PH336

Voters and Elections

Lecture Plan

1. The “paradox of voting”
2. The Downsian model of electoral competition
3. The Citizen-Candidate model of electoral competition
4. Evidence from close elections

The Public Choice Perspective on Voting

- If the citizen is rational, votes iff p.B>C, where:
  - C is the cost of voting (becoming informed, going to the polling station, etc.)
  - p is the probability of being able to change the outcome of the election (probability of *being pivotal*)
  - B is the benefit from changing the outcome, given that the voter is pivotal

- Paradox of Voting; for most elections, (e.g. US presidential, UK general), p is very, very small, while C is clearly positive, so why does anybody vote?

The Paradox of Voting: Possible Solutions

1. *Rational choice explanation*. p depends on other voters' behaviour. If no-one else votes, then no matter how large the electorate, you will be pivotal with probability 1, so you have an incentive to vote! When the details are worked out, the Nash equilibrium involves some members of the electorate voting with positive probability. Turnout can be high in such equilibria (see Aldrich(1993)).

2. *Voting as an expressive act*. Citizens are motivated to vote because they see it as a duty, to express support for their party, etc. Can be accommodated in the calculus of voting by adding an additional "payoff" D>0 i.e. the citizen votes iff p.B+D-C>0.

The Paradox of Voting: Possible Solutions

3. “behavioural economics” explanation: *minimax regret theory* can explain voting. If you do vote, the worst that can happen is that your vote was wasted, i.e. you lose C by voting. If you do not vote, the worst that can happen is that if you had voted, you would have been pivotal. In this case, by not voting, you lose B-C. So, if you act to minimize the maximum loss from a course of action, you will vote if C< B-C or C<B/2.

- In reality, probably a combination of all of these
- Public choice theory does not claim to explain the level of turnout, but how it varies with the probability of being pivotal: turnout should be higher when the probability of being pivotal is higher

Testing the Public Choice theory of Voting

- The public choice theory of voting predicts that the voter is more likely to be vote, and thus in the aggregate, turnout will be higher, the higher is p, the probability of being pivotal.

- Usual measure of turnout : number of voters/(population of voting age) or number of voters/(population eligible to vote)

- Usual measure of p: p=1-percentage vote gap, where the percentage vote gap =lead of the winner over the runner up, as fraction of the total votes cast. (e.g. if leader has 52%, runner up 40%, then p= 1-(0.52-0.40)=0.88.)

- Researchers look for:

  - a positive relationship between p and turnout at the level of the electoral district

  - A positive effect of p on the probability that any individual votes in the election

Testing the Public Choice theory of Voting: Pattie and Johnston

- Results for UK general elections show at the aggregate (constituency) level, p and turnout are positively correlated (Pattie and Johnston(1998)).

Year Of election	1959	1964	1966	1970	1974 (Oct)	1979	1987	1992
Correlation	0.33	0.23	0.46	0.44	0.48	0.51	0.18	0.18

- This positive link remains after other constituency-level variables affecting turnout are allowed for via multiple regression (Pattie and Johnston(1998)).

- However, when individual level data on turnout from the British Electoral Survey is studied, the link between p and turnout is positive but no longer significant

Testing the Public Choice theory of Voting: Surveys

- Gey's (2006) survey of aggregate (electoral district level) studies finds that out of 362 tests in 52 studies, 206 find p positively related to turnout, 19 find it negatively related, and 137 find no effect.

- But Smets and Van Ham's (2013) survey of 90 individual voter level studies finds:

  - 36 studies meaured p by the closeness of the election at the national level, and p had a positive effect on the probability of voting in 15 studies, and no effect in the remanider.

  - 13 studies meaured p by the closeness of the election at the electoral district level, but p was insignificant in all cases.

The Median Voter

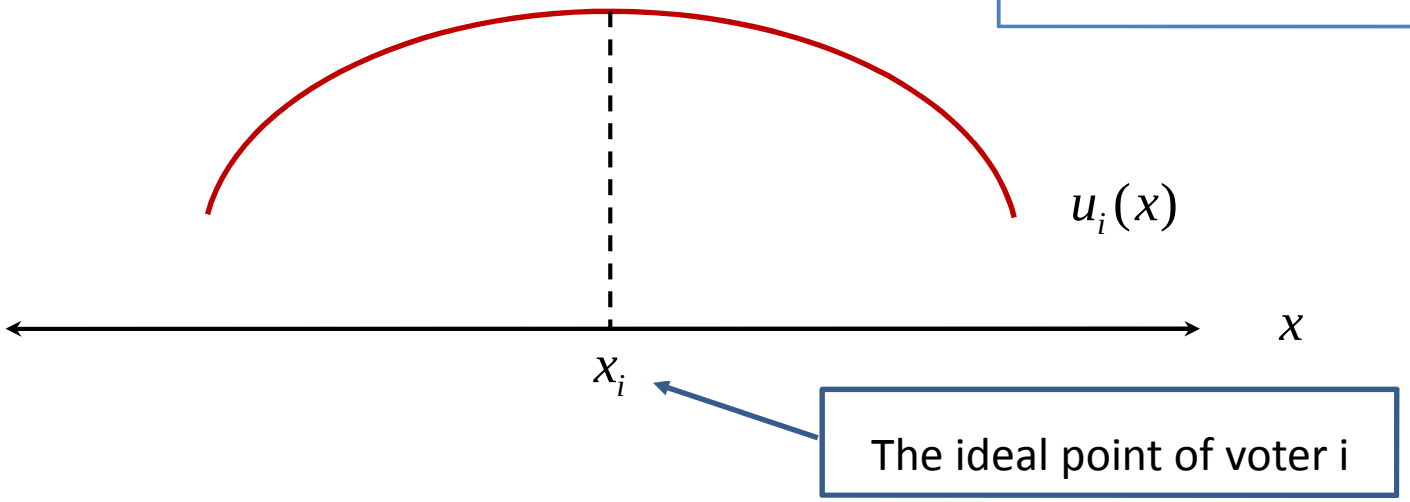
- Assume an odd number of voters i=1,..n, and a one-dimensional set of alternatives x in X (the real line or a subset of it)
- Every voter has preferences over X represented by a utility function:

$$u_i(x), \quad x \in X$$

- The utility function is assumed *single-peaked*:

