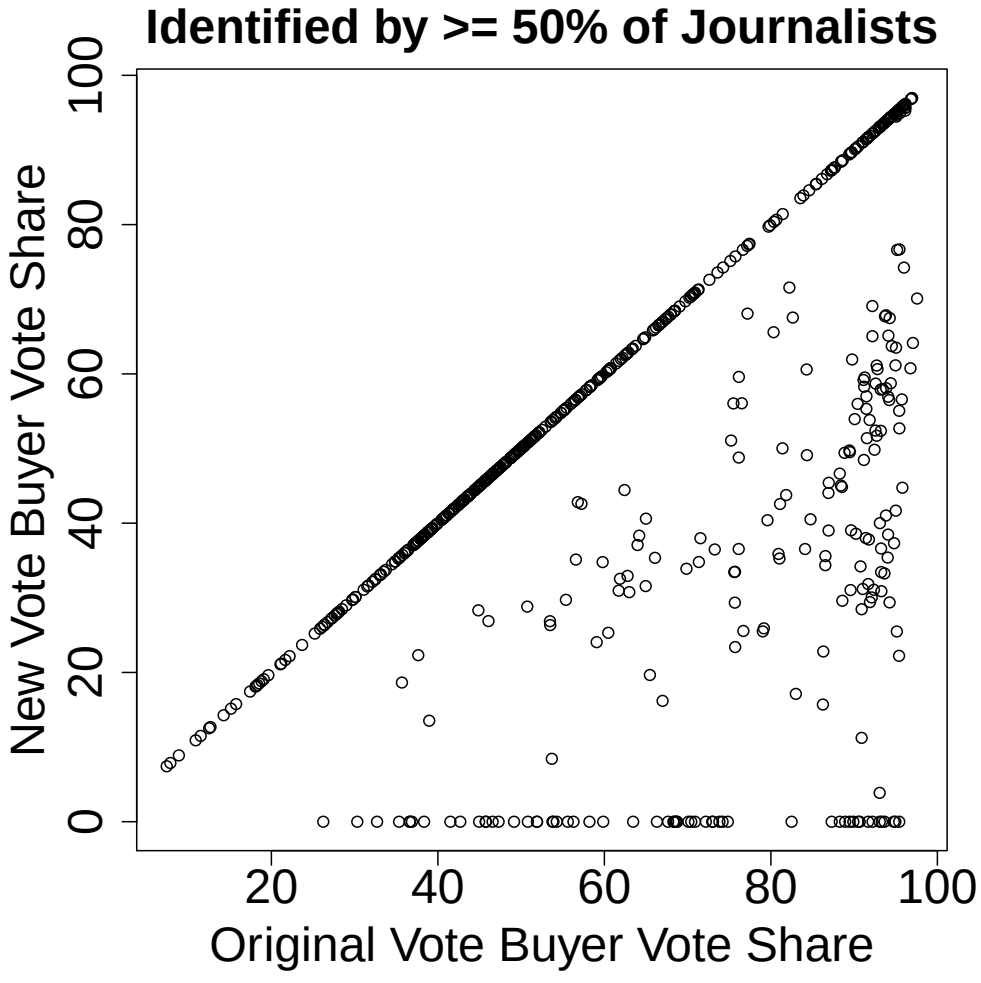
What does it mean to be a vote buying party?



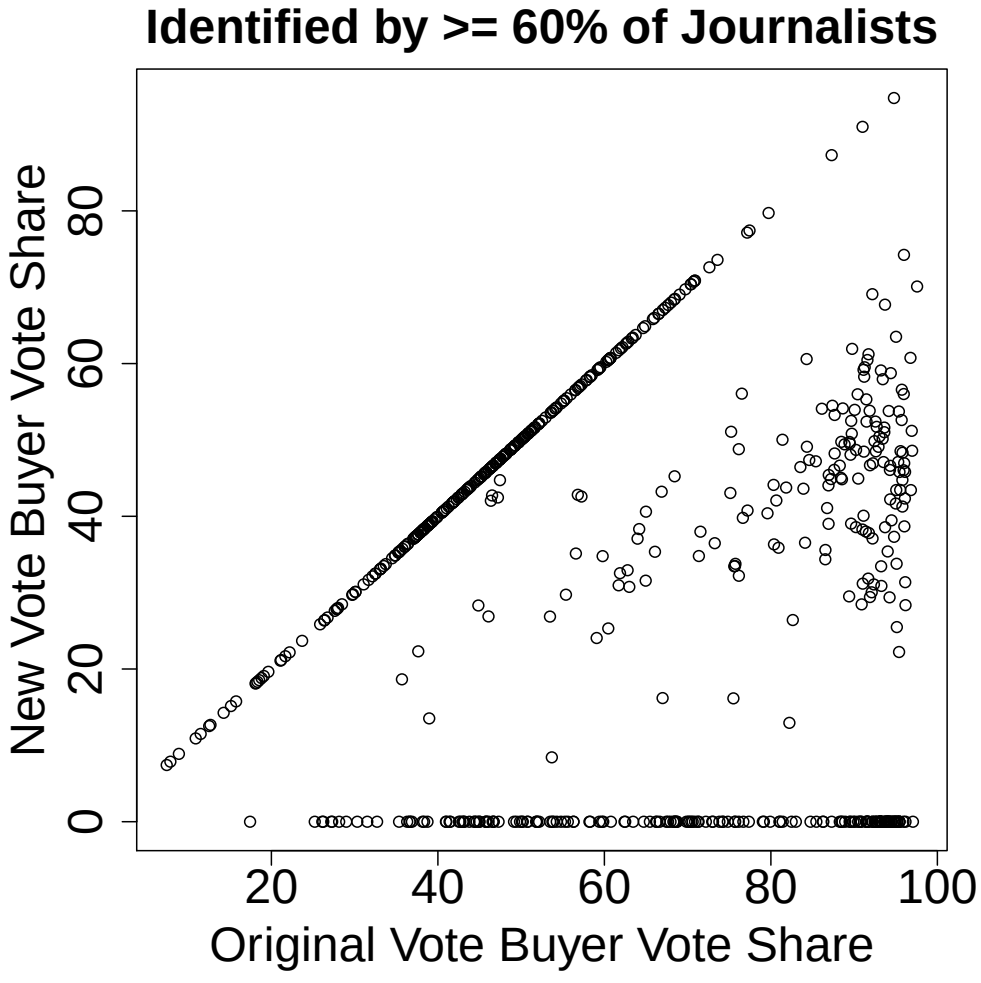
What does it mean to be a vote buying party?



