14.770: Introduction to Political Economy Lectures 4 and 5: Voting and Political Decisions in Practice

Daron Acemoglu

MIT

September 18 and 20, 2017.

Introduction

- How does voting work out in practice?
- The answer is: in a much more complicated way than the simplest theory would suggest — perhaps not surprisingly.
- In this lecture, I will focus on three aspects of this problem:
  1. To what extent do voters vote strategically? Why do they turn out?
  2. To what extent do the Dowsian prediction of convergence to the middle/the median voter's preferences work out in practice?
  3. Do political decisions reflect the preferences of the median voter/voters?

Why Do Voters Vote?

- As we have seen, it is difficult to get people to turn out if they are voting to be pivotal (unless voting is costless or pleasurable).
- So this means there are three sets of reasons why people might be turning out:
  1. They enjoy voting.
  2. They are subject to social pressure.
  3. They vote because of some moral/ethical considerations.

Do Voters Enjoy Voting?

- This is not an easy question to answer.
- So instead we can look at whether once you induce people to start voting (exogenously), they continue to vote.
- This is the so-called "habit-formation" hypothesis, for which you can go back as far as Aristotle (on ethical behavior feeding into further ethical behavior).
- More recently advocated by Brody and Sniderman (1977). We know that there are significant persistent differences in likelihood of voting across groups and individuals. But a huge identification problem.
- Gerber, Green and Shachar (2003) provide evidence using the vote-canvassing RCT in Connecticut.
- They encourage voting with door-to-door canvassing and phone calls before the 1998 general election, and then look at the effects on voting behavior in 1998, and then in a subsequent election in 1999.

Habit-Formation

- The results are consistent with this hypothesis, though not overwhelming.

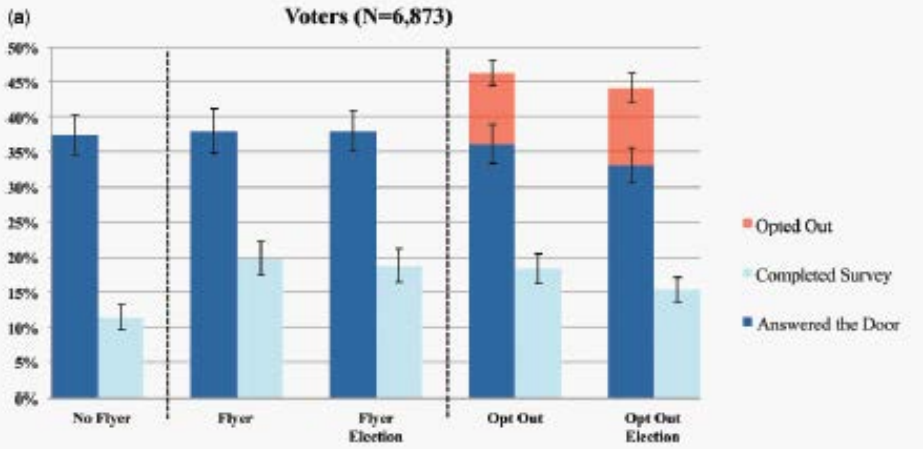
TABLE 2 Voter Turnout in 1998 and 1999, by Treatment Prior to the 1998 Election

	Percentage Voting in 1998	Percentage Voting in 1999	Number of Observations
Personal Canvassing Experiment			
Subjects in the control group	48.1%	39.2%	20,250
Subjects in the treatment group	51.1	40.3	4,950
Direct Mail Experiment			
Subjects in the control group	48.5	39.2	12,565
Subjects sent one piece of mail	47.7	38.3	4,087
Subjects sent two pieces of mail	49.0	39.3	4,341
Subjects sent three pieces of mail	50.0	41.1	4,207

Social Pressure

- By social pressure, I mean the fact that voters do not really enjoy voting, but feel compelled to do so because others will shun or ostracize them if they are seen not to vote.
- This idea is investigated in a recent creative paper by Della Vigna et al. (2017).
- They design a field experiment with door-to-door canvassing in Chicago following the 2010 congressional elections.
- The creative new element is that the arrival of canvassers is preannounced to one of the treatment groups, and incentives to lie about past voting are manipulated.
- *Design*: no flyer group receives no flyers, treatment groups receive flyers that do or do not mentioned election, and the opt out groups receive a flyer with a box to check if they do not want to be disturbed.

Social Pressure (continued)



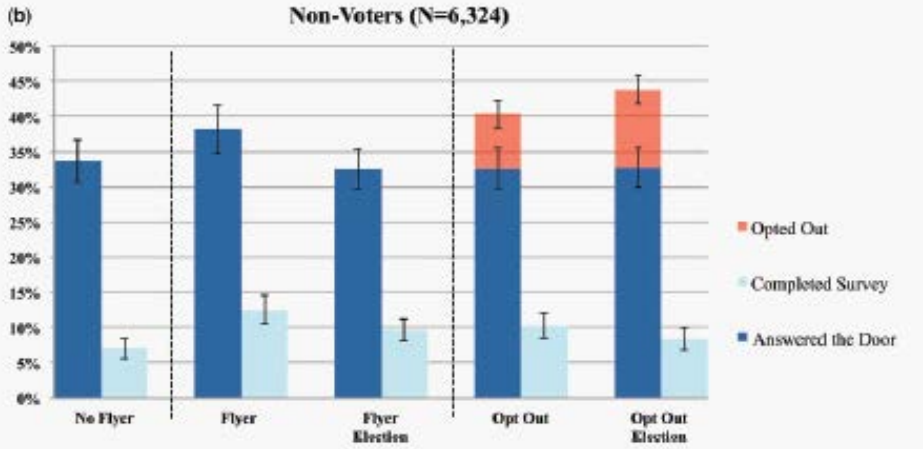


FIGURE 5

Response to information about election in flyer

Note: Figure 5 presents the share of households answering the door, the (unconditional) share completing the survey, and (when applicable) the share opting out, separately for each of the five flyer treatments and separately for voting households and non-voting households. The averages are pooled across the three different payment and duration treatments featured in Figure 3. Standard errors are clustered at the solicitor-date level.

Social Pressure (continued)

- The results indicate significant "social image" considerations: people avoid the canvassing when they are informed that there will be questions about past voting.
- They are also willing to pay significantly to avoid this.
- In particular, rates of answering the door and completing the survey are lower among non-voters if the flyer mentions the election, and non-voters given the opt-out option that mentions election are significantly less likely to answer the door.

Rule Utilitarianism

- John Harsanyi proposed the idea of rule utilitarianism, whereby individuals vote taking their group's interest, rather than their own interest, into account.
- Thus individuals may turn out even if it is costly for them because they are adopting a rule that they want others to adopt also ("do unto others as you would like them to do unto you...")
- Some argue that this perspective is most useful for thinking about voting behavior, but direct empirical evidence is difficult to generate.

Do Voters Vote Strategically?

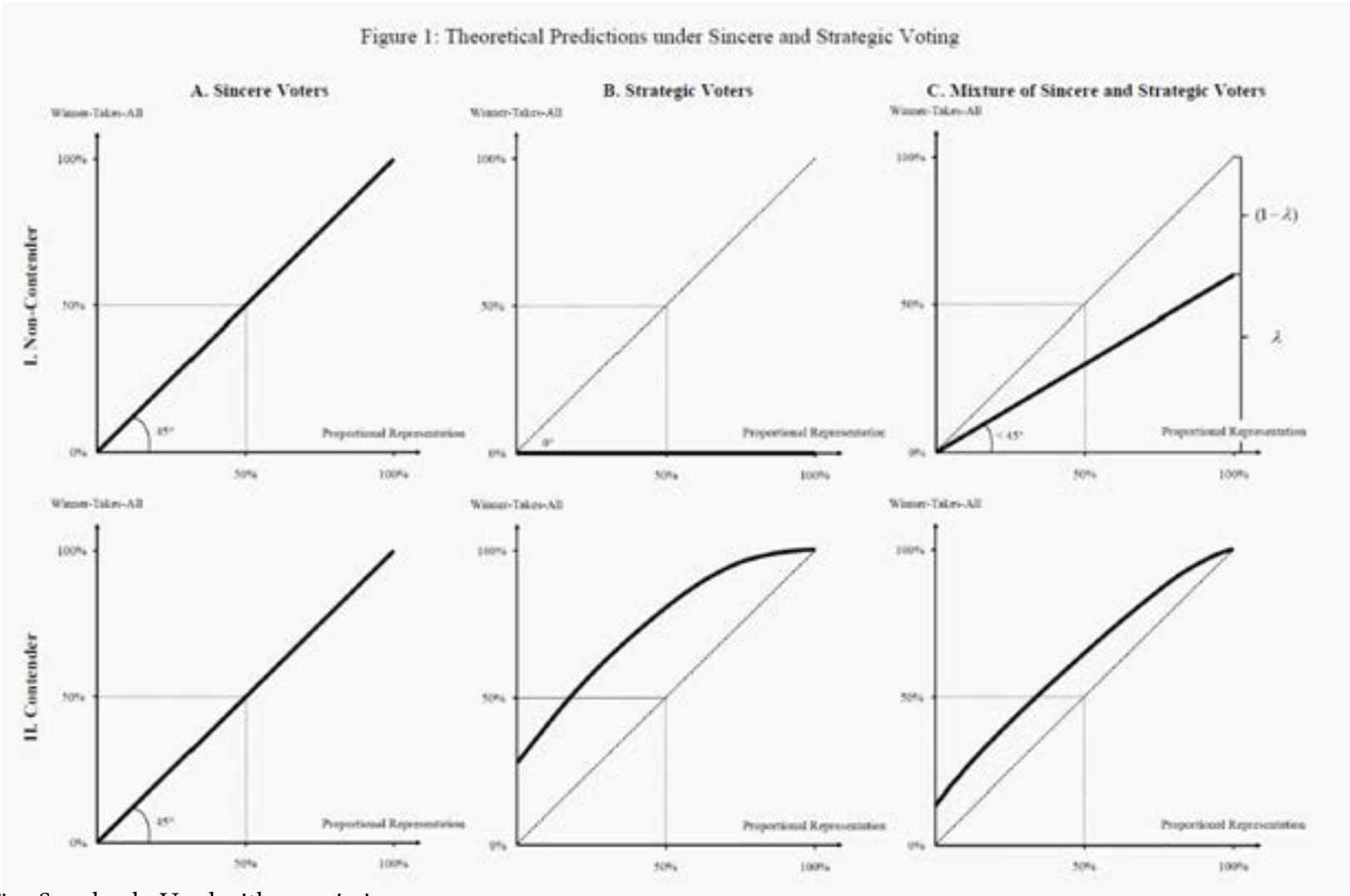
- We have seen that whether voters vote sincerely or strategically matters in the presence of common-interest policy choices and incomplete information.
- Even more simply, the same issues arise when there are more than two candidates/options in an election.
- Why should you vote for somebody who is your first choice but sure to lose when you can support somebody that has a chance to win?
- The problem is that we know people do support sure losers, so either not everybody votes strategically or there are other considerations (direct utility?).
- Part of the literature investigates whether there is any evidence for strategic voting and how important it is.