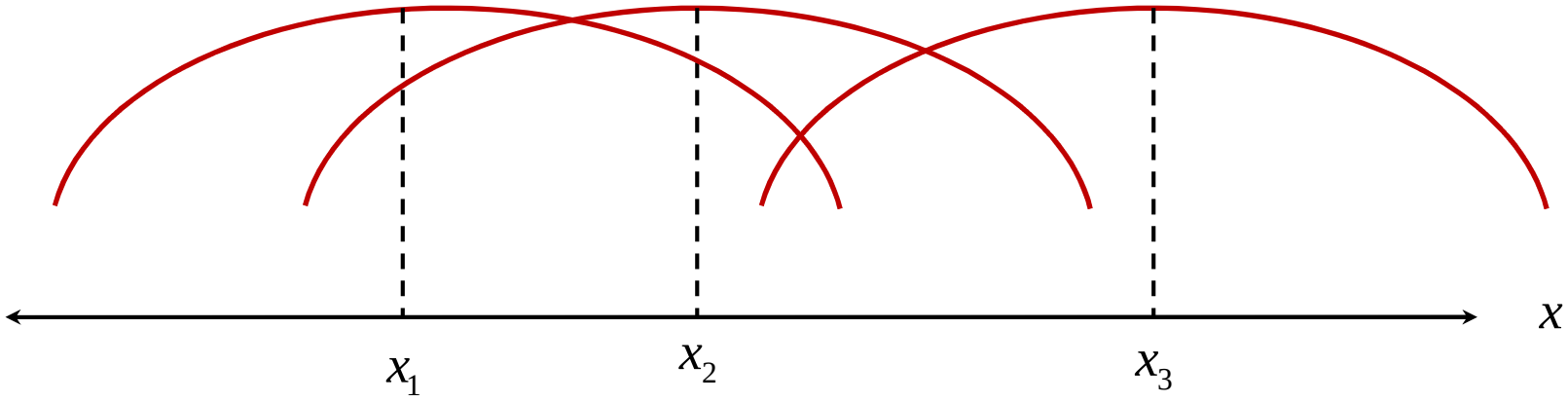
For example,

$$u_i(x) = -(x - x_i)^2$$



The Median Voter

- Assume that all the ideal points are different, and order the voters so that $x_1<x_2<..x_n$
- Then, voter m=(n+1)/2 is the *median voter*
- The median voter a key concept in the theory of voting



Down's Model of Political Competition

- An odd number of voters i=1,..n, and a one-dimensional set of alternatives X.
- Every voter has single-peaked preferences on X.
- Two parties, A and B. Parties have no policy preferences, but each gets a payoff E>0 from holding office (ego-rent) : so, they both wish to maximize the chance of winning.

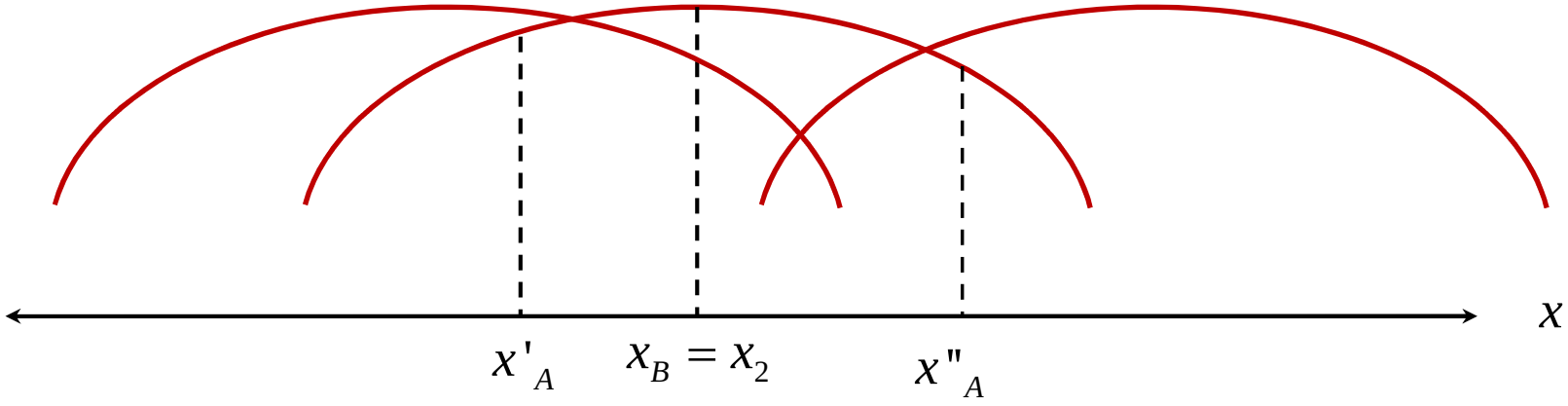
Order of events:

1. *Both parties simultaneously choose policies x_A, x_B in X:*
2. *Voters all simultaneously vote for A or B. The party with the most votes wins the election (plurality rule)*
3. *The winning party w implements its policy x_W chosen at stage 1 (precommitment to policies)*

[Tie-breaking rule: if both parties get the same numbers of votes, then each wins with probability ½: implies that if $x_A=x_B$, both win with probability 1/2.]