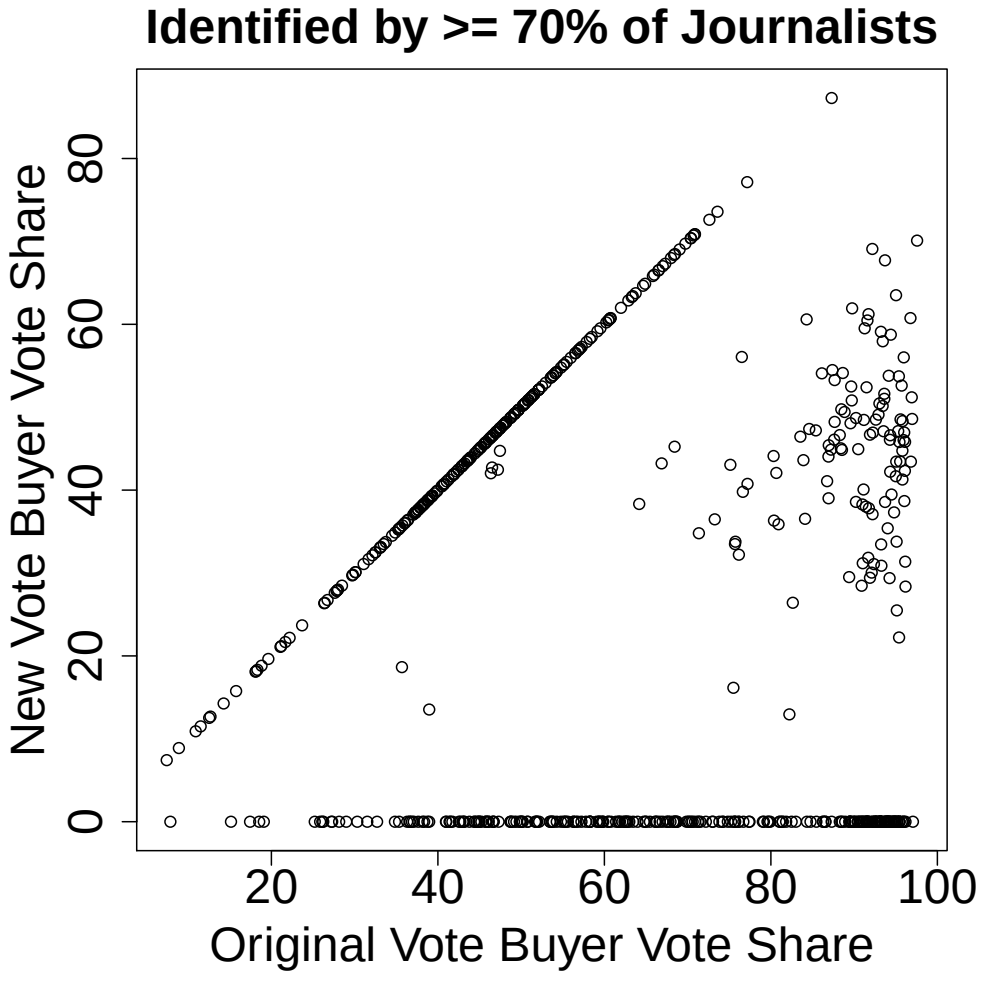
What does it mean to be a vote buying party?