Testing Strategic Voting Using Two Linked Elections

- Spenkuch (2017) uses the German voting system, where each individual has two votes — a *list vote* for a party, counted that the national level, which approximates a proportional voting system; and a *candidate vote*, counted at the district level in a first-past-the-post electoral system.
- As is well known, in proportional voting systems (barring issues about strategic thinking on legislative bargaining etc.), individuals have incentives to vote sincerely.
- In first-past-the-post elections, there are reasons for deviating from sincere voting.

Simple Theory

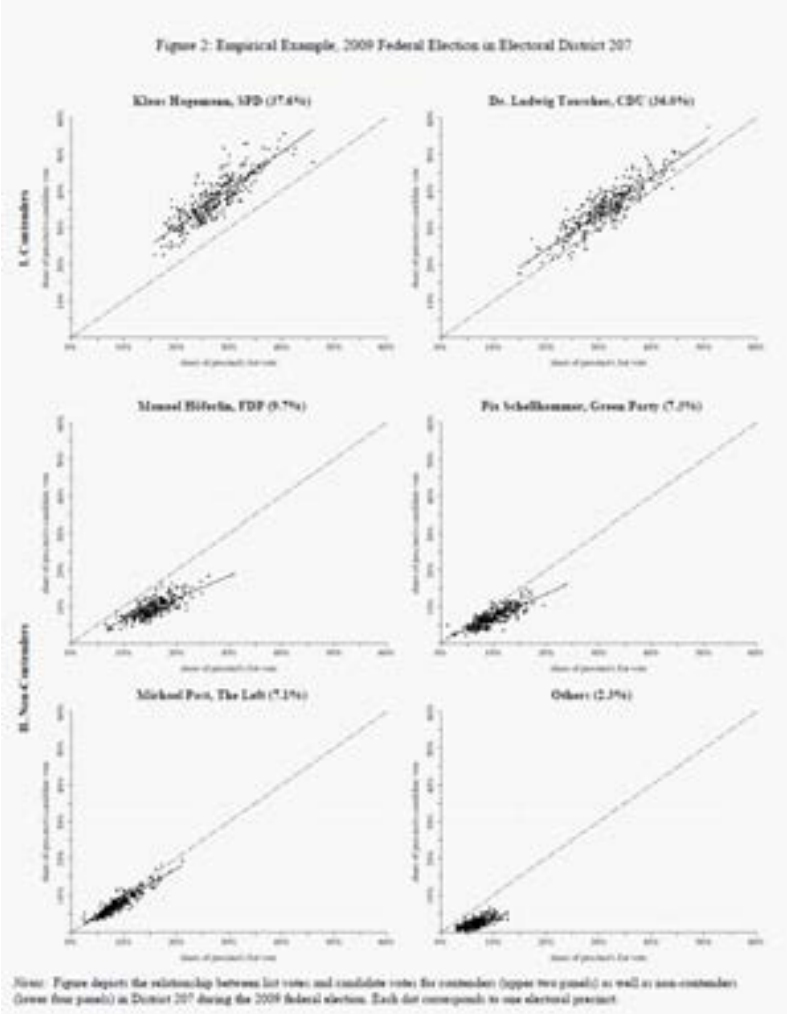
- Summarizing the previous theoretical expectation:



Courtesy of Jörg Spenkuch. Used with permission.

How Do the Data Lineup?

- In one electoral district

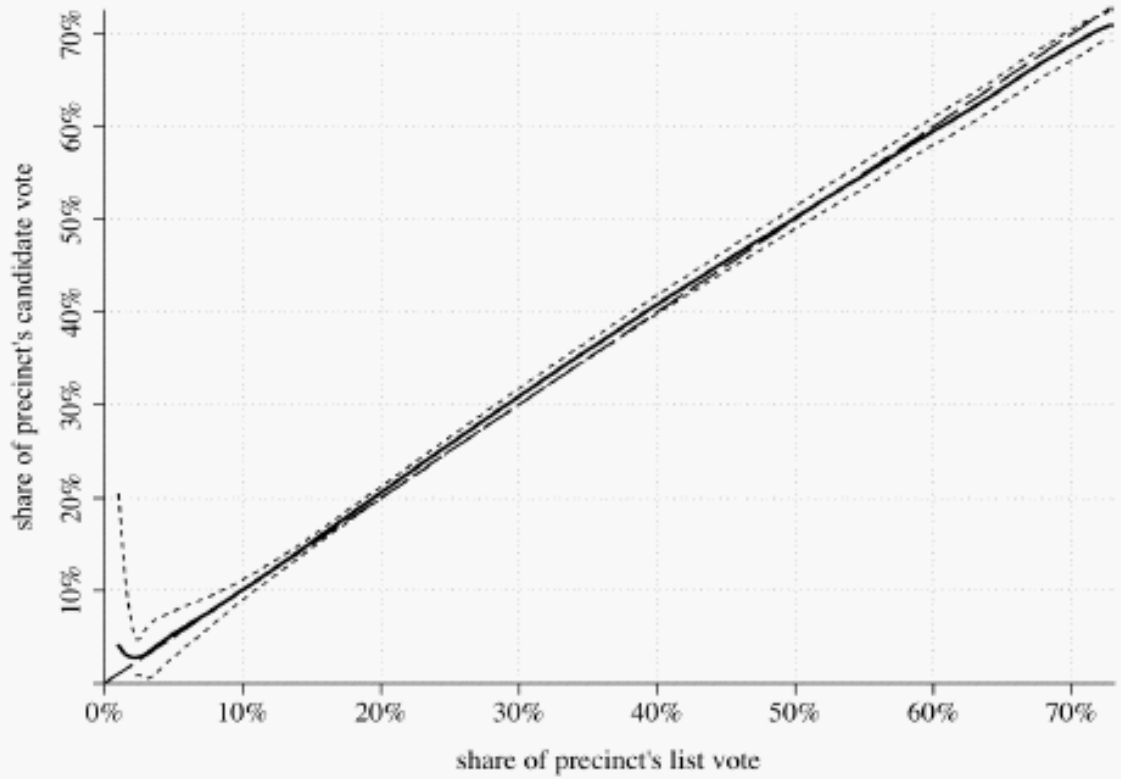


Courtesy of Jörg Spenkuch. Used with permission.

How Do the Data Line up? (continued)

- On average for contenders:

Figure 4: Relationship between List and Candidate Votes for Contenders

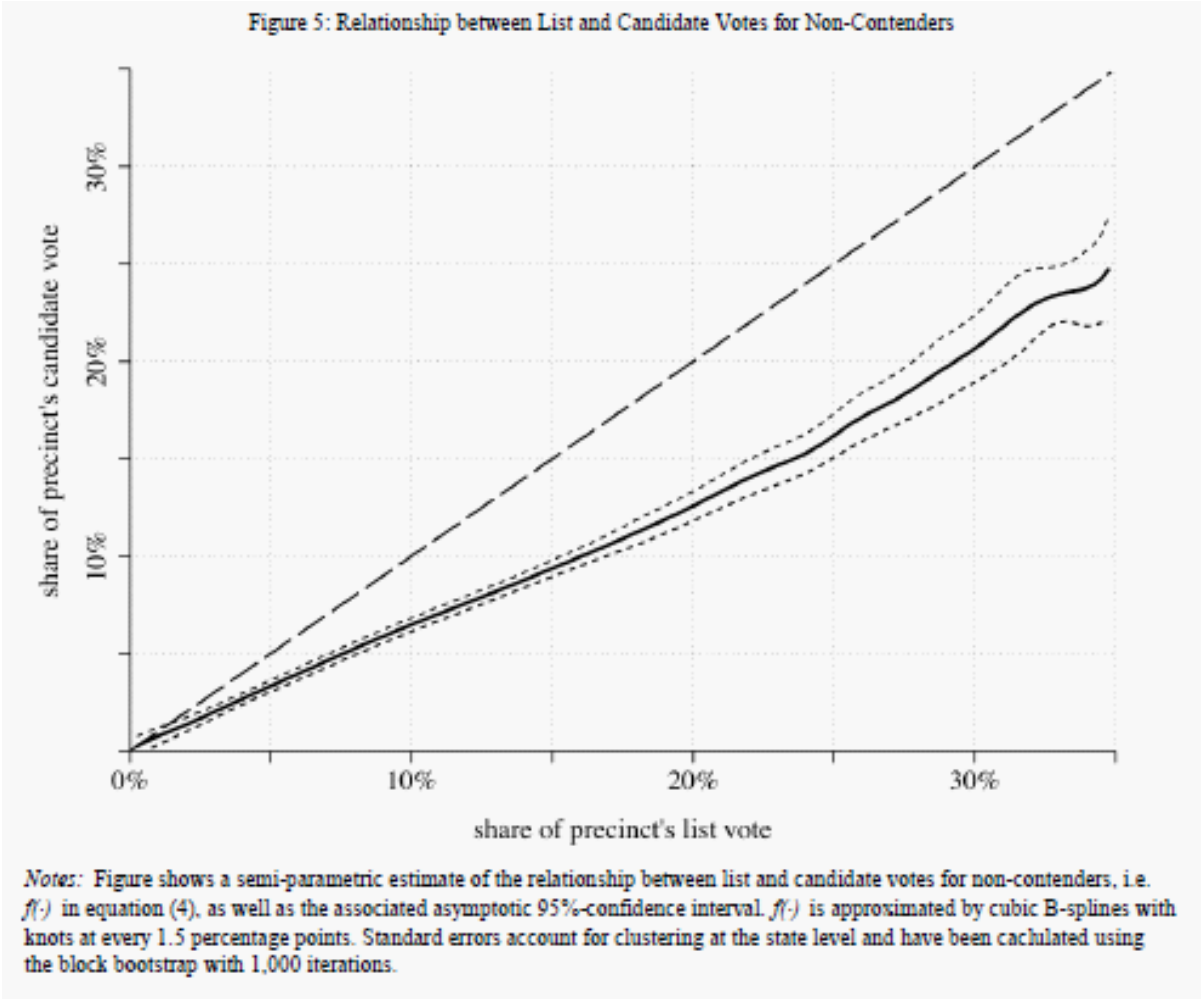


Notes: Figure shows a semi-parametric estimate of the relationship between list and candidate votes for district winners and runner-ups, i.e. $f(\cdot)$ in equation (4), as well as the associated asymptotic 95%-confidence interval. $f(\cdot)$ is approximated by cubic B-splines with knots at every 3 percentage points. Standard errors account for clustering at the state level and have been caclulated using the block bootstrap with 1,000 iterations.

Courtesy of Jörg Spenkuch. Used with permission.

How Do the Data Line up? (continued)

- On average for non-contenders:



Courtesy of Jörg Spenkuch. Used with permission.

Duverger's Law and Strategic Voting

- One implication of strategic voting is Duverger's Law, which claims that with simple majority, single-ballot elections, there will be a strong tendency towards a two-party system, because strategic voting considerations will make voters shun non-contender parties. In contrast, proportional representation or dual-ballot system can support multiple parties.
- Fujiwara (2011) tests this implication using a regression-discontinuity design based on different voting systems in place in Brazilian municipalities based on population.
- In municipal elections (for mayors), a single ballot or the dual ballot system is used below and above the cutoff of 200,000.

Evidence for Duverger's Law

RD estimate:

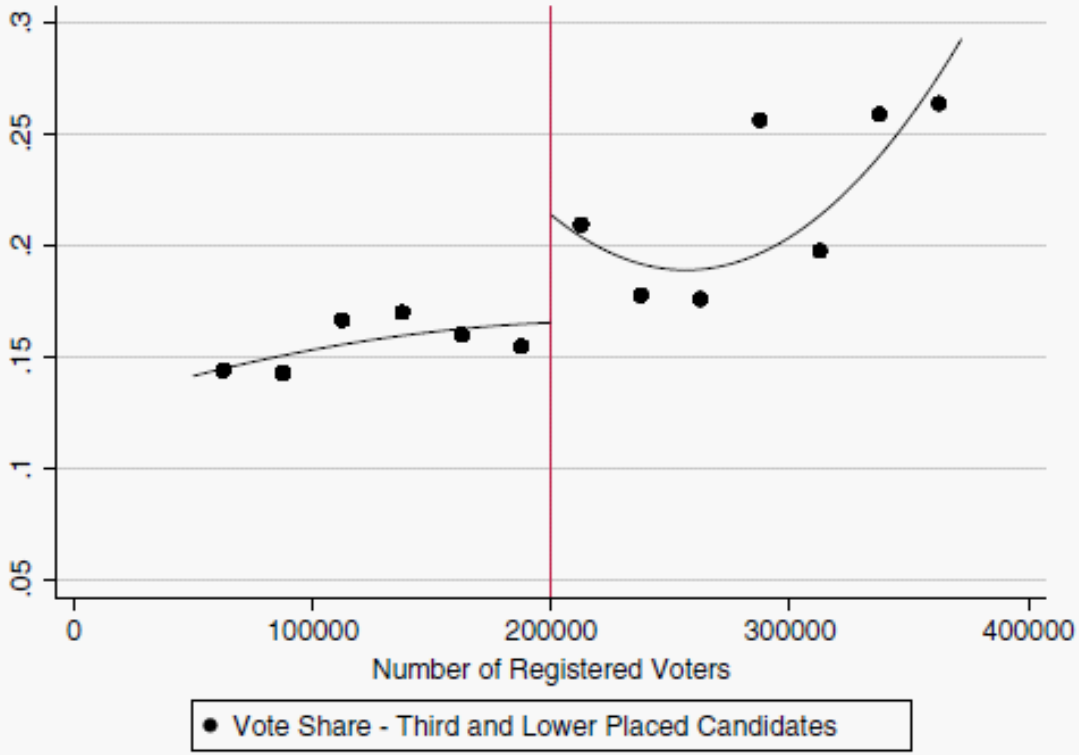


Figure 1. Vote share of third and lower placed candidates — local averages and parametric fit.

Courtesy of Thomas Fujiwara. Used with permission.

Evidence for Duverger's Law (continued)

Table 1. Treatment effects on electoral outcomes.

Specification/ bandwidth	Single-ballot mean	Linear 50,000	Linear 25,000	Linear 75,000	Quad. 50,000	Quad. 75,000
Dependent variable		(1)	(2)	(3)	(4)	(5)
Vote share — 3rd and lower placed candidates	0.155	0.088	0.093	0.069	0.104	0.113
		(0.040)	(0.056)	(0.033)	(0.058)	(0.046)
Vote Share — 4th and lower placed candidates	0.041	0.043	0.046	0.036	0.057	0.055
		(0.024)	(0.030)	(0.021)	(0.031)	(0.028)
Vote Share — 5th and lower placed candidates	0.012	0.015	0.017	0.015	0.022	0.021
		(0.010)	(0.012)	(0.009)	(0.012)	(0.011)
Registration rate	0.638	0.011	0.016	0.021	0.031	0.014
		(0.019)	(0.030)	(0.016)	(0.029)	(0.024)
Turnout rate	0.851	0.003	−0.004	0.002	−0.003	−0.002
		(0.007)	(0.011)	(0.007)	(0.01)	(0.009)
Observations	—	175	81	282	175	282

Robust standard errors clustered at the municipality level in parenthesis. Each figure in the table is from a separate local linear/quadratic regression with the specified bandwidth. The level of observation is a municipal election. The estimated treatment effect is of a change from SB to DB. All estimates include year effects. Details on the dependent variables are presented in the text.

Evidence for Duverger's Law (continued)

- Results driven by elections predicted to be contested:

Table 3. Treatment effects in contested and uncontested elections.

Specification/ bandwidth	SB mean	Linear 50,000	Linear 25,000	Linear 75,000	Quad. 50,000	Quad. 75,000
		(1)	(2)	(3)	(4)	(5)
Panel A: Elections predicted to be contested						
Vote share — 3rd and lower placed candidates	0.148	0.157 (0.076)	0.145 (0.107)	0.144 (0.061)	0.145 (0.081)	0.177 (0.083)
Observations	—	64	25	109	64	109
Panel B: Elections predicted to be uncontested						
Vote share — 3rd and lower placed candidates	0.138	0.015 (0.049)	0.001 (0.075)	0.011 (0.039)	0.003 (0.075)	0.032 (0.057)
Observations	—	80	40	123	80	123

Robust standard errors clustered at the municipality level in parenthesis. Each figure in the table is from a separate local linear/quadratic regression with the specified bandwidth. The level of observation is a municipal election. All estimates include year effects. Details on the dependent variables are presented in the text.

Summary

- Overall, quite a bit of evidence that there is some strategic voting, and perhaps quite a bit of it.
- But this evidence doesn't really speak to whether people are very sophisticated or just so-so strategic.
- Strategic voting may not be inconsistent with rule-utilitarianism either.

Testing Dowsian Convergence

- The Dowsian convergence result, discussed in the first two lectures, is viewed as iconic of basic voting theory.
- It has attracted considerable attention from social scientists and beyond.
- As these things go, it is also a relatively easy theory to test.
- One approach is to use regression discontinuity design: holding the ideology of the electorate constant, which party gets elected shouldn't matter if we are indeed in the Dowsian world.
- Several papers have attacked this problem.

Importance of Party Identity in the US

- Lee et al. (2004) do this using US Congress elections.
- They focus on basic regression discontinuity estimates and look at nominate scores as a summary of the voting record (from rollcall votes) of U.S. House members.
- They also look at likelihood of voting the same way as the Democratic Party leader.