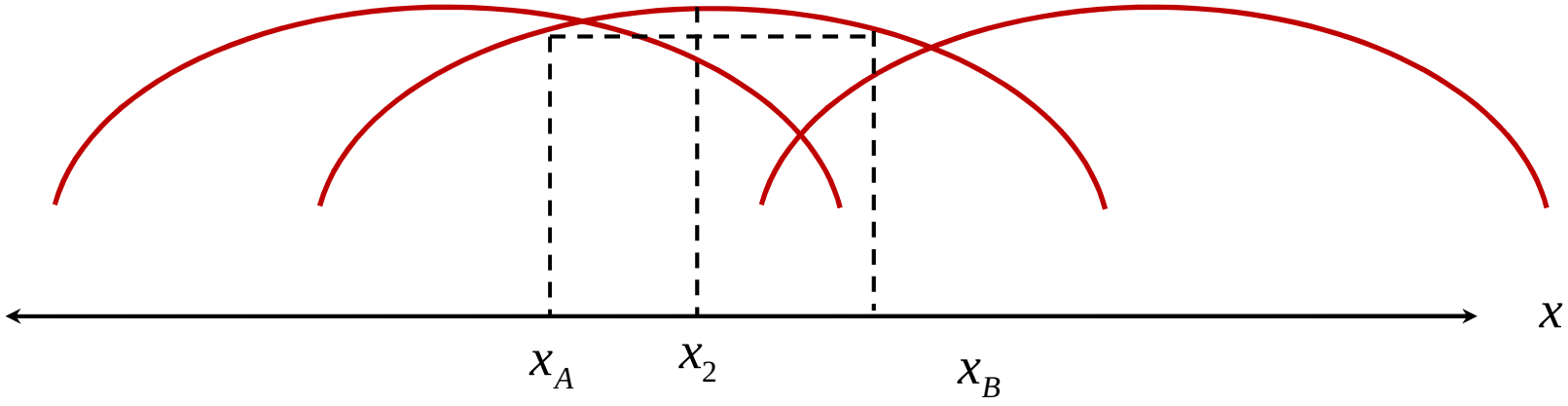
Equilibrium in the Downsian Model

- The only Nash equilibrium in the Downsian model is where $x_A=x_B=x_2$.
- To check that it is an equilibrium: if $x_B=x_2$, then party A will win with probability 0.5 if it sets $x_A=x_2$, and lose otherwise (and vice versa) i.e. $x_A=x_B=x_2$ are mutual best responses

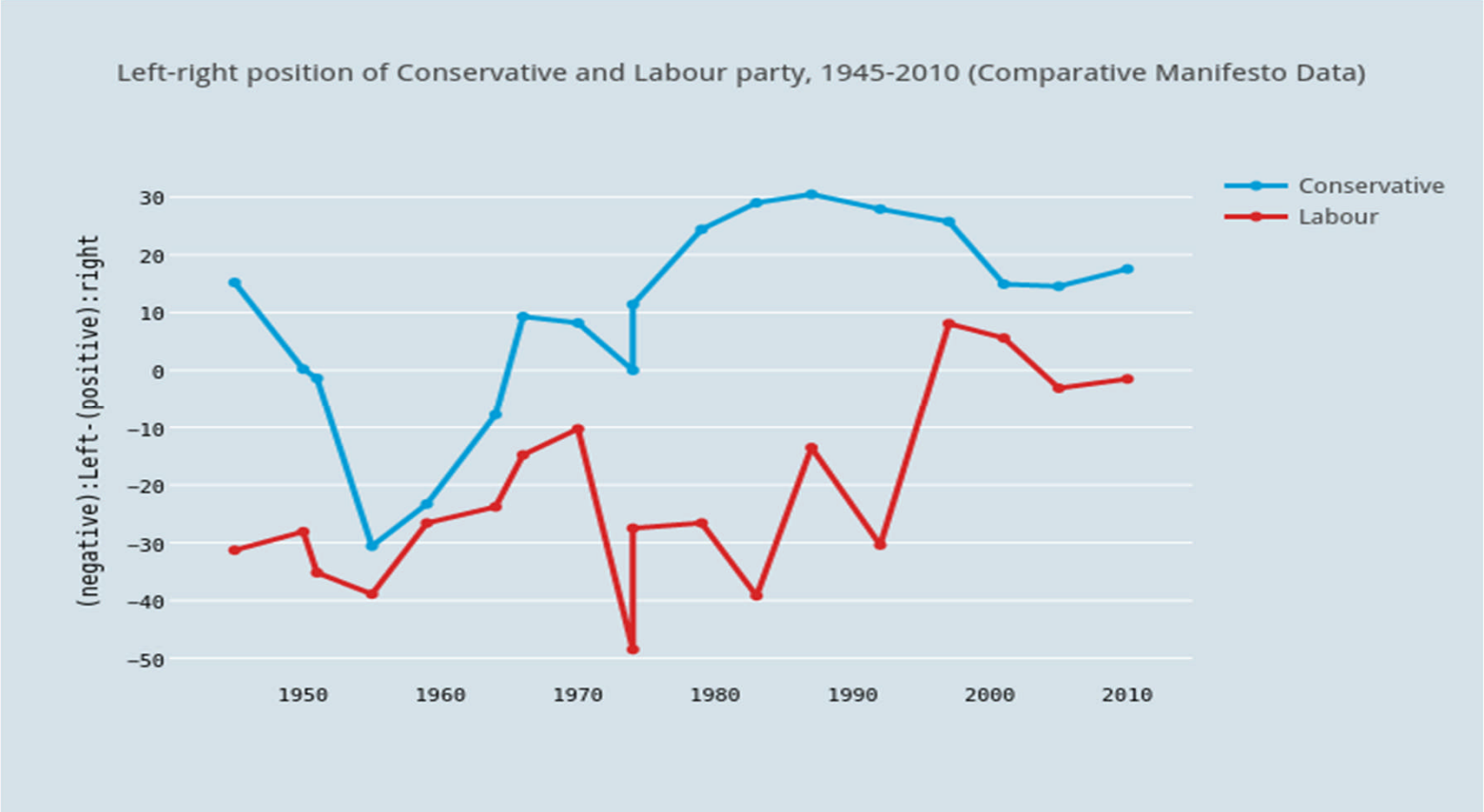


Equilibrium in the Downsian Model

- To check that there is no other Nash equilibrium, suppose the contrary i.e. There is an equilibrium $x_A, x_B \neq x_2$.
- Then, one of the parties can always strictly increase his chance of winning by moving to x_2
- As shown, the median voter 2 will vote for A at platforms x_A, x_B ,so in this case, party B loses the election and can win by moving to x_2.