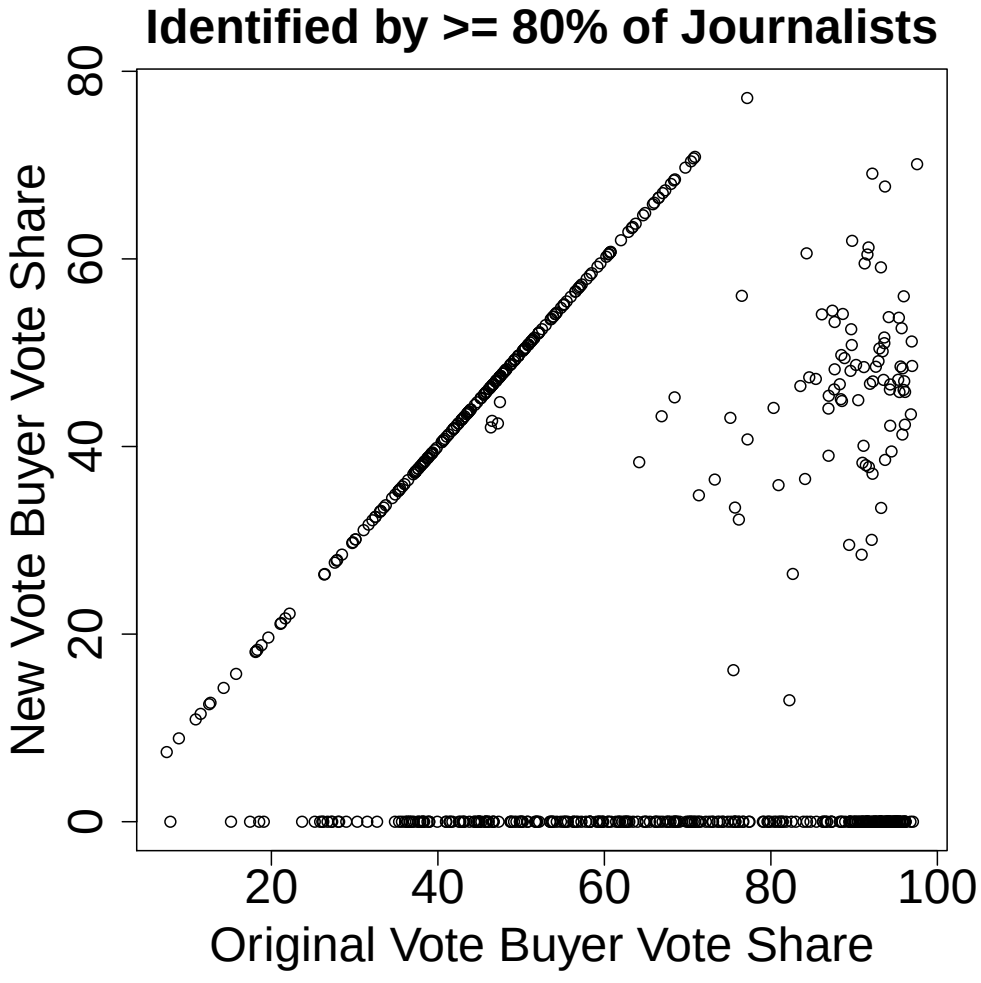
What does it mean to be a vote buying party?



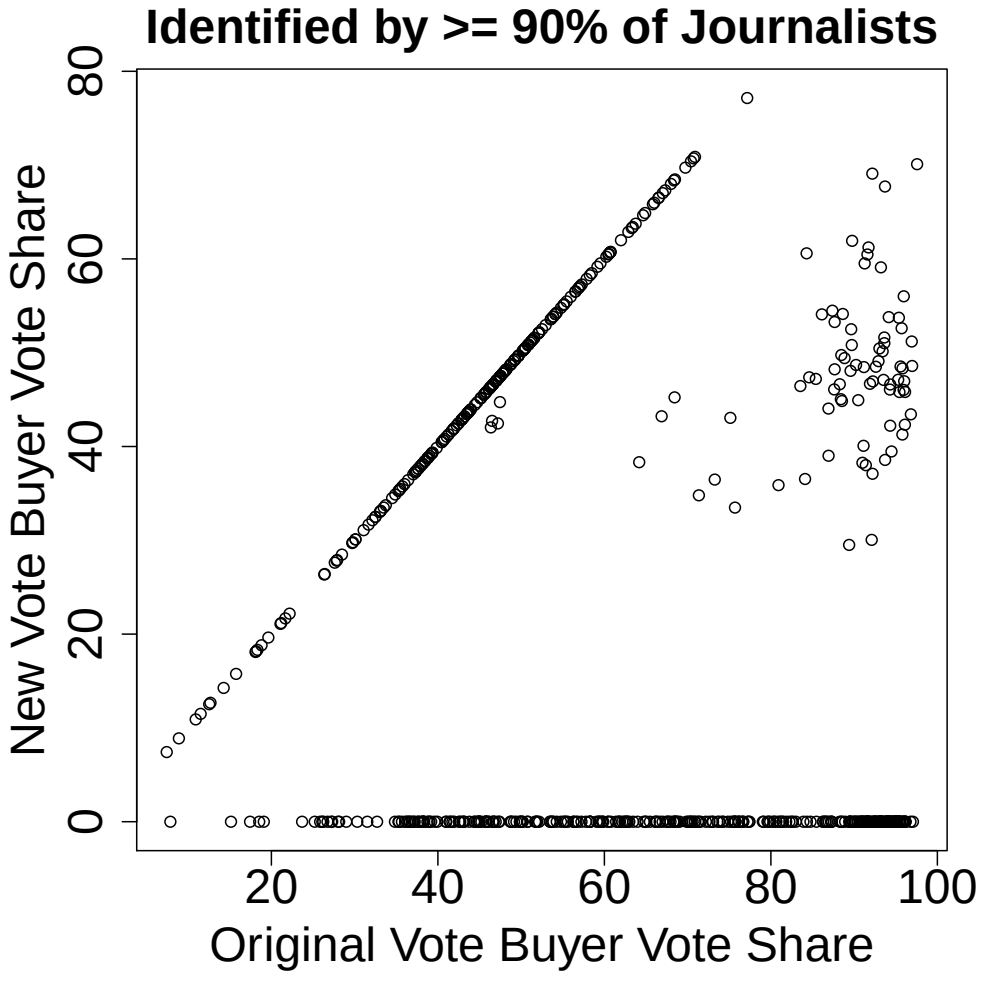
What does it mean to be a vote buying party?