No Dowsian Convergence in the US

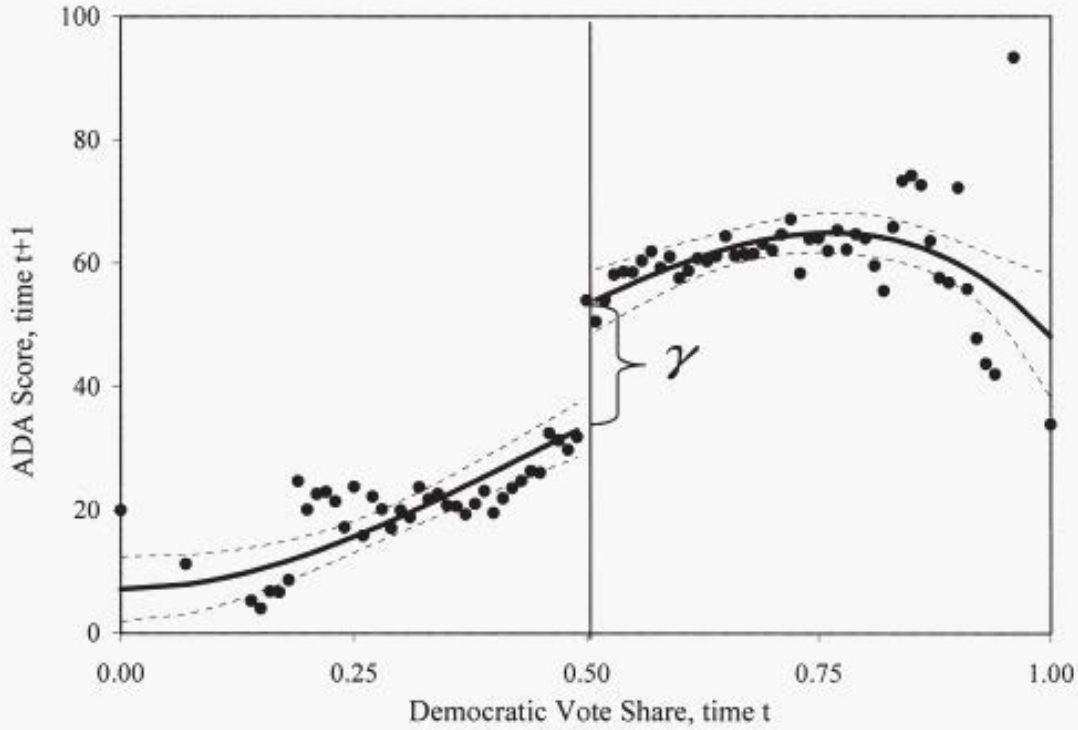


FIGURE I

Total Effect of Initial Win on Future ADA Scores: γ

This figure plots ADA scores after the election at time $t + 1$ against the Democrat vote share, time t. Each circle is the average ADA score within 0.01 intervals of the Democrat vote share. Solid lines are fitted values from fourth-order polynomial regressions on either side of the discontinuity. Dotted lines are pointwise 95 percent confidence intervals. The discontinuity gap estimates

$$\gamma = \underbrace{\pi_0(P^{*D}_{t+1} - P^{*R}_{t+1})}_{\text{“Affect”}} + \underbrace{\pi_1(P^{*D}_{t+1} - P^{*R}_{t+1})}_{\text{“Elect”}}.$$

No Dowsian Convergence in the US (continued)

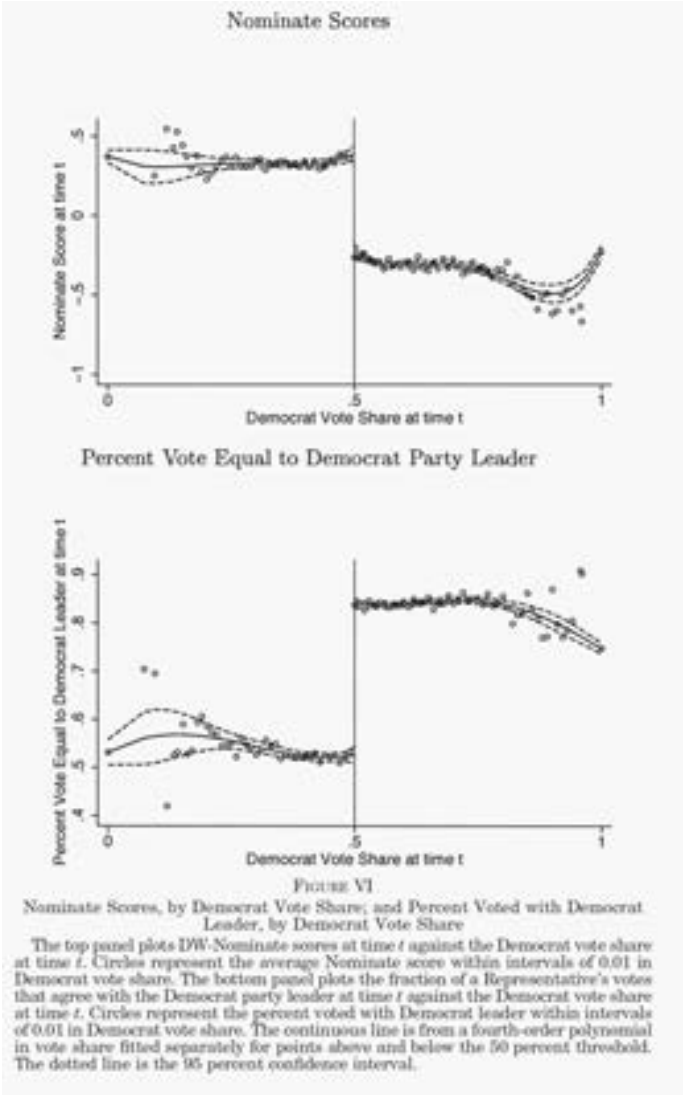


FIGURE VI

Nominate Scores, by Democrat Vote Share; and Percent Voted with Democrat Leader, by Democrat Vote Share

The top panel plots DW-Nominate scores at time t against the Democrat vote share at time t. Circles represent the average Nominate score within intervals of 0.01 in Democrat vote share. The bottom panel plots the fraction of a Representative's votes that agree with the Democrat party leader at time t against the Democrat vote share at time t. Circles represent the percent voted with Democrat leader within intervals of 0.01 in Democrat vote share. The continuous line is from a fourth-order polynomial in vote share fitted separately for points above and below the 50 percent threshold. The dotted line is the 95 percent confidence interval.

Interpretation

- Clear partisan behavior from marginally elected Democrats and/or Republicans.
- Does this clearly reject Dowsian policy convergence?
- Yes and no — elected representatives are clearly not the same regardless of which party they come from; but they are not determining policy (they may be non-pivotal in the House).
- The pure Dowsian framework requires policy to be convergent — and thus its rejection requires that we show party identity to matter for policy.
- This is what Pettersson-Lidbom (2008) does using data from Swedish municipalities, and finds candidates from the social democrats to lead to higher spending and taxes, and more government employees, and lower unemployment..

No Dowsian Convergence in Sweden

TABLE 7. Party effect: Fiscal policies.

	1	2	3	4	5	6	7
Log (Total spending per capita)	0.024**	0.027***	0.023**	0.021**	0.024*	0.020**	0.022**
	(0.009)	(0.009)	(0.010)	(0.010)	(0.013)	(0.0009)	(0.010)
Log (Total spending as a share of income)	0.021**	0.025**	0.024**	0.025**	0.034*	0.021**	0.024***
	(0.010)	(0.010)	(0.010)	(0.011)	(0.018)	(0.009)	(0.009)
Log (Current spending per capita)	0.024**	0.027***	0.027**	0.026**	0.019	0.025**	0.027**
	(0.010)	(0.010)	(0.011)	(0.011)	(0.013)	(0.010)	(0.011)
Log (Current spending as a share of income)	0.022*	0.025**	0.028**	0.030***	0.029	0.026***	0.029***
	(0.011)	(0.011)	(0.012)	(0.012)	(0.018)	(0.009)	(0.010)
Log (Total revenues per capita)	0.024***	0.027***	0.019**	0.017*	0.015	0.017*	0.014
	(0.009)	(0.009)	(0.009)	(0.009)	(0.013)	(0.009)	(0.010)
Log (Total revenues as a share of income)	0.021**	0.025**	0.020**	0.021**	0.025	0.018**	0.017*
	(0.010)	(0.010)	(0.010)	(0.010)	(0.018)	(0.009)	(0.009)
Log (Proportional income tax rate)	0.012***	0.013***	0.012***	0.013***	0.011	0.013***	0.014***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.008)	(0.004)	(0.004)
Sample	Full	Full	Full	Full	±2	Full	Full
Left vote share polynomial	First	Second	Third	Fourth	None	Fourth	Fourth × time
Controls	No	No	No	No	No	Yes	Yes

Note: Standard errors clustered at the local government's term in office level are within parentheses. Each entry is a separate regression. All regressions also include, but do not report, municipality specific effects, time effects, and an indicator for undefined majority governments. The full sample includes 5,913 observations and the ±2 sample include all observations that are in the range of [48, 52] of the left vote share and there are 828 such observations.

*Significant at 10%; **significant at 5%; ***significant at 1%.

No Dowsian Convergence in Sweden (continued)

TABLE 8. Party effect: Economic policies.