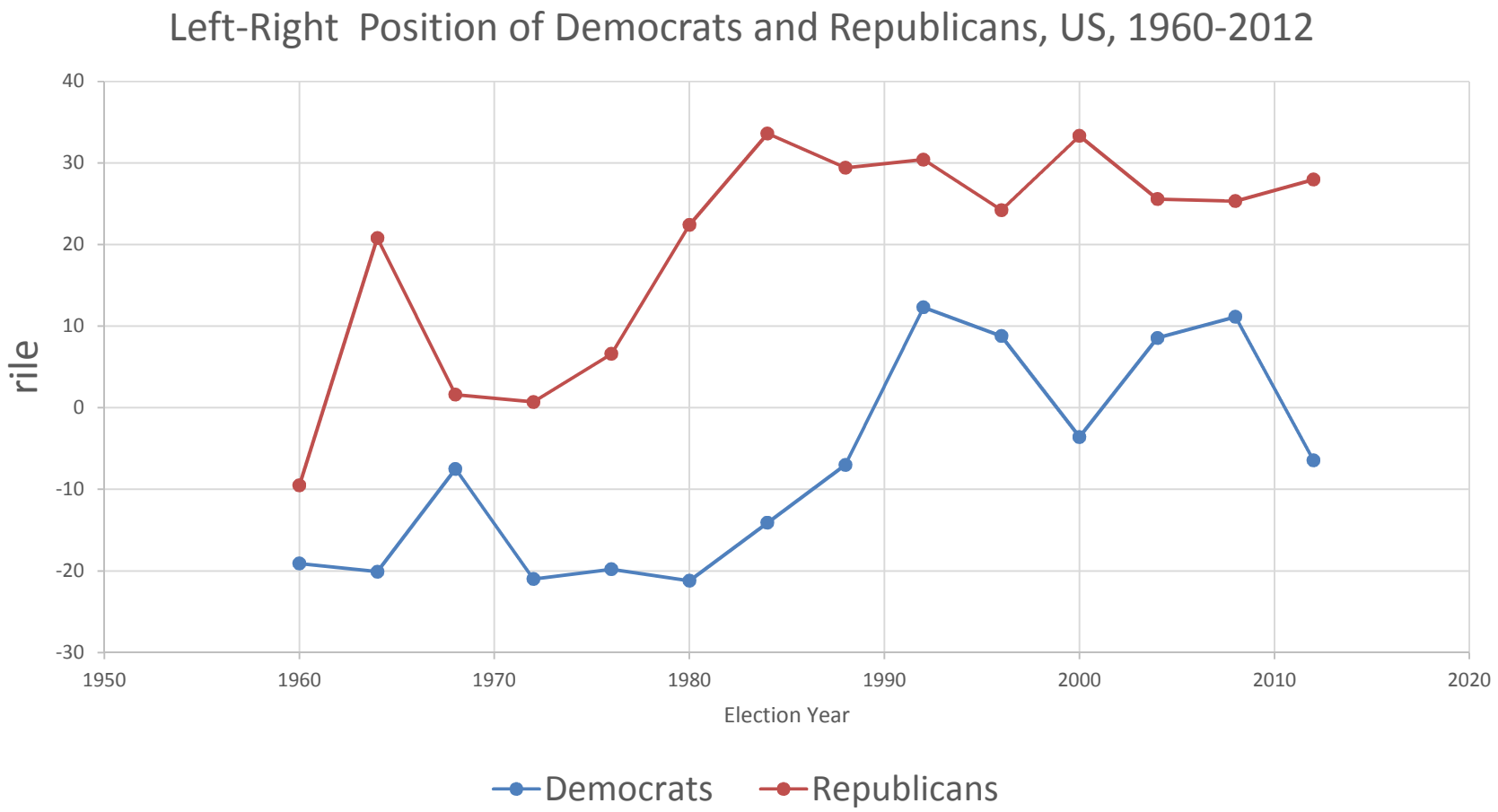


Platform Divergence in Practice



Source: Comparative Manifesto Project (variable =rile): manifestoproject.wzb.eu

Platform Divergence in Practice



Explaining Policy Divergence

- Large literature on extensions to Downs' basic model: see Grofman (2004) for a survey.
  - The convergence result also occurs with: alternative objectives for the parties (vote share, policy preferences), multiple electoral districts
  - The convergence result may not occur with more than two parties, but this gets a bit technical: see Cox, "Centripetal and centrifugal incentives in electoral systems", *American Journal of Political Science* (1990): 903-935 for a survey
  - See Appendix to these notes for more

- Introduce asymmetric information; i.e. the parties are not sure where the median voter is located, and combine with policy preferences for parties:
  - gives an interesting theory of the sources of policy divergence: known as the Calvert-Wittman model

Uncertainty about the Position of the Median Voter

- Extend Downs' model to allow for:
  - *policy preferences* by two parties i.e. parties L,R with ideal points z_L,z_R
  - *Uncertainty about the position of the median voter*: both parties believe that the median voter has an ideal point $z_L<x_m<z_R$, where x_m is a random variable with mean zero
  - Also, assume quadratic preferences for both voters and parties
  - And, policy space X=[-1,1]

- Then, easy to show that under some conditions, the Nash equilibrium must involve policy divergence i.e. equilibrium platforms $x_L<x_R$
  - Intuition: parties face a trade-off between (i) converging to the median → higher probability of winning, and (ii) being closer to the party ideal point.

Uncertainty about the Position of the Median Voter

- Assume that for simplicity party ideal points are symmetric around zero i.e. $z_R=z$, $z_L=-z$ (symmetry). Then, can focus on party R.
- Also, assume uniform distribution of x_m on [-σ, σ] i.e. F(x)=1/2+x/2σ: σ measures parties' uncertainty about location of MV
- The median voter will vote for R iff distance from x_m to x_R is smaller than distance from x_m to x_L i.e.

$$x_m - x_L \geq x_R - x_m \Rightarrow x_m \geq \frac{x_R + x_L}{2}$$

- Probability that party R wins is thus $p_R = \Pr\left(x_m \geq \frac{x_L + x_R}{2} \right) = 1 - F\left(\frac{x_L + x_R}{2} \right) = \frac{1}{2} - \frac{x_L + x_R}{4\sigma}$
- Note as party R moves away from the expected position of the median voter, 0, and toward 1, p_R decreases at rate 1/4σ

.