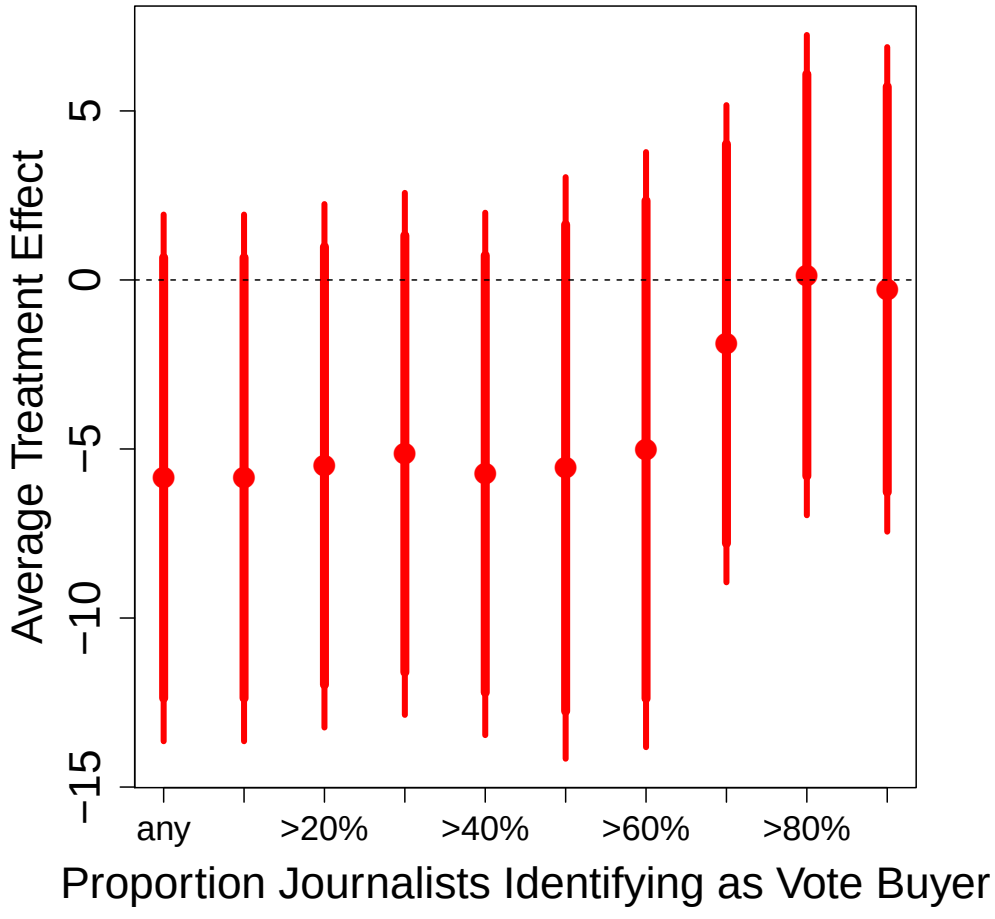
What does it mean to be a vote buying party?