	1	2	3	4	5	6	7
Log (Unemployment rate)	−0.017	−0.032	−0.056*	−0.056*	−0.121	−0.048	−0.070**
	(0.033)	(0.031)	(0.032)	(0.032)	(0.089)	(0.031)	(0.033)
Log (Government employees per capita)	0.030**	0.033***	0.035***	0.036***	0.039***	0.032***	0.036***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.016)	(0.011)	(0.012)
Sample	Full	Full	Full	Full	±2	Full	Full
Left vote share polynomial	First	Second	Third	Fourth	None	Fourth	Fourth × time
Controls	No	No	No	No	No	Yes	Yes

Note: Standard errors clustered at the local government's term in office level are within parentheses. Each entry is a separate regression. All regressions also include, but do not report, municipality specific effects, time effects, and an indicator for undefined majority governments. The full sample includes 5,913 observations for government employment and 4520 for unemployment. The ±2 sample include all observations that are in the range of [48, 52] of the left vote share and there are 828 such observations for government employment and 603 for unemployment.

*Significant at 10%; **significant at 5%; ***significant at 1%.

No Dowsian Convergence in India

- Another implication of non-convergence is that the identity of the politician will matter.
- There is a subliterature investigating this issue with politicians 'gender or other characteristics.
- One example is Chattopathyay and Duflo (2004), looking at women brought to power at the panchayat level in India because of political reservations based on gender.

No Dowsian Convergence in India: Participation

TABLE III

EFFECT OF WOMEN'S RESERVATION ON WOMEN'S POLITICAL PARTICIPATION

Dependent Variables	Mean, Reserved GP (1)	Mean, Unreserved GP (2)	Difference (3)
West Bengal			
Fraction of Women Among Participants in the Gram Samsad (in percentage)	9.80 (1.33)	6.88 (.79)	2.92 (1.44)
Have Women Filed a Complaint to the GP in the Last 6 Months	.20 (.04)	.11 (.03)	.09 (.05)
Have Men Filed a Complaint to the GP in the Last 6 Months	.94 (.06)	1.00	.06 (.06)
Observations	54	107	
Rajasthan			
Fraction of Women Among Participants in the Gram Samsad (in percentage)	20.41 (2.42)	24.49 (3.05)	−4.08 (4.03)
Have Women Filed a Complaint to the GP in the Last 6 Months	.64 (.07)	.62 (.06)	.02 (.10)
Have Men Filed a Complaint to the GP in the Last 6 Months	.95 (.03)	.88 (.04)	.073 (.058)
Observations	40	60	

Notes: 1. Standard errors in parentheses. 2. Standard errors are corrected for clustering at the GP level in the West Bengal regressions, using the Moulton (1986) formula.

No Dowsian Convergence in India: Policy Preferences

TABLE IV

ISSUES RAISED BY WOMEN AND MEN IN THE LAST 6 MONTH

	West Bengal						Rajasthan					
	Women			Men	Average	Difference	Women			Men	Average	Difference
	Reserved	Unreserved	All				Reserved	Unreserved	All			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Other Programs												
Public Works	.84	.84	.84	.85	.84	−.01	.60	.64	.62	.87	.74	−.26
Welfare Programs	.12	.09	.10	.04	.07	.06	.25	.14	.19	.03	.04	.16
Child Care	.00	.02	.01	.01	.01	.00	.04	.09	.07	.01	.02	.06
Health	.03	.04	.04	.02	.03	.02	.06	.08	.07	.04	.03	.03
Credit or Employment	.01	.01	.01	.09	.05	−.08	.06	.06	.05	.04	.09	.01
Total Number of Issues	153	246	399	195			72	88	160	155		
Breakdown of Public Works Issues												
Drinking Water	.30	.31	.31	.17	.24	.13	.63	.48	.54	.43	.49	.09
Road Improvement	.30	.32	.31	.25	.28	.06	.09	.14	.13	.23	.18	−.11
Housing	.10	.11	.11	.05	.08	.05	.02	.04	.03	.04	.04	−.01
Electricity	.11	.07	.08	.10	.09	−.01	.02	.04	.03	.02	.02	.01
Irrigation and Ponds	.02	.04	.04	.20	.12	−.17	.02	.02	.02	.04	.03	−.02
Education	.07	.05	.06	.12	.09	−.06	.02	.07	.05	.13	.09	−.09
Adult Education	.01	.00	.00	.01	.00	.00	0	0	.00	.00	.00	.00
Other	.09	.11	.10	.09	.09	.01	.19	.21	.20	.12	.28	.05
Number of Public Works Issues	128	206	334	166			43	56	99	135		
Public Works												
Chi-square		8.84		71.72				7.48		16.38		
p-value		.64		.00				.68		.09		

Notes: 1. Each cell lists the number of times an issue was mentioned, divided by the total number of issues in each panel. 2. The data for men in West Bengal comes from a subsample of 48 villages. 3. Chi-square values placed across two columns test the hypothesis that issues come from the same distribution in the two columns.

No Dowsian Convergence in India: Outcomes

TABLE V

EFFECT OF WOMEN'S RESERVATION ON PUBLIC GOODS INVESTMENTS

	West Bengal			Rajasthan		
Dependent Variables	Mean, Reserved GP (1)	Mean, Unreserved GP (2)	Difference (3)	Mean, Reserved GP (4)	Mean, Unreserved GP (5)	Difference (6)
A. *Village Level*						
Number of Drinking Water Facilities	23.83	14.74	9.09	7.31	4.69	2.62
Newly Built or Repaired	(5.00)	(1.44)	(4.02)	(.93)	(.44)	(.95)
Condition of Roads (1 if in good	.41	.23	.18	.90	.98	−.08
condition)	(.05)	(.03)	(.06)	(.05)	(.02)	(.04)
Number of Panchayat Run	.06	.12	−.06			
Education Centers	(.02)	(.03)	(.04)			
Number of Irrigation Facilities	3.01	3.39	−.38	.88	.90	−.02
Newly Built or Repaired	(.79)	(.8)	(1.26)	(.05)	(.04)	(.06)
Other Public Goods (ponds, biogas,	1.66	1.34	.32	.19	.14	.05
sanitation, community buildings)	(.49)	(.23)	(.48)	(.07)	(.06)	(.09)
Test Statistics: Difference Jointly Significant			4.15			2.88
(*p*-value)			(.001)			(.02)
B. *GP Level*						
1 if a New Tubewell Was Built	1.00	.93	.07			
		(.02)	(.03)			
1 if a Metal Road Was Built or Repaired	.67	.48	.19			
	(.06)	(.05)	(.08)			
1 if There Is an Informal Education	.67	.82	−.16			
Center in the GP	(.06)	(.04)	(.07)			
1 if at Least One Irrigation Pump Was Built	.17	.09	.07			
	(.05)	(.03)	(.05)			
Test Statistics: Difference Jointly Significant			4.73			
(*p*-value)			(.001)			

Notes: 1. Standard errors in parentheses. 2. In West Bengal, there are 322 observations in the village level regressions, and 161 in the GP level regressions. There are 100 observations in the Rajasthan regressions. 3. Standard errors are corrected for clustering at the GP level in the village level regressions, using the Moulton (1986) formula, for the West Bengal regressions.

Interpretation

- Overall, the evidence is fairly clear that at least the strong form of Dowsian policy convergence doesn't hold (reality check, think of the US at the moment).
- But how do we make sense of this?
- So let's think about theory again.

Non-Convergence in Theory

- One possibility, which is not unrealistic even if it's not exciting theoretically, is that parties are unable to make binding commitments to policies.
- If so, then voters will choose candidates based on what they expect they will do once in office.
- This is a perspective adopted in "citizen-candidate" type models, such as Osborne and Slivinski (1996) or Besley and Coate (1997), whose main focus is the modeling of entry decisions of candidates.

What Happens with Policy-Motivated Politicians?

- Instead, suppose that parties/politicians can commit to policies, but have policy preferences.
- For example, one party may prefer right-wing policies the other one left-wing ones.
- What happens in this case?

Theorem

Suppose we are in the baseline model with single-peaked or single-crossing preferences, and the two parties have their own policy platforms, one to the left of the median the other one to the right of the median. The unique equilibrium is Dowsian policy convergence.

- Why?

Add Frictions

- The previous result is no longer true if there are "frictions".
- The most obvious friction is idiosyncratic party preferences as in the probabilistic voting model. In this case, each politician has a captured audience, and will be able to push his policies (at least a little bit) in the direction of his bliss point. (What is a simple proof of this?)
- Another friction might be preventing certain types of parties from entering. For example, parties representing the interests of certain ethnic groups or worker groups are banned from elections in many countries (e.g., Turkey, Burma).
- If we have that both parties are to the right of the median voter, then the above theorem doesn't work.

Policy Responsiveness

- The lack of Dowsian policy convergence does not imply that policies are not, on average, responsive to voter preferences.
- At some level whether this is the case or not is much more important.
- Investigating this issue is made complicated by the fact that we don't generally know what voters want. But there is one setting in which we infer changes in voter preferences — de jure or de facto changes in the voting franchise.
- In contrast to comparative statics with respect to inequality, which we saw not to be robust in the second lecture, comparative statics with respect to changes in the voting franchise are fairly straightforward.

Democracy and Redistribution

- Consider a model similar to that discussed in the second lecture, where each individual has income y_i and the only fiscal tools are a linear income tax and lump-sum redistribution.
- As a result, the most preferred tax rate of a richer individual is lower than that of a poorer individual (holding everything else including the distribution of income constant).
- Suppose that individuals are ranked according to income, and only those above the qth percentile are enfranchised. An extension of the franchise — a democratization — is a decline in this percentile.

Theorem

Consider an extension of the franchise. This always increases taxes and redistribution.

