Inferring Strategic Voting

Kei Kawai
Northwestern

Yasutora Watanabe
Northwestern

Jan 2011

Strategic Voting

- This paper addresses the extent to which voters vote sincerely or strategically
- Sincere Voting
  - Voting naively according to preferences
- Strategic Voting
  - Voting knowing that one should take into account the tie probabilities.

Strategic Voting

- Strategic voting is important in political economy
  - Different models make different assumptions
    - models of strategic voting
      comparing performance of different electoral rules (e.g. Myerson and Weber [1993])
      information aggregation (e.g. Feddersen and Pesendorfer)
    - models of sincere voting
      Multi-party competition (e.g.Palfrey 1984, Osborne [2000], Callander [2005])
      Citizen Candidate: Osborne and Slivinski [1996]
  - To what extent do voters vote strategically is an empirical question
- Strategic voting is important in practice
  - The extent to which voters vote strategically affects election outcomes: e.g., 2000 U.S. Presidential Election

Literature

- Previous empirical studies of strategic voting
  - Alvarez and Nagler (2000), Blais et al (2001)
  - Identifies misaligned voting, which is a subset of the set of strategic voting.
- Distinction between misaligned voting and strategic voting is important: *equilibrium object* v.s. *primitive*
- Other recent work: Ideological voting: Myatt (2007), Degan and Merlo (2009)

Setting

- 2005 Japanese House of Representatives Election
  - 3 or more candidates.
  - 175 plurality rule elections.
    - Variation in the district is important for identification.
- Model: Adaptation of Myerson and Weber (1993) with sincere voters.

What we do

- First, identify and estimate the preferences of voters in a setting where
  - Voting of strategic voters depend on (unknown) tie probabilities
  - The extent of strategic voting is unknown.
- Second, identify and estimate the extent of strategic voting
- Set identification and bounds estimation.

This paper

- we find large fraction [75.3%, 80.3%] of *strategic voters* on average.
- we find small fraction [2.4%, 5.5%] of *misaligned voting* on average.
  - this is close to the existing estimates of "strategic voting" (3% to 15%)
- In our counterfactual experiment
  - we find that absent any strategic voting, one party would gain [17,40] seats and another would lose [20,45] seats (out of 175).
  - Large change due to the fact that often races are close.

Outline

- Introduction
- Brief Intuition of Identification
- Model
- Data
- Identification
- Estimation
- Results and Counterfactual

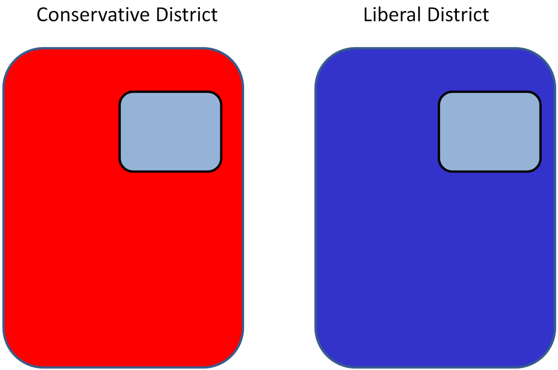
Identification: Preference

- Discrete choice setup: utility of voters as function of demographic and candidate characteristics
- Problem: No one-to-one correspondence between choice and preference for strategic voters: Depends on beliefs.
  - cf. Automobile choice for consumers
- Use the fact that it is a (weakly) dominated strategy to vote for least preferred candidate: partial identification

Identification: Extent of Strategic Voters

- Once preference is identified, possible to predict outcome if all voters vote according to preferences
- No strategic voters $\Longrightarrow$ match predicted outcome apart from random shocks
- With strategic voters, not necessarily the case
  - Strategic Voters $\Longrightarrow$ Systematic Deviation
- Then difference between predicted outcome and actual outcome identifies strategic voting.
- Source of identification in the data
  - liberal municipality in a liberal district (both sincere and strategic are likely to vote according to preferences)
  - liberal municipality in a conservative district (strategic voters likely to change their votes)