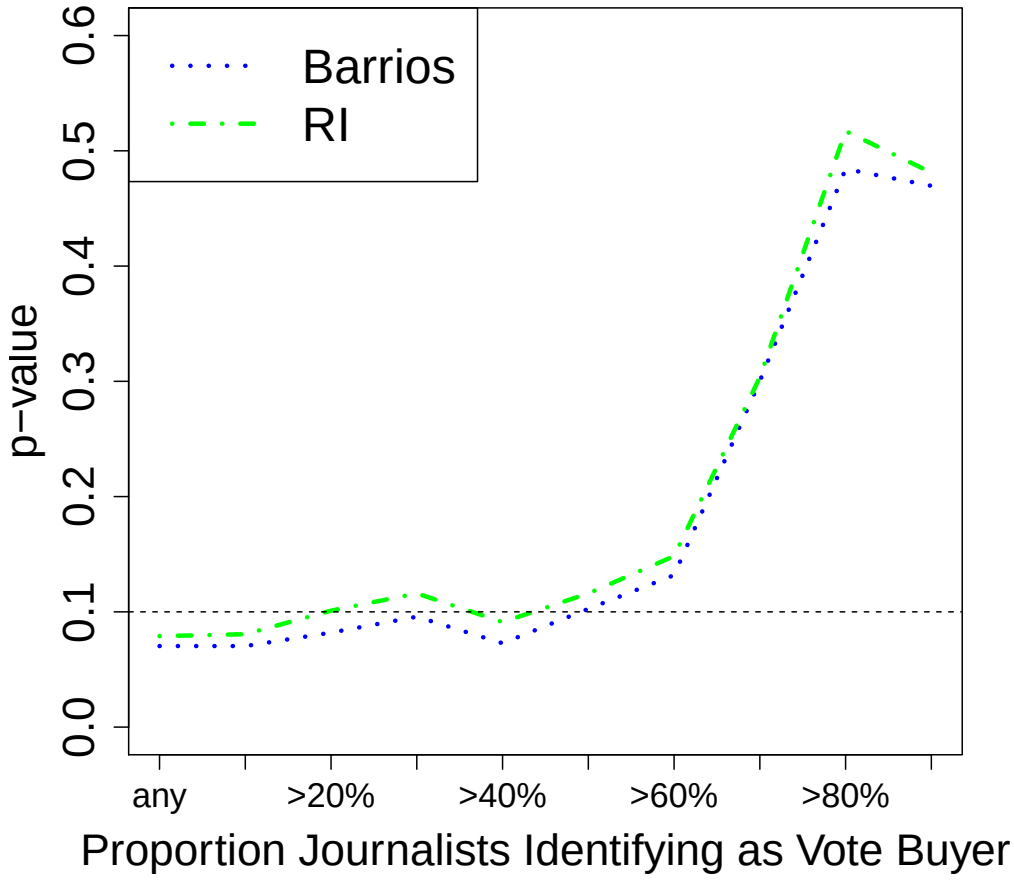
What does it mean to be a vote buying party?



Robustness to the definition of vote buying party



Robustness to the definition of vote buying party



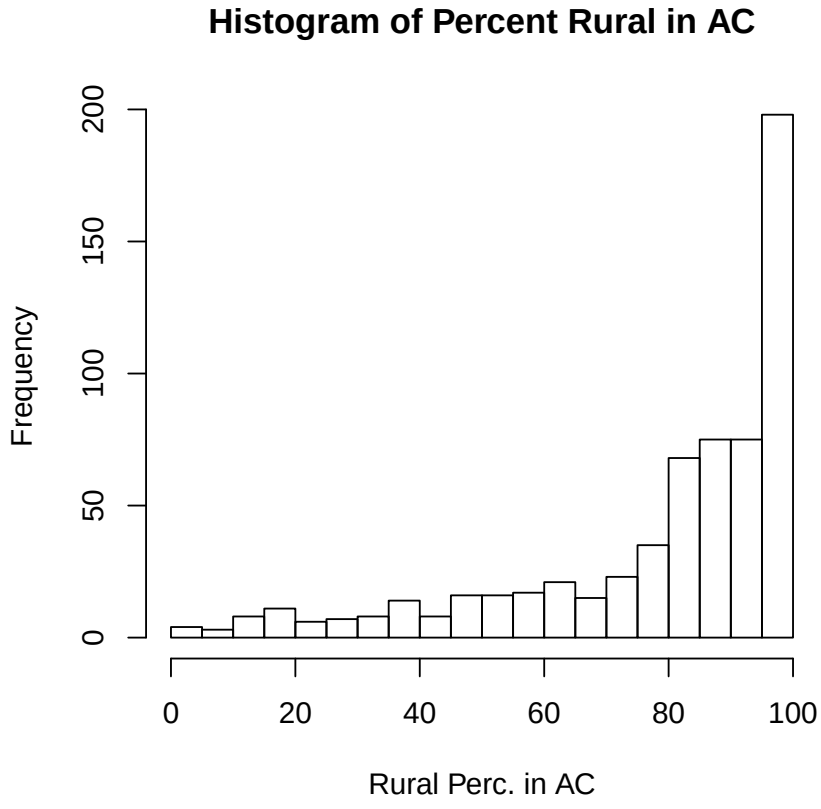
Heterogeneous effects: Urban

Dummy: More than 90% Rural

	Coef.	SE	p
Treat	-4.68	3.6	0.1
Rural >90 pc	1.69	2.55	0.25
Treat:Rural90	-3.16	3.83	0.2
R squared	0.44		

Continuous Rural

	Coef.	SE	p
Treat	1.79	6.79	0.4
Rural pc	-0.01	0.05	0.45
Treat:Rural pc	-0.1	0.06	0.06
R squared	0.44		