Measuring Democracy

- To test this prediction, we need to measure of democracy or democratizations. This is in general tricky.
- Acemoglu, Naidu, Restrepo and Robinson (2014, 2017) developed a binary index based on several sources.
- Using this annual measure of democracy, they investigate the effects of democratizations on taxes, revenues and inequality.
- there are several econometric issues one has to be careful about (serial correlation, endogeneity, Nickell bias, etc.)
- All the same, for our focus here, the robust result is that democratizations leads to higher taxes and government revenues.

Democracy and Taxes

Courtesy of Daron Acemoglu, Suresh Naidu, Pascual Restrepo, and James A. Robinson. Used with permission.

Table 2: Effects of democratization on the log of tax revenue as a percentage of GDP.

						Assuming AR(1) coefficient				
			GMM			$\rho = 0$	$\rho = 0.25$	$\rho = 0.5$	$\rho = 0.75$	$\rho = 1$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Democracy lagged	15.00***	11.71***	11.27	18.68**	14.63**	15.00***	11.92***	8.84***	5.77**	2.69
	(4.33)	(3.38)	(7.23)	(8.78)	(5.98)	(4.33)	(3.27)	(2.55)	(2.48)	(3.11)
Dep. Var lagged		0.27***	0.27***	0.29***	0.33***					
		(0.06)	(0.10)	(0.07)	(0.08)					
Observations	944	944	816	816	816	944	944	944	944	944
Countries	128	128	125	125	125	128	128	128	128	128
Numer of moments			81	61	61					
Hansen p-value			0.12	0.05	0.06					
AR2 p-value			0.92	0.83	0.78					
Democracy changes in the sample	92	92	82	82	82	92	92	92	92	92
Long run effect of democracy	15.00	15.97	15.49	26.35	21.97	15.00	15.89	17.68	23.06	.
P-value for the long run effect	0.00	0.00	0.11	0.03	0.01	0.00	0.00	0.00	0.02	.

Note:- Dependent variable: log of tax revenue as a percentage of GDP. OLS estimates (Columns 1-2) include a full set of country and year fixed effects. Arellano and Bond's GMM estimators of the dynamic panel model (Columns 3-4) remove country fixed effects by taking first differences of the data, or by taking forward orthogonal differences (Column 5) and then construct moment conditions using predetermined lags of the dependent variable and democracy as instruments. Columns 4 and 5 use only up to the fifth lag of predetermined variables to create moments. Columns 6-10 impose different values for the autocorrelation coefficient in the dependent variable series, and estimates the effect of democracy including a full set of country and year fixed effects. All models control for lagged GDP per capita but this coefficient is not reported to save space. Robust standard errors, adjusted for clustering at the country level, are in parentheses. For the GMM models, significance levels for the Hansen J-test and test for lack of second-order serial correlation in residuals are reported at the bottom.

- NB: GDP is controlled for on the RHS, so these are effects on taxes.

Democracy and Inequality

- But no effect on inequality:

Table 6: Effects of democratization on inequality.

						Assuming AR(1) coefficient				
			GMM			$\rho = 0$	$\rho = 0.25$	$\rho = 0.5$	$\rho = 0.75$	$\rho = 1$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Dependent variable: Gini coefficient, net income.									
Democracy lagged	0.62	-0.74	-2.01	-2.60	-1.60	-0.42	-0.67	-0.92	-1.17	-1.42
	(0.78)	(0.88)	(1.59)	(1.63)	(1.51)	(0.93)	(0.89)	(0.89)	(0.93)	(1.00)
Dep. Var lagged		0.32***	0.35***	0.39***	0.32***					
		(0.07)	(0.10)	(0.12)	(0.12)					
Observations	657	537	420	420	424	537	537	537	537	537
Countries	127	113	100	100	100	113	113	113	113	113
Numer of moments			81	61	61					
Hansen p-value			0.60	0.69	0.30					
AR2 p-value			0.02	0.03	0.01					
Democracy changes	65	47	31	31	31	47	47	47	47	47
Long run effect	0.62	-1.10	-3.12	-4.28	-2.36	-0.42	-0.90	-1.84	-4.67	.
P-value	0.43	0.40	0.21	0.12	0.30	0.65	0.45	0.31	0.21	.
	Dependent variable: Gini coefficient, gross income.									
Democracy lagged	-1.22	-1.50	-1.45	-1.88	-1.22	-1.51	-1.50	-1.50*	-1.49*	-1.49
	(0.99)	(0.90)	(1.44)	(1.59)	(1.27)	(1.15)	(1.00)	(0.90)	(0.87)	(0.92)
Dep. Var lagged		0.50***	0.64***	0.64***	0.76***					
		(0.06)	(0.11)	(0.11)	(0.11)					
Observations	657	537	420	420	424	537	537	537	537	537
Countries	127	113	100	100	100	113	113	113	113	113
Numer of moments			81	61	61					
Hansen p-value			0.54	0.29	0.37					
AR2 p-value			0.59	0.57	0.48					
Democracy changes	65	47	31	31	31	47	47	47	47	47
Long run effect	-1.22	-2.98	-3.99	-5.26	-5.15	-1.51	-2.00	-3.00	-5.97	.
P-value	0.22	0.11	0.36	0.30	0.42	0.19	0.14	0.10	0.09	.

Note:- Dependent variables: Gini coefficient, net income (top panel) and gross income (bottom panel). OLS estimates (Columns 1-2) include a full set of country and year fixed effects. Arellano and Bond's GMM estimators of the dynamic panel model (Columns 3-4) remove country fixed effects by taking first differences of the data, or by taking forward orthogonal differences (Column 5) and then construct moment conditions using predetermined lags of the dependent variable and democracy as instruments. Columns 4 and 5 use only up to the fifth lag of predetermined variables to create moments. Columns 6-10 impose different values for the autocorrelation coefficient in the dependent variable series, and estimates the effect of democracy including a full set of country and year fixed effects. All models control for lagged GDP per capita but this coefficient is not reported to save space. Robust standard errors, adjusted for clustering at the country level, are in parentheses. For the GMM models, significance levels for the Hansen J-test and test for lack of second-order serial correlation in residuals are reported at the bottom.

Courtesy of Daron Acemoglu, Suresh Naidu, Pascual Restrepo, and James A. Robinson. Used with permission.

Why Democracy May Not Impact Inequality?

- There are several possibilities:
  - Democracy is captured and is not responsive. But if so, why are taxes going up?
  - Democracy is responsive to the middle class, and the middle class may want lower redistribution towards the poor when the poor are added to the franchise.
  - Democratizations may change the structure of the economy, creating more inequality-generating opportunities (e.g., the fall of apartheid in South Africa).
- Acemoglu, Naidu, Restrepo and Robinson provide evidence consistent with the second and third channels. But nothing definitive.

Women's Enfranchisement

- Similar issues come into action when those being enfranchised aren't the poor but women.
- Miller (2008) looks at this in the context of the United States — US states enfranchised women between 1869 and 1920.
- He finds greater municipality based on spending following women's enfranchisement and significant impacts on one of the issues about which women care — child survival/mortality.

The Effects of Women's Enfranchisement: Spending

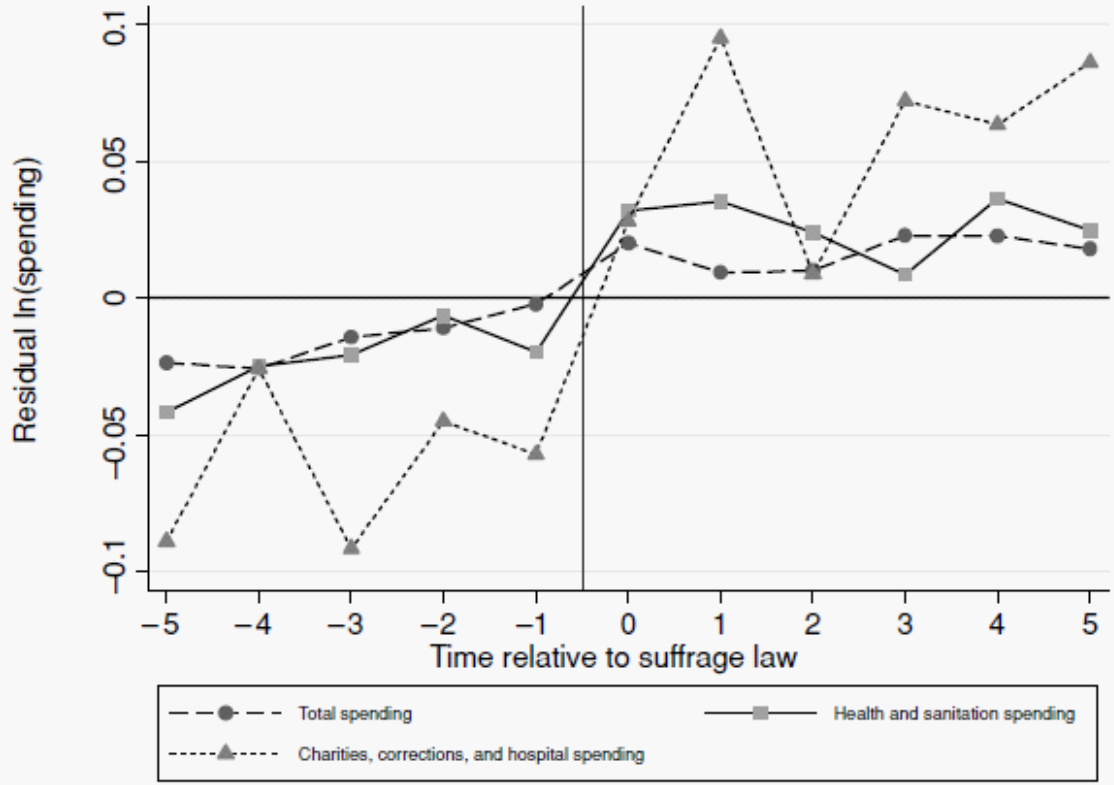


FIGURE II

Municipal Public Spending and Women's Suffrage Law Timing

Municipal public finance data from the U.S. Bureau of the Census's *Statistics of Cities Having a Population of Over 30,000* and *Financial Statistics of Cities Having a Population of Over 30,000*. Residual means shown relative to the year of women's suffrage laws in each state (year 0) obtained by estimating equation (1) without the suffrage dummy variable and with city rather than state fixed effects.