Uncertainty about the Position of the Median Voter

- Objective for party R is $p_R\left(E-(x_R-z)^2\right)-(1-p_R)(x_L-z)^2$
- To find the Nash equilibrium platforms , we maximise this expression w.r.t. x_R , impose the equilibrium condition $x_R=x^*$, $x_L=-x^*$, and then solve the FOC for x*
- This is done in the Appendix and gives $x^*=\frac{z-E/4\sigma}{1+z/\sigma}$
- So, there is greater convergence (i.e. smaller x*):
- *The greater is E, rent from office;*
- *The smaller is z, i.e. the less extremist are the parties;*
- *The smaller is σ, the party uncertainty about the median voter*

Citizen-Candidate Theory of elections

- Developed by Besley and Coate (1997) and Osborne and Slivinski (1996)

- Addresses two limitations of Downs' model:
  - Where do parties come from?
  - What happens of parties cannot precommit to platforms?

- Key assumptions/changes to Down's model
  - Every voter can also be a candidate
  - If a voter stands for election and wins, he will set policy x equal to his ideal point x_i (platforms/manifestos, campaign promises are completely non-credible)

The Citizen-Candidate Model

- Sequence of events:

  1. All voters choose simultaneously whether to stand as a candidate at some cost c>0 or not
  2. Every voter votes for one of the candidates
  3. The winner i sets policy $x=x_i$, his ideal point
  4. If nobody stands, a default policy x_0 is chosen

- The CC model makes the opposite assumption to the Downs model about policy "platforms" – the truth is somewhere in between

- The CC model does not really describe political parties very well (but can be extended to do so).

Non-Convergence in The Citizen-Candidate Model

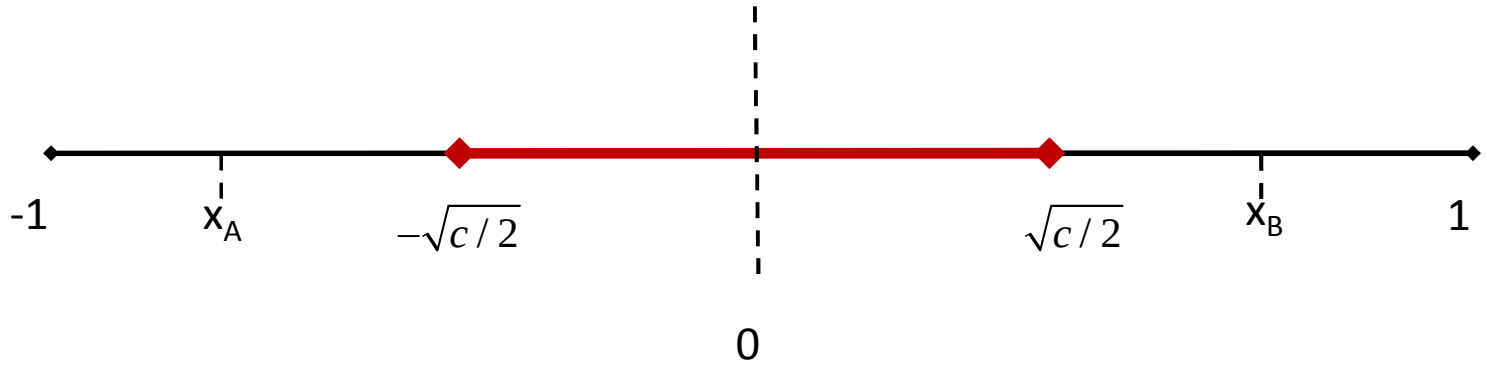
- The “problem” for the CC model is that there are many equilibria with different numbers of candidates
-
- We focus on the case that is comparable to the Downsian model, with two candidates

- Here, any two candidate equilibrium has the following properties:
  - one candidate is to the left of the median voter, and one to the right
  - There is non-convergence; the two candidate positions will be difference, and the larger the entry cost c, the larger this difference is
  - Intuition; a voter will not incur the cost c of fighting the election if there is already another candidates whose ideal point is very close to his own

Non-Convergence in The Citizen-Candidate Model

- Make assumptions for clarity/convenience:
  - Policy space is X=[-1,1]
  - Voters have quadratic utility i.e. suffer a loss of the squared deviations from their ideal point i.e. $-(x-x_i)^2$
  - Voter ideal points x_i distributed uniformly on [-1,1], so median voter has ideal point $x_m=0$.

- Then, equilibrium positions of candidates i.e. x_A, x_B must be centred around median, as below:



Minimum distance between candidate positions is $\sqrt{2c}$

Testing the Convergence Hypothesis

- Two ways of looking at convergence of party “positions” (i) election manifestos (platforms) and (ii) policies.

- Two problems with election manifestos:
  - hard to quantify: so far, only done at the national level (CMP)
  - manifestos may not be fully implemented by parties once in power

- New approach in the economics literature – regression discontinuity (RD):

- RD approach is simplest when there are only two major parties, L and R e.g. US system