Identification: Extent of Strategic Voters



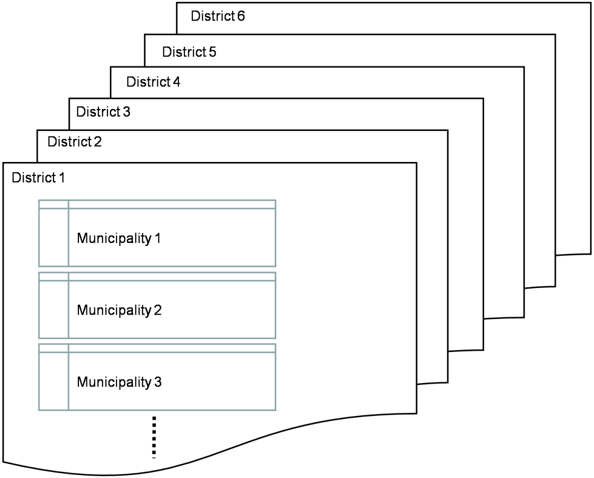
Model: Voter Preference

- In each election $d \in \{1, ..., D\}$,
- $K \geq 3$ candidates
- M municipalities $m_1, m_2, ..., m_M$
- N_m voters in municipality m
- Voter n's utility of having candidate k in office is

$$u_{nk} = u(\mathbf{x}_n, \mathbf{z}_k) + \xi_{km} + \varepsilon_{nk}$$

  where

  - $\mathbf{x}_n$: voter characteristics
  - $\mathbf{z}_k$: candidate characteristics
  - ξ_{km} : candidate-municipality shock
  - ε_{nk} : voter idiosyncratic shock


District 6
District 5
District 4
District 3
District 2
District 1
Municipality 1
Municipality 2
Municipality 3

Model: Voter Preference

- Sincere voter votes according to preference:

$$\text{vote for candidate } k \iff u_{nk} \geq u_{nl}, \forall l$$

- Strategic voter takes into consideration tie probabilities.

$$\text{vote for candidate } k \iff \bar{u}_{nk}(T_n) \geq \bar{u}_{nl}(T_n), \forall l$$

  - $T_n = \{T_{n,kl}\}_{kl}$ voter n's beliefs that k and l are going to be tied for 1st place
  - Expected utility from voting for k:

$$\bar{u}_{nk}(T_n) = \frac{1}{2} \sum_{l \in \{1,..,K\}, l \neq k} T_{n,kl}(u_{nk} - u_{nl})$$

  - Note that depending on the value of T_n, possible that voter votes for any candidate except the candidate k with the lowest u_{nk}.
  - Assume that for some $\{k, l\}$, $T_{n,kl} > 0$.
    - Then can normalize $\sum_k \sum_{l>k} T_{n,kl} = 1$, or $T_n \in \Delta^K$.

Model: Voter Type

- Voter type: r.v. $\alpha_{nm} \in \{0, 1\}$
  - $\alpha_{nm} = 0$ is sincere and $\alpha_{nm} = 1$ is strategic
  - α_{nm} drawn from a binomial.
  - Municipal specific prob. of success: α_m.
- Probability that voter n in municipality m is strategic:

$$\Pr(\alpha_{nm} = 1 | \alpha_m) = \alpha_m$$

- Probability that voter n in municipality m is sincere:

$$\Pr(\alpha_{nm} = 0 | \alpha_m) = 1 - \alpha_m$$

where

-
  - α_m: r.v. drawn from a Beta distribution

Common Belief Assumption

- **Assumption** Beliefs $T_n = \{T_{n,kl}\}_{kl}$ are common across voters in the same electoral district:

$$\{T_{n,kl}\} = \{T_{n',kl}\} \text{ for } \forall n, n'$$

Vote Share Outcome

- $V_{k,m}^{SIN}$: Fraction of votes cast by sincere voters to candidate k.
- $V_{k,m}^{STR}(T)$: Fraction of votes cast by strategic voters to candidate k.