The Effects of Women's Enfranchisement: Public Finances

TABLE II
WOMEN'S SUFFRAGE LAWS AND MUNICIPAL AND STATE PUBLIC FINANCE

Dependent variable	Estimate (standard error)	N	R^2
Panel A: Municipal public finance			
ln(total spending)	0.079***	3,661	0.97
	(0.029)		
ln(health conservation and sanitation spending)	0.061*	3,661	0.94
	(0.036)		
ln(charities, hospitals, and corrections spending)	0.360***	3,454	0.92
	(0.105)		
ln(total infrastructure investment)	0.012	3,658	0.85
	(0.086)		
ln(health conservation and sanitation infrastructure investment)	0.152	3,629	0.70
	(0.114)		
ln(charities, hospitals, and corrections infrastructure investment)	0.580**	1,462	0.71
	(0.276)		
Panel B: State public finance			
ln(total revenue)	0.010	673	0.89
	(0.084)		
ln(property tax revenue)	0.070	579	0.94
	(0.209)		
ln(total spending)	−0.057	688	0.87
	(0.088)		
ln(highway spending)	0.300	667	0.90
	(0.215)		
ln(education spending)	0.137	689	0.75
	(0.157)		
ln(social service spending)	0.206***	688	0.84
	(0.071)		

The Effects of Women's Enfranchisement: Infant Health

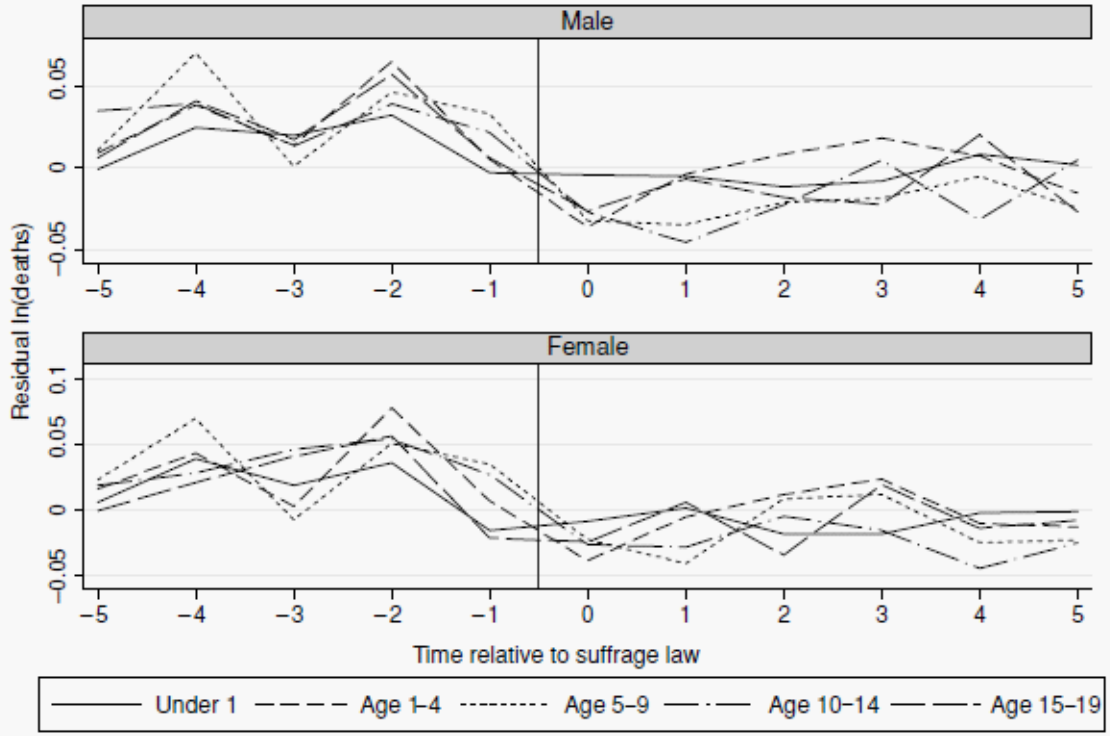


FIGURE IV
Deaths by Age and Sex and the Timing of Suffrage Laws
Mortality data from the U.S. Bureau of the Census's annual *Mortality Statistics*. Residual means shown relative to the year of women's suffrage laws in each state (year 0) obtained by estimating equation (1) without the suffrage dummy variable.

Though some pre-trends perhaps.

De Facto Enfranchisement

- Certain voters may be disenfranchised because of practices or their inability to effectively express their voice.
- Fujiwara (2015) investigates a setting, whereby the introduction of new technology (electronic voting) enables previously de facto disenfranchised low-education voters to increase their voting and their influence.
- In Brazil, before this voting technology, the complicated nature in which voting would have to take place meant that the ballots of a large fraction of low-education, poor voters were spoilt. (25% of adults at the time were unable to read or write a simple note).
- This changed with the introduction of electronic voting in the mid-1990s, and did so above a threshold, enabling a regression discontinuity design.
- One expectation might be that these low-education voters would be ineffective voters even after the change in technology.
- This is not what Fujiwara finds.

No Differences in Turnout

- Threshold for electronic voting in 1998 was 40,000.

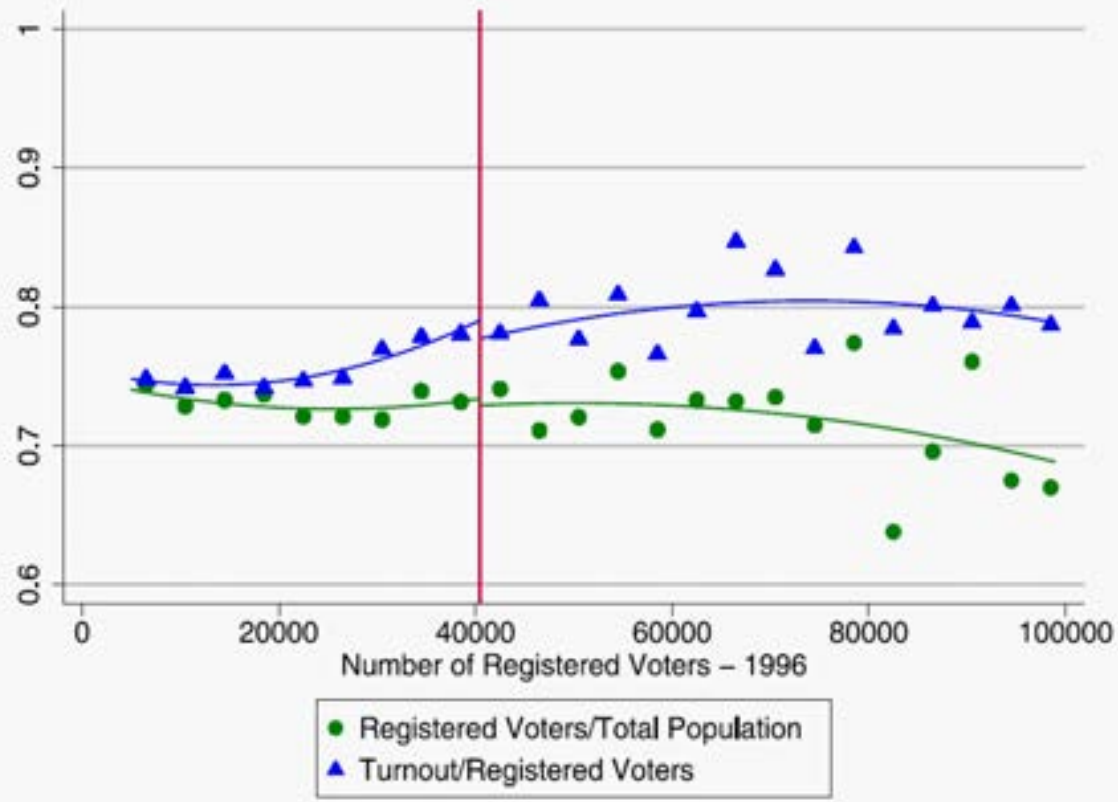


FIGURE 3.—Registration and turnout—local averages and parametric fit.