The RD Approach

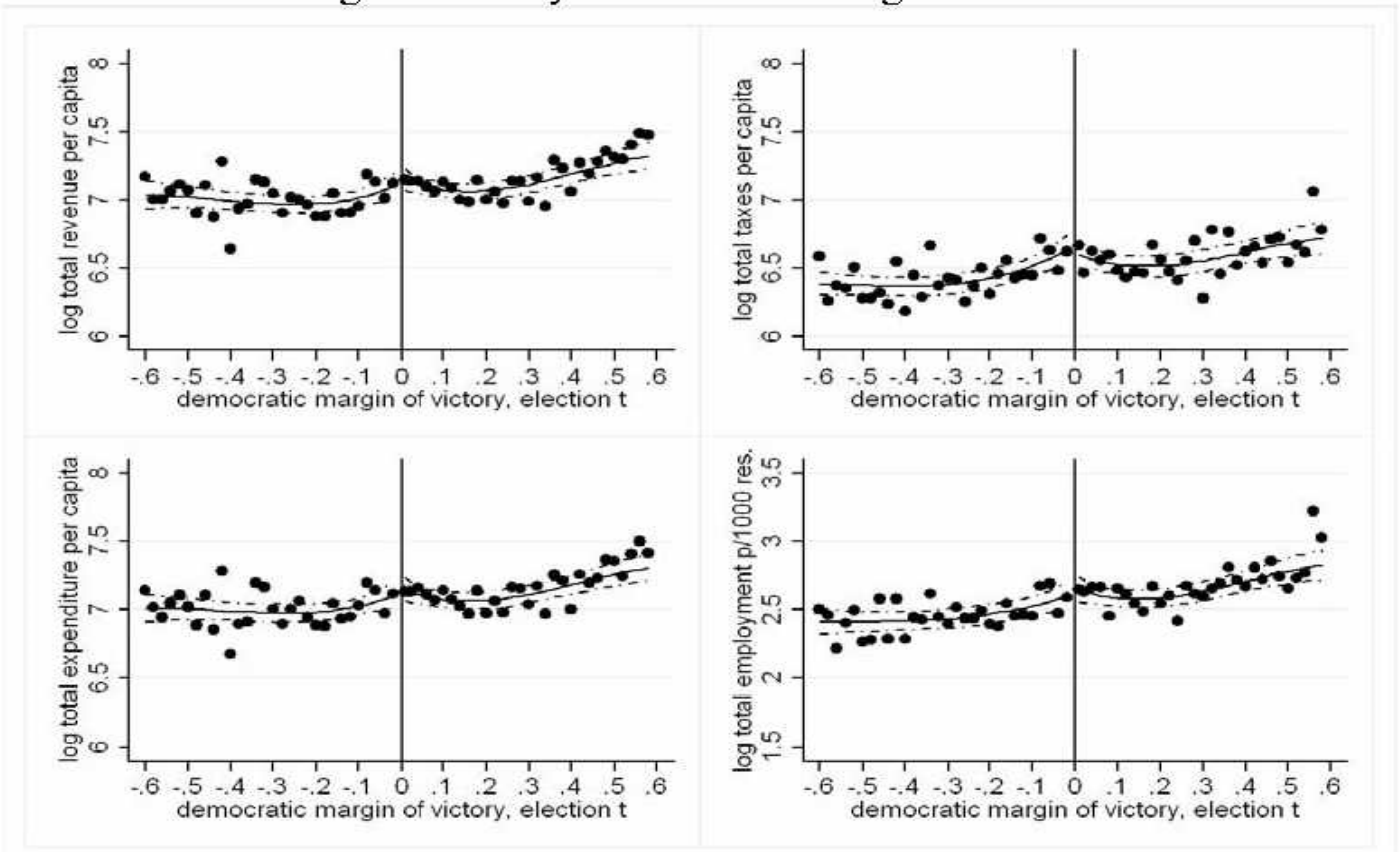
- RD compares two kinds of districts: one where the L candidate “just” wins, and one where the R candidate “just” wins
- Maintained assumption; the preferences of the median voter are similar in these two types of districts
- But then, *if the convergence hypothesis holds, the policies (or voting record) of the elected representatives following the election should be similar in these two types of districts*
- “Cleanest test”: local government in the US with directly elected mayors, as here (i) there are really only two parties; (ii) candidates are least likely to be constrained by national “party discipline”.

Study 1: Ferreira and Gyourko

- 413 US cities that directly elect a mayor, studied over the period 1950-2005
- Only studied elections where both winner and runner-up were Democrat or Republican
- Indicators of policy on left-right scale: total city government revenues, taxes, expenditure per capita, and city government employment per 1000 residents
- Main finding: no discontinuity in policy at zero margin of victory for Democrat candidates

Study 1: Ferreira and Gyourko

Figure 4. Party effect on size of government



Notes: The panels (from top left to bottom right) refer to the following outcomes: log total revenues per capita, log total taxes per capita, log total expenditures per capita and log total employment per 1,000 residents. All monetary measures are as of year 2000. Each dot corresponds to the unconditional mean within intervals of 0.02 Democratic margin of victory in election *t*. The solid line represents the predicted values of a third-order polynomial fit described in equation (3'), with the dashed lines showing the 95% confidence intervals.

Study 2: Lee, Moretti and Butler

- All US Congressional districts (for choosing members of the House of Representatives) over the period 1946–1995
- Only study elections where both winner and runner-up were Democrat, Republican
- Indicator of policy on left-right scale: ADA scores. *"For each Congress, the ADA chooses about twenty high-profile roll-call votes, and creates an index that varies between 0 and 100 for each Representative of the House. Higher scores correspond to a more "liberal" voting record."*
- Graphed Democrat vote shares at congressional elections in a given year (e.g. 1992) against ADA scores for elected Representatives in the following Congress (1993-4).
  - A clear discontinuity at zero: reject the convergence hypothesis
  - Congress members from different parties representing similar districts vote differently
  - Some party discipline even in US Congress!