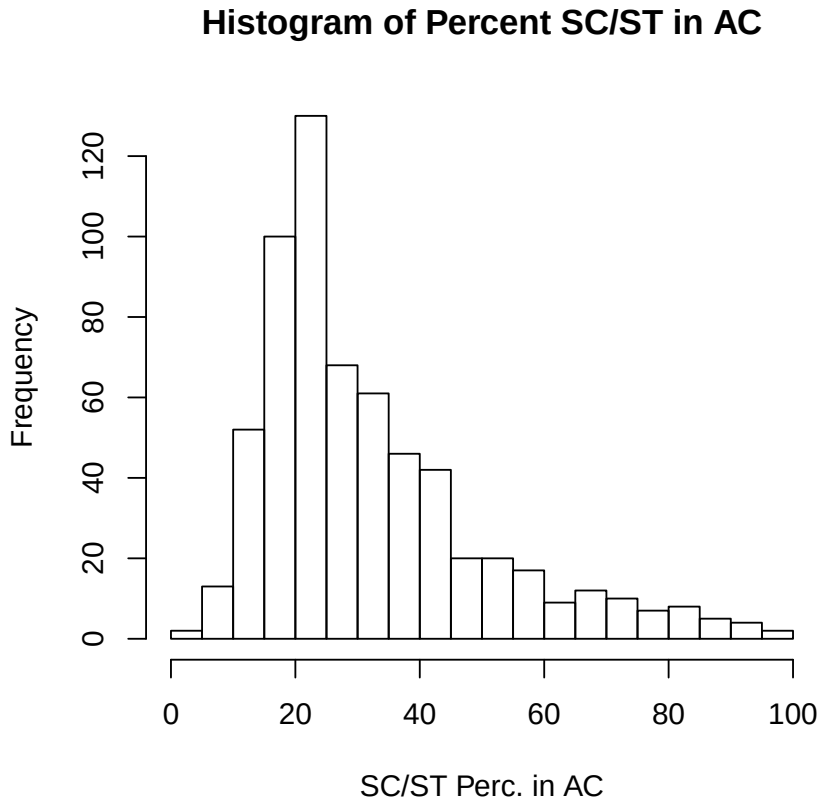
Heterogeneous effects: Minority voters

Dummy: >50% SC/ST

	Coef.	SE	p
Treat	-6	4.4	0.09
SC/ST >50 pc	-4.88	3.66	0.09
Treat:SC/ST50	1.87	5.06	0.36
R squared	0.44		

Continuous SC/ST

	Coef.	SE	p
Treat	-6.03	6.36	0.17
ST/SC pc	-0.06	0.09	0.26
Treat:SC/ST pc	0.01	0.11	0.48
R squared	0.44		



Heterogeneous effects: Date of election

Table : Elec Date by State

	Andhra Pradesh	Bihar	Chattisgarh	Jharkhand	Karnataka
2014-04-10	0	6	0	2	0
2014-04-17	0	8	0	30	75
2014-04-24	0	0	42	0	0
2014-04-30	63	0	0	0	0
2014-05-07	50	0	0	0	0
2014-05-12	0	0	0	0	0

	Madhya Pradesh	Maharashtra	Orissa	Rajasthan	Uttar Pradesh
2014-04-10	17	19	30	0	0
2014-04-17	10	45	19	64	24
2014-04-24	18	34	0	32	4
2014-04-30	0	0	0	0	5
2014-05-07	0	0	0	0	30
2014-05-12	0	0	0	0	1

Heterogeneous effects: Date of election

	Dependent variable:
	Vote Share VB 2014
Treat	−17.979*** (3.543)
Poll 2014-04-17	−3.271 (2.818)
Poll 2014-04-24	0.678 (3.120)
Poll 2014-04-30	−15.522*** (3.316)
Poll 2014-05-07	32.088*** (3.730)
Vote Share VB 2009	0.634*** (0.025)
Num Radio 1	4.927 (15.262)
Num Radio 2	5.406 (15.391)
Treat:Poll 2014-04-17	17.296*** (4.003)
Treat:Poll 2014-04-24	12.344*** (4.445)
Treat:Poll 2014-04-30	22.415*** (5.438)
Treat:Poll 2014-05-07	−13.280*** (4.940)
Constant	26.732* (15.507)
Observations	627
R^2	0.590
Note:	*p<0.1; **p<0.05; ***p<0.01
	Omitted Date 2014-04-10, Excludes 2014-05-12