$$V_{k,m}^{SIN} = \frac{\sum_{n=1}^{N_m}(1-\alpha_{nm})\cdot \mathbf{1}\{u_{nk} \geq u_{nl},\ \forall l\}}{\sum_{n=1}^{N_m}(1-\alpha_{nm})},$$

$$V_{k,m}^{STR}(T) = \frac{\sum_{n=1}^{N_m}\alpha_{nm}\cdot \mathbf{1}\{\overline{u}_{nk}(T) \geq \overline{u}_{nl}(T),\ \forall l\}}{\sum_{n=1}^{N_m}\alpha_{nm}}.$$

- Total vote share for candidate k:

$$\begin{aligned} V_{k,m}(T) &= \frac{\sum_{n=1}^{N_m}(1-\alpha_{nm})}{N_m} V_{k,m}^{SIN} + \frac{\sum_{n=1}^{N_m}\alpha_{nm}}{N_m} V_{k,m}^{STR}(T) \\ &\approx (1-\alpha_m)\iint \mathbf{1}\{u_{nk} \geq u_{nl}, \forall l\}]g(\varepsilon)d\varepsilon f_m(\mathbf{x})d\mathbf{x} \\ &+ \alpha_m \iint \mathbf{1}\{\overline{u}_{nk}(T) \geq \overline{u}_{nl}(T), \forall l\}g(\varepsilon)d\varepsilon f_m(\mathbf{x})d\mathbf{x} \end{aligned}$$

Solution Concept

- **C1** : $V_k > V_l \Rightarrow T_{kj} \geq T_{lj} \quad \forall k, l, j \in \{1, ..., K\}$
  - Pivot probabilities involving candidates with high vote shares are larger than those with low vote shares: e.g., $V_1 > V_2 > V_3 \Rightarrow T_{12} \geq T_{13} \geq T_{23}$
  - A weaker version of Myerson and Weber (1993) voting equilibrium
- **C2** : $V_{k,m} = \frac{\sum_{n=1}^{N_m}(1-\alpha_{nm})}{N_m} V_{k,m}^{SIN} + \frac{\sum_{n=1}^{N_m}\alpha_{nm}}{N_m} V_{k,m}^{STR}(T)$
  - Given beliefs T, voters vote optimally
  - embodies the restriction that no voter votes for his least preferred candidate

Definition *A set of solution outcomes W is all pairs of beliefs (T) and vote shares, $(\{V_{km}\}_{km})$,* $W = \left\{T, \{\{V_{k,m}\}_{k=1}^{K}\}_{m=1}^{M}\right\}$ *such that* **C1** and **C2** are satisfied.

- W is non-empty, not a singleton

Data

- Vote share data
  - 2005 Japanese House of Representatives Election
  - 175 plurality rule elections
  - Breakdown of votes for 1830 Municipalities
- Demographics data
  - Municipality level data on income distribution, years of schooling, population above 65.
- Candidate data
  - Party affiliation, hometown, previous political experience

Data Structure

In each electoral district, municipality-level breakdown is available.

District	Municipality	cand 1	cand 2	cand 3	cand 4
		Narumiya (JCP)	Kitagami (DPJ)	Nakagawa (LDP)	Tanaka (YUS)
Kyoto 4	Ukyo-ku	16,929	30,967	31,021	17,945
	Sakyo-ku	10,571	26,743	26,518	13,341
	Kameoka	4,613	10,732	8,629	27,573
	Miyama	581	498	808	1,554
	Sonobe	923	1641	1,961	4,923
⋮	⋮	⋮	⋮	⋮	⋮
	TOTAL	35,705	73,550	75,192	75,036

Data: Descriptive Statistics

	mean	st.d.	min	max	# obs
# of municipalities per district	9.3	7.1	2	36	175
winner's vote share (%)					
3 candidate district	52.9	5.6	36.0	73.6	158
4 candidate district	41.2	6.8	29.0	55.9	17
winning margin (%)					
3 candidate district	14.2	10.1	0.2	53.9	158
4 candidate district	9.4	9.7	0.1	35.5	17
margin between 2nd and 3rd (%)					
3 candidate district	30.4	7.4	4.7	43.3	158
4 candidate district	10.7	8.0	0.6	23.3	17

Identification: Overview

- Our identification argument follows two steps
  1. set-identify preference
  2. set-identify extent of strategic voting

Identification: Preference

Use the restriction that weakly-dominated to vote for the least preferred candidate

- Fix a preference parameter θ.