Study 2: Lee, Moretti and Butler

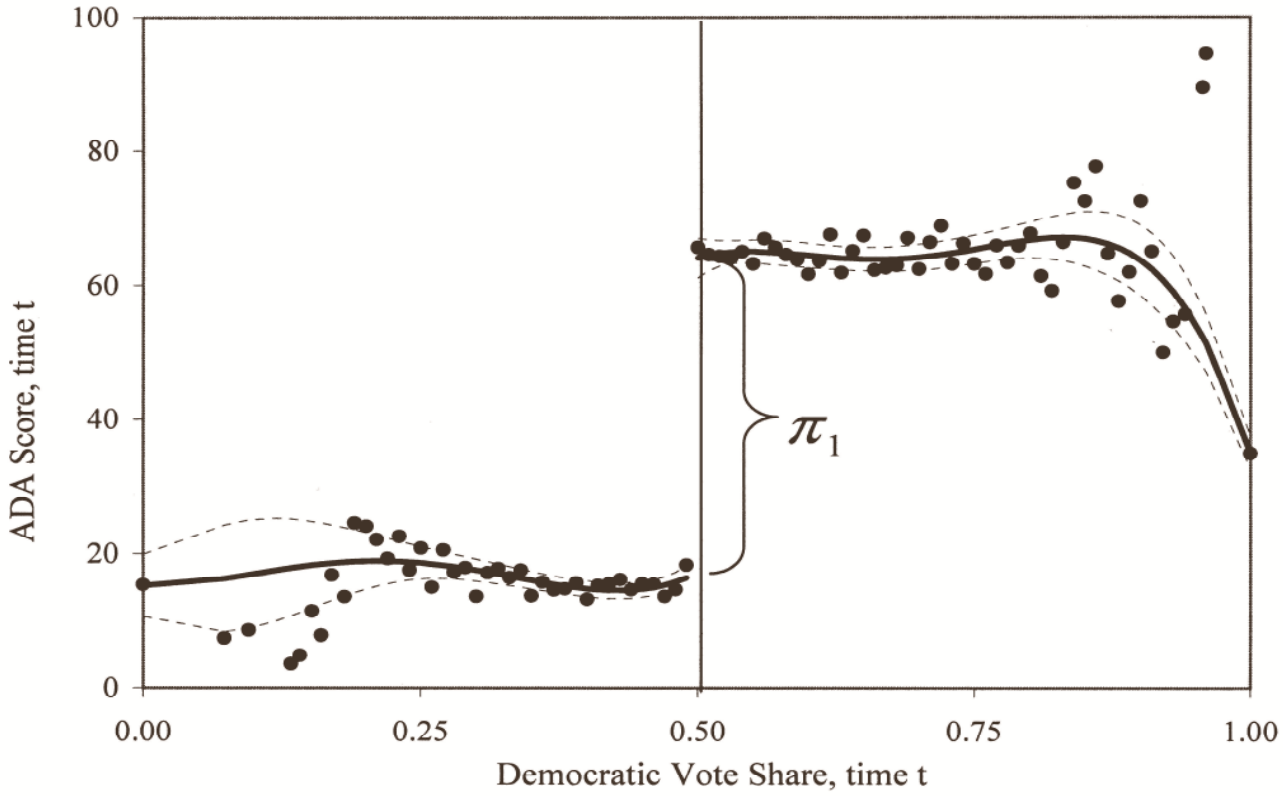


FIGURE IIa
Effect of Party Affiliation: π_1

Study 3: Lockwood and Porcelli

- Studies the effect of party control on local government expenditure and taxation in England and Wales over the period 1993-2012

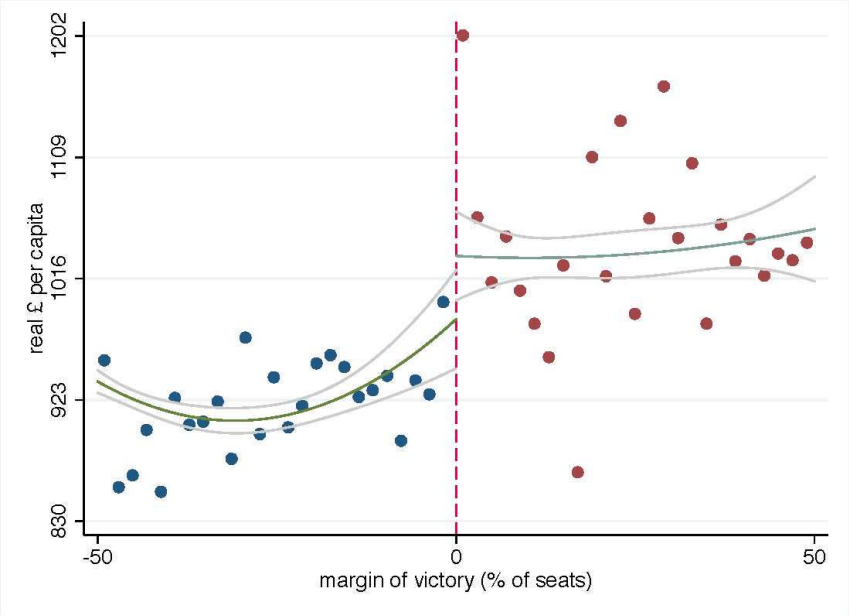
- Following FG, we only consider local authorities where (i) one of the two major parties in the UK system (Labour, Conservatives) has a majority on the council, and where (ii) the second-largest party on the council is the other of these two parties.
  - Labour governments appear to spend more and tax less than Conservative ones

- BUT, there is a party effect in grants: Labour local governments get between £183 and £280 more per capita in grants than unaligned ones

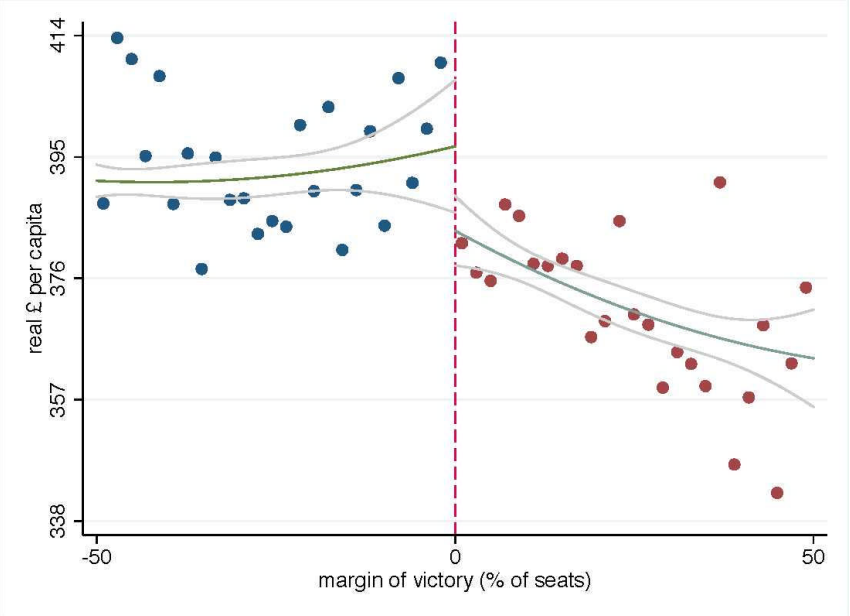
- Once this effect is allowed for, political control (Labour or Conservative) has no effect on council on current expenditure, the tax requirement, and the standardised council tax rate (consistent with FG)