Change in Votes

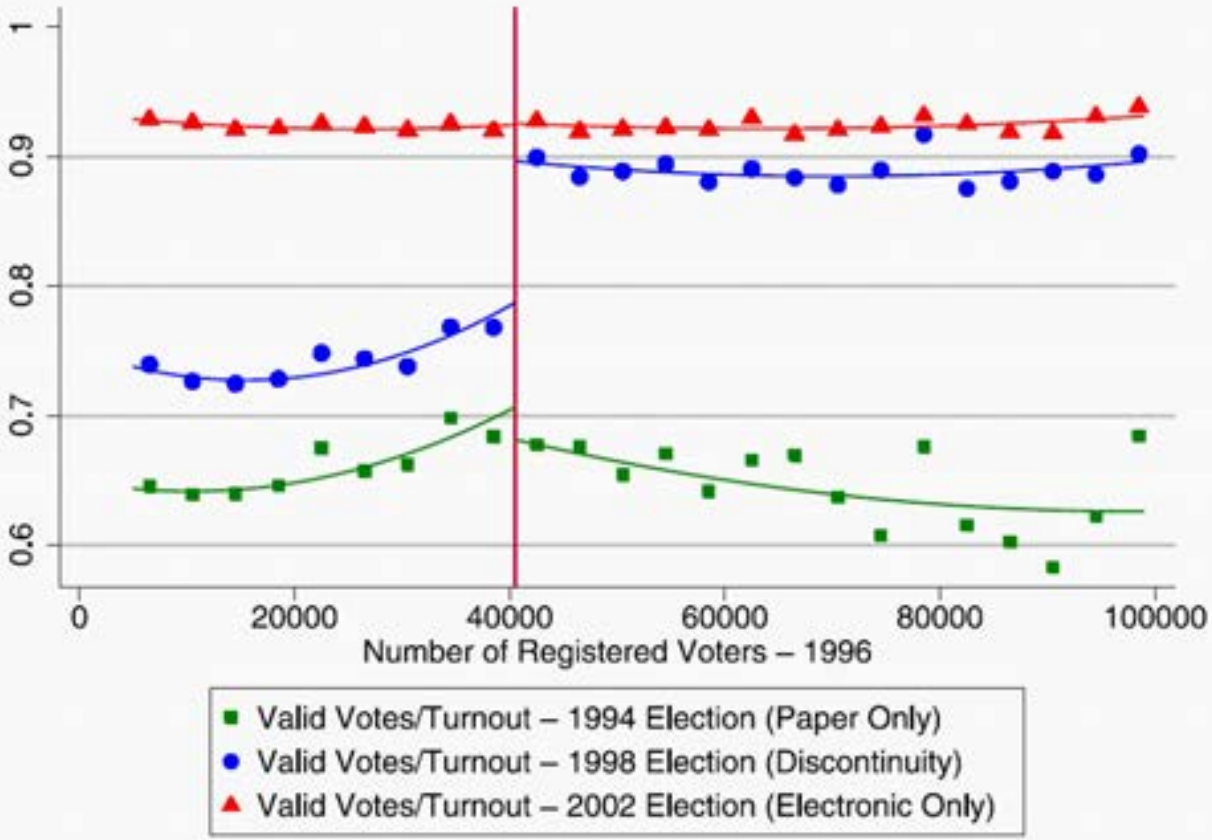


FIGURE 2.—Valid votes/turnout—local averages and parametric fit. Each marker represents the average value of the variable in a 4000-voter bin. The continuous lines are from a quadratic fit over the original ("unbinned") data. The vertical line marks the 40,500-voter threshold.

Where Do the Votes Go?

TABLE II

TREATMENT EFFECTS OF ELECTRONIC VOTING[a]

	Full Sample Mean	Pre-Treat. Mean	IKBW {Obs.}	(1)	(2)	(3)
Panel A: Baseline Results						
Valid Votes/Turnout (1998 Election)	0.755 [0.087]	0.780 (0.013)	11,873 {265}	0.118 (0.015)	0.121 (0.016)	0.124 (0.025)
Turnout/Reg. Voters (1998 Election)	0.765 [0.091]	0.785 (0.011)	12,438 {283}	−0.005 (0.019)	0.013 (0.021)	0.007 (0.033)
Reg. Voters/Population (1998 Election)	0.748 [0.141]	0.737 (0.010)	15,956 {388}	−0.004 (0.027)	0.010 (0.034)	0.032 (0.044)
Panel B: Placebo Tests (Election Years Without Discontinuous Assignment)						
Valid Votes/Turnout (1994 Election)	0.653 [0.099]	0.697 (0.011)	17,111 {433}	−0.013 (0.019)	−0.008 (0.023)	0.006 (0.032)
Valid Votes/Turnout (2002 Election)	0.928 [0.026]	0.921 (0.002)	17,204 {437}	0.005 (0.005)	0.008 (0.006)	0.009 (0.010)
Panel C: Do Left-Wing Parties Benefit Disproportionately From Electronic Voting?						
Vote-Weighted Party Ideology (1998 Elec.)	5.397 [0.692]	5.162 (0.094)	20,000 {558}	−0.222 (0.100)	−0.250 (0.081)	−0.108 (0.170)
Bandwidth				IKBW	10,000	5000
Specification				Linear	Linear	Linear
N	5281			—	229	116

[a]Robust standard errors in parentheses, standard deviations in brackets, number of observations in curly brackets—{}. The unit of observation is a municipality. Each figure in columns (1)–(3) is from a separate local linear regression estimate with the specified bandwidth. The pre-treatment mean is the estimated value of the dependent variable for a municipality with 40,500 registered voters that uses paper ballot (based on the specification on column (1)). The IKBW column provides the Imbens and Kalyanaraman (2012) optimal bandwidth (capped at 20,000) and the associated number of observations. Details on the dependent variables in the text.

The Illiterate Benefited

TABLE III

TREATMENT EFFECTS OF ELECTRONIC VOTING, BY ILLITERACY RATE[a]

	Pre-Treat. Mean	IKBW (Obs.)	(1)	(2)	(3)	(4)
Panel A: Municipalities With Above-Median Illiteracy						
Valid Votes/Turnout	0.759	11,873	0.147	0.150	0.152	0.176
	(0.017)		(0.019)	(0.015)	(0.020)	(0.031)
N	—	—	116	279	103	49
Panel B: Municipalities With Below-Median Illiteracy						
Valid Votes/Turnout	0.799	11,873	0.092	0.113	0.096	0.089
	(0.018)		(0.020)	(0.016)	(0.022)	(0.032)
N	—	—	149	279	126	67
Test of Equality in TEs (*p*-Value)	—	—	0.049	0.090	0.056	0.054
Bandwidth	—	—	IKBW	20,000	10,000	5000

[a]Robust standard errors in parentheses, standard deviations in brackets. The unit of observation is a municipality. Each figure in columns (1)–(4) is from a separate local linear regression estimate with the specified bandwidth. The pre-treatment mean is the estimated value of the dependent variable for a municipality with 40,500 registered voters that uses paper ballot (based on the specification on column (1)). The IKBW column provides the Imbens and Kalyanaraman (2012) optimal bandwidth. Details on the dependent variables in the text. Estimates on Panel A (Panel B) use only municipalities where the adult illiteracy rate is above (below) 25.43%.

State-Level Results

TABLE IV

MAIN OUTCOMES AND THE SIGN-SWITCH PATTERN[a]

Parameter:		θ^{98}	θ^{02}	Linear Combinations	
Sample (Terms):		1994–1998 (Paper–Disc.)	1998–2002 (Disc.–Electr.)	$(\theta^{98}-\theta^{02})/2$	$(\theta^{98}+\theta^{02})/2$
	Sample Avg.	(1)	(2)	(3)	(4)
Panel A: Electoral Outcomes					
Valid Votes/Turnout	0.829	0.092	−0.111	0.102	−0.009
	[0.112]	(0.033)	(0.010)	(0.017)	(0.018)
		{0.102}	{0.002}	{0.008}	{0.630}
Seat-Weighted Policy Position	4.623	−0.112	0.299	−0.206	0.094
	[0.601]	(0.641)	(0.167)	(0.350)	(0.302)
		{0.842}	{0.154}	{0.574}	{0.800}
Panel B: Fiscal Outcomes (Health Care Spending)					
log(Total Spending)	—	−0.004	−0.257	0.127	−0.131
		(0.093)	(0.156)	(0.097)	(0.082)
		{0.946}	{0.274}	{0.254}	{0.228}
Share of Spending in Health Care	0.099	0.039	−0.029	0.034	0.005
	[0.037]	(0.017)	(0.013)	(0.008)	(0.013)
		{0.104}	{0.044}	{0.000}	{0.678}
log(Health Spending p.c.)	—	0.428	−0.677	0.552	−0.125
		(0.264)	(0.262)	(0.096)	(0.242)
		{0.200}	{0.034}	{0.000}	{0.628}
Panel C: Birth Outcomes (Mothers Without Primary Schooling)					
Share With 7+ Visits	0.362	0.122	−0.023	0.069	0.047
	[0.123]	(0.065)	(0.033)	(0.040)	(0.039)
		{0.154}	{0.558}	{0.182}	{0.320}
Share With Low-Weight Births (×100)	7.721	−0.370	0.528	−0.529	0.201
	[1.110]	(0.304)	(0.269)	(0.246)	(0.236)
		{0.266}	{0.104}	{0.044}	{0.450}
N (State-Terms)	—	54	54	—	—
N (States/First-Diffs)	—	27	27	—	—

[a]Standard errors clustered at the state level in parentheses. Standard deviations in brackets. *p*-values based on Cameron, Gelbach, and Miller (2008) cluster-robust wild-bootstrap in curly brackets—{}. The unit of observation is a state-electoral term. Each row reports the estimation of equations (6) and (7) using the specified dependent variable. Each figure in columns (1) and (2) is from a separate regression, providing the coefficient on the share of voters living above the cutoff (S_i) on the 1998 and 2002 first-differences, respectively (θ^{98} and θ^{02}). Columns (3) and (4) report the specified linear combination of these parameters. Region-time effects are included.

Summary

- Though there are much more complicated patterns in practice, the evidence is consistent with the idea that voters to vote in line with (some) of their interests, and policies to change in response to major changes such as the enfranchisements
- But at the same time, we are far from the Dowsian framework.

14.770 Introduction to Political Economy

Fall 2017