Table : Treatment by Date

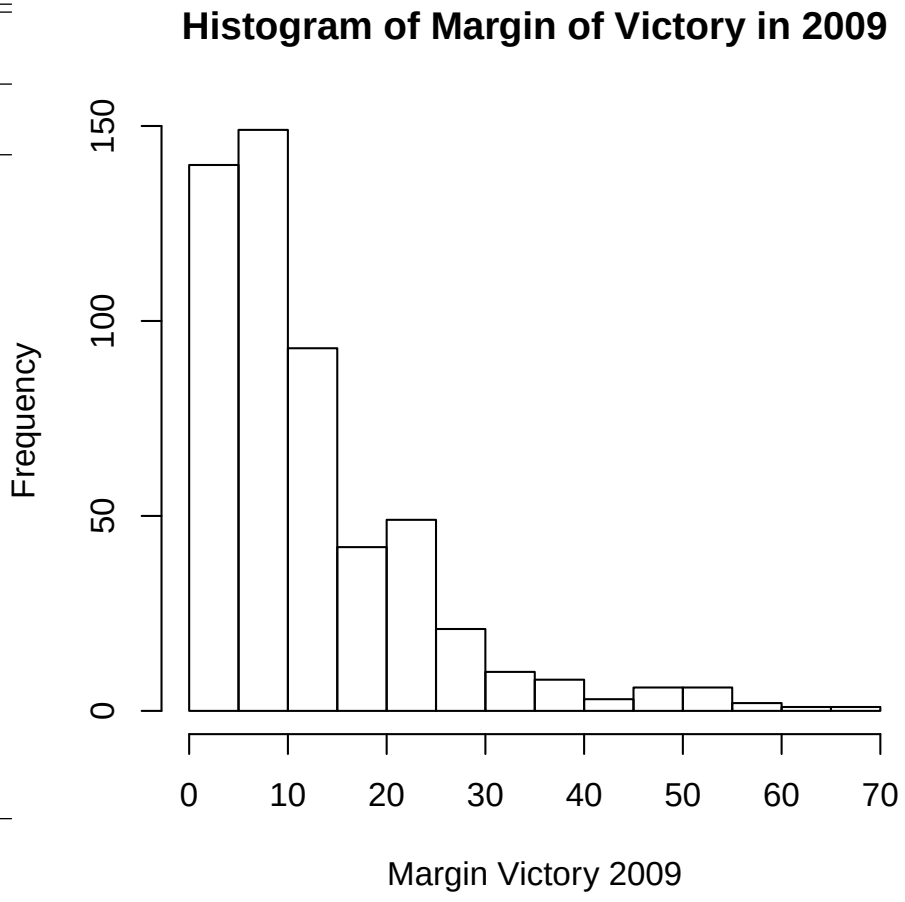
	C	T
2014-04-10	37	37
2014-04-17	131	144
2014-04-24	65	65
2014-04-30	49	19
2014-05-07	33	47
2014-05-12	0	1

** SEs and p-values not adjusted for station-level treatment **

Heterogeneous effects: Competitiveness of election

Table

	Dependent variable:
	VB Share 2014
Treat	−4.213 (3.111)
Margin 5-10	2.896 (3.058)
Margin 10-20	−1.788 (3.042)
Margin 20-30	−0.441 (3.811)
Margin 30+	1.873 (5.742)
VB share 2009	0.557*** (0.031)
1 Station	6.856 (18.436)
2 Stations	9.066 (18.595)
Treat:Margin 5-10	−1.423 (4.333)
Treat:Margin 10-20	1.862 (4.436)
Treat:Margin 20-30	−5.958 (5.393)
Treat:Margin 30+	−5.392 (7.175)
Constant	27.781 (18.620)
Observations	531
R^2	0.407
Note:	*p<0.1; **p<0.05; ***p<0.01



Heterogeneous effects: State

Table : Treatment Status of ACs by State

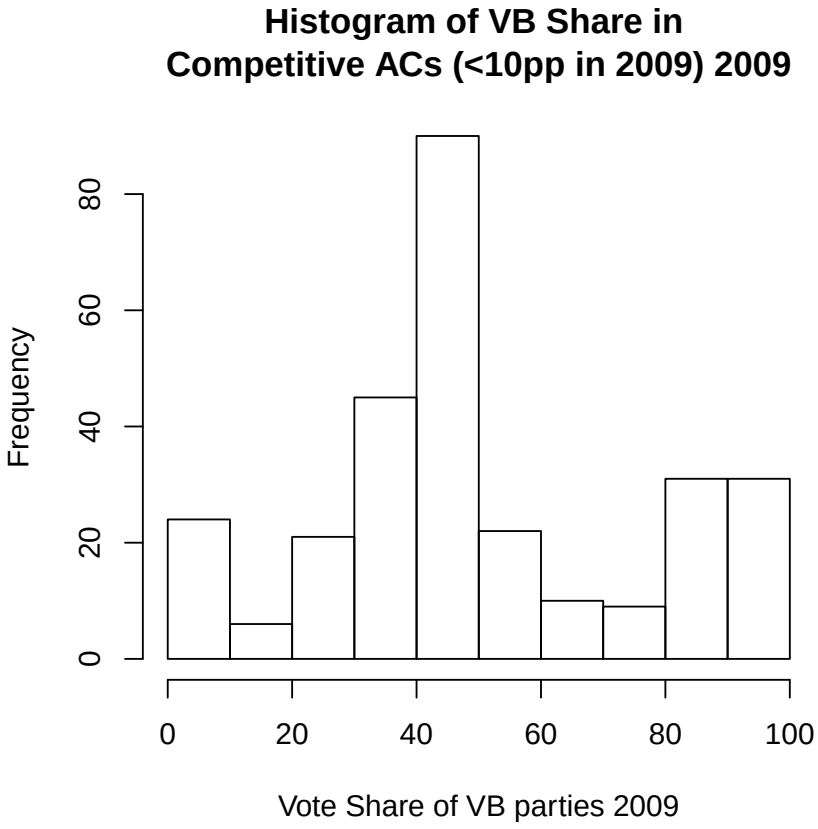
	Control AC	Treated AC
Andhra Pradesh	82	31
Bihar	0	14
Chattisgarh	15	27
Jharkhand	15	17
Karnataka	50	25
Madhya Pradesh	27	18
Maharashtra	60	38
Orissa	23	26
Rajasthan	42	54
Uttar Pradesh	1	63