Study 3: Lockwood and Porcelli

Expenditure per capita



Tax per capita

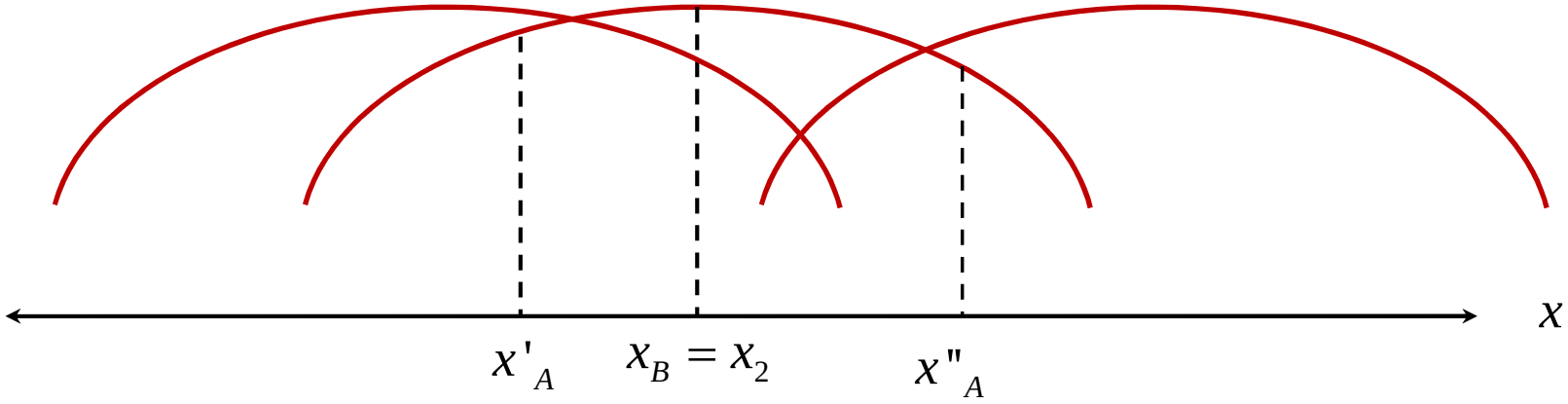


Appendix

Alternative Party Objectives

- In PR electoral systems with many parties and coalition government, parties will care about vote shares rather than just winning, because ministerial appointments, etc. (and thus power over setting policy) will depend on the vote share

- But, in the Downsian model, if the parties care about vote shares, the outcome is still complete convergence of policies to the median voter's preferred policy x_m.
  - If a party is not setting its platform at x_m, it can always increase its vote share by switching to x_m.



A's vote share is 1/3 at either x'_A or x''_A, can increase it to ½ by moving to x_2

Multiple Electoral Districts

- Nearly all electoral systems have multiple districts (even pure PR ones, such as the Netherlands)
- Assume k=1,..D districts, and let m,k be the median voter in district k, and $x_{m,k}$ his/her ideal point.
  - Order the districts so that $x_{m,1}<x_{m,2}<..x_{m,D}$
  - Let d=(D+1)/2 be the median district
- Assume that parties A,B are constrained to set the same policy platforms x_A,x_B in all districts (realistic?)
- Assume also that to win the election, the party must win in a majority of districts
- Then, the Nash equilibrium is where both parties "target" *the median voter in the median district* i.e. $x_A=x_B= x_{m,d}$
  - Reason: if $x_B= x_{m,d}$ and e.g. $x_A< x_B= x_{m,d}$, then party A will lose the median district and all districts "to the right" of the median and the election overall

More than Two Parties

- See Cox, "Centripetal and centrifugal incentives in electoral systems“, *American Journal of Political Science* (1990): 903-935 for a survey

- The basic insight is that “platform convergence” can break down

- The exact results depend on:
  - The number of parties
  - whether parties just care about winning or care about vote share
  - whether electoral competition is just between pre-existing parties or whether they also face the threat of entry by a new party.

- E.g. potential entry by a third party

  - Here, policy convergence cannot be an equilibrium, as the third party can locate just on one side of x_m and steal some of the votes of the first two parties, and win
  - Palfrey(1984): under some conditions, the two pre-existing parties will differentiate their policies to prevent a third party entering (but a pure-strategy equilibrium may not exist)

Appendix

Objective for party R is $p_R\left(B-(x_R-z)^2\right)-(1-p_R)(x_L-z)^2$.

Recalling that $\frac{\partial p_R}{\partial x_R}=-\frac{1}{4\sigma}$, the FOC for a maximum of this expression w.r.t. x_R is

$$-2p_R(x_R-z)-\frac{1}{4\sigma}\left[B-(x_R-z)^2+(x_L-z)^2\right]=0 \qquad (1)$$

At equilibrium, $p_R=0.5,\ x_L=-x_R=x^*$, which gives from (1) that

$$-(x_R-z)-\frac{1}{4\sigma}\left[B-(x_R-z)^2+(-x_R-z)^2\right]=0 \qquad (2)$$

Multiplying out squared terms and cancelling gives

$$-(x_R-z)-\frac{1}{4\sigma}\left[B+4x_Rz\right]=0$$

Solving (3) for x_R gives $x_R=\frac{z-B/4\sigma}{1+z/\sigma}$