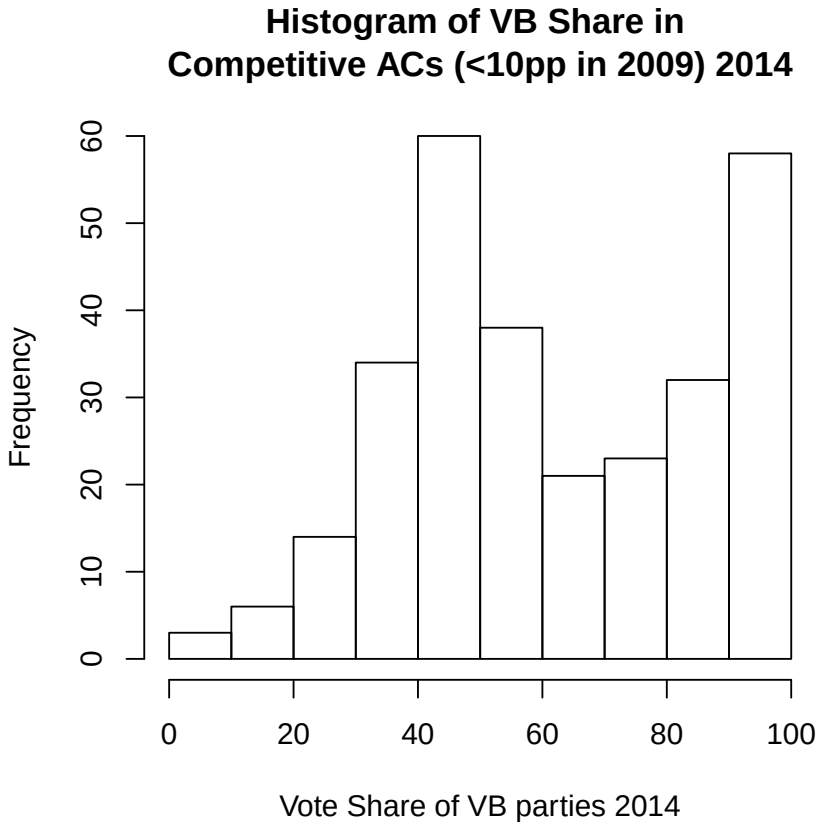
Heterogeneous effects: State

	Dependent variable:
	2014 Vote Share
	Vote Buying Parties
State Bihar	−26.287*** (5.571)
State Chattisgarh	−5.774 (4.893)
State Jharkhand	−3.946 (4.916)
State Karnataka	−8.440*** (3.115)
State Madhya Pradesh	−2.123 (3.875)
State Maharashtra	−4.945* (2.909)
State Orissa	−4.407 (4.084)
State Rajasthan	1.235 (3.420)
State Uttar Pradesh	−61.526*** (17.276)
Vote Share 2009	0.588*** (0.030)
Num Radio 1	2.224 (17.458)
Num Radio 2	1.392 (17.560)
Constant	35.029** (17.569)
Treat	4.353 (3.702)
Treat:Bihar	
Treat:Chattisgarh	−9.903 (6.618)
Treat:Jharkhand	0.761 (7.062)
Treat:Karnataka	−3.484 (5.559)
Treat:Madhya Pradesh	−11.592* (6.357)
Treat:Maharashtra	−8.632* (5.113)
Treat:Orissa	−8.085 (6.116)
Treat:Rajasthan	−14.242*** (5.046)
Treat:Uttar Pradesh	43.523** (17.650)
Constant	35.029** (17.569)
Observations	628
R^2	0.485
Adjusted R^2	0.467
Residual Std. Error	17.111 (df = 606)
F Statistic	27.158*** (df = 21; 606)
Note:	*p<0.1; **p<0.05; ***p<0.01

Interpretation of the results

Common to switch parties and punish incumbents in India.
Among 289 comp. ACs in 2009, 179 switched parties in 2014.





Interpretation of the results

Are people fleeing major parties and voting for minor parties? Does this change the results of elections?

Next step:

- Among ACs competitive in 2009, would the treatment have changed election outcome?
- Check if winner, runner-up parties in 2014 were vote-buyers
- Is the winner non-VB party while runner-up is VB party?
- Check margin of victory in 2014 - smaller than ATE?