Ordering		$\Pr(Ordering \mid \theta)$
$A \succ B \succ C$	$\cdots$	$1/6$
$A \succ C \succ B$	$\cdots$	$1/6$
$B \succ A \succ C$	$\cdots$	$1/6$
$B \succ C \succ A$	$\cdots$	$1/6$
$C \succ A \succ B$	$\cdots$	$1/6$
$C \succ B \succ A$	$\cdots$	$1/6$

$$V_A \in [0, 2/3]$$
$$V_B \in [0, 2/3]$$
$$V_C \in [0, 2/3]$$

- No restriction on T.

Identification: Preferences

- Common Beliefs within a municipality

$$
\begin{array}{lcl}
A \succ B \succ C & \cdots & 1/6 \\
A \succ C \succ B & \cdots & 1/6 \\
B \succ A \succ C & \cdots & 1/6 \\
B \succ C \succ A & \cdots & 1/6 \\
C \succ A \succ B & \cdots & 1/6 \\
C \succ B \succ A & \cdots & 1/6
\end{array}
$$

$$
V_A \in [0, 1/2]
$$
$$
V_B \in [0, 1/2]
$$
$$
V_C \in [0, 1/2]
$$

Identification: Preferences

- Common Beliefs within a district

$A \succ B \succ C \cdots 1/6$		$A \succ B \succ C \cdots 0$	
$A \succ C \succ B \cdots 1/6$	$V_A^1 \in [0,1/2]$	$A \succ C \succ B \cdots 0$	$V_A^2 \in [0,1/4]$
$B \succ A \succ C \cdots 1/6$	$V_B^1 \in [0,1/2]$	$B \succ A \succ C \cdots 1/4$	$V_B^2 \in [0,3/4]$
$B \succ C \succ A \cdots 1/6$	$V_C^1 \in [0,1/2]$	$B \succ C \succ A \cdots 1/4$	$V_C^2 \in [0,3/4]$
$C \succ A \succ B \cdots 1/6$		$C \succ A \succ B \cdots 1/4$	
$C \succ B \succ A \cdots 1/6$		$C \succ B \succ A \cdots 1/4$	

- $(V_A^1, V_B^1, V_C^1) = (0, 1/2, 1/2)$ and $(V_A^2, V_B^2, V_C^2) = (1/4, 3/4, 0)$ are individually not rejected.
  - $(V_A^1, V_B^1, V_C^1) = (0, 1/2, 1/2)$ requires $T_{BC} \approx 1$.
  - $(V_A^2, V_B^2, V_C^2) = (1/4, 3/4, 0)$ requires $T_{AB} \approx 1$.
- Jointly rejected

Identification Extent of Strategic Voting

Assume (for now) that we have identified the preference parameters.

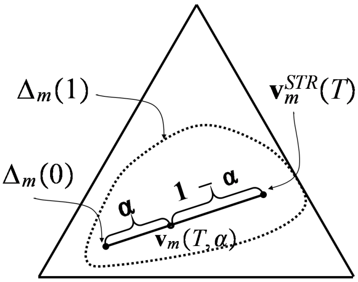
- Recall that if we knew T, and for given preference parameters, then we could predict outcome as a function of α_m (and ξ_m) as:

$$\begin{aligned} V_{k,m}(T,\alpha_m) &\approx (1-\alpha_m)v^{SIN}_{k,m} + \alpha_m v^{STR}_{k,m}(T) \\ \text{where } v^{SIN}_{k,m} &\equiv \iint \mathbf{1}\{u_{nk} \geq u_{nl}, \forall l\}]g(\varepsilon)d\varepsilon f_m(\mathbf{x})d\mathbf{x} \\ v^{STR}_{k,m}(T) &\equiv \iint \mathbf{1}\{\overline{u}_{nk}(T) \geq \overline{u}_{nl}(T), \forall l\}g(\varepsilon)d\varepsilon f_m(\mathbf{x})d\mathbf{x} \\ & \quad g \text{ and } f_m \text{ are dist. of } \varepsilon \text{ and characteristics } \mathbf{x} \end{aligned}$$

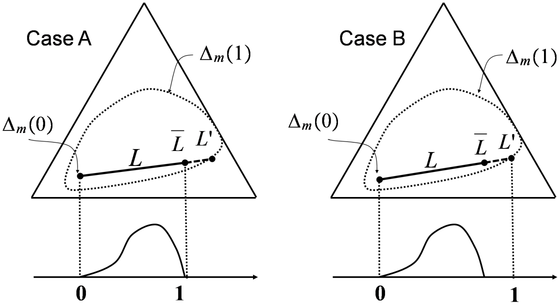
  - $u_{nk} = u(\mathbf{x}_n, \mathbf{z}_k) + \xi_{km} + \varepsilon_{nk}$
  - $\bar{u}_{nk}(T_n) = \frac{1}{2}\sum_{l\in\{1,..,K\},l\neq k} T_{n,kl}(u_{nk} - u_{nl})$
  - $V_{k,m}(T,\alpha_m)$ is the point which divides $v^{SIN}_{k,m}$ and $v^{STR}_{k,m}(T)$ with ratio $(1-\alpha_m)$ and α_m.
- However, we do not know T
  - $\bigcup_T V_m(T,\alpha_m)$ is the predicted set of vote shares

Identification: Extent of Strategic Voting

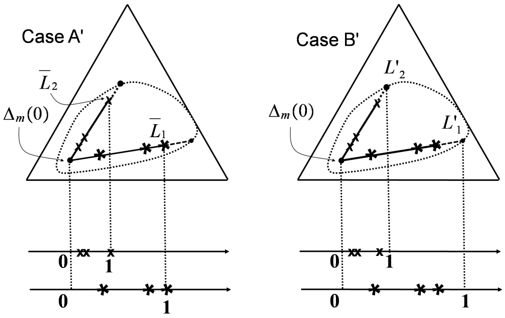
- Vote share can be expressed as a point in the simplex $v_{1m} + v_{2m} + v_{3m} = 1$.
  - Define $\Delta_m(\alpha_m)$ as the possible set of vote shares when the fraction of strategic voters is α_m: $\Delta_m(\alpha_m) = \bigcup_T V_m(T, \alpha_m)$
  - $\Delta_m(0)$ is the vote share corresponding to the what the sincere voters vote: $\Delta_m(0) = v_{k,m}^{SIN}$



- First: Ignore randomness from $\boldsymbol{\xi}$ ($\boldsymbol{\xi} \approx \mathbf{0}$)
- Infinite # of municipalities within district
  - i.e. One election and many municipalities.
- Fix demographic characteristics.
  - Then V_m line up along a line segment between $v_{k,m}^{SIN}$ and $v_{k,m}^{STR}(T)$
  - $v_{k,m}^{SIN}$ and $v_{k,m}^{STR}(T)$ same for all m: Differences among V_m arise from differences in α_m.
  - observed vote share map differently into different values of α depending on the position of $v_m^{STR}(T)$, e.g, it can be $\overline{L}$ or L'
- Case A corresponds to UB, and Case B to LB.



Second: No randomness, infinite # of districts, finite # of municipalities within district



- Position of $v_m^{STR}(T)$ changes across districts because T are different. Still, a vote share realization can be mapped to a value of α.
- Case A$'$: the upper bound, and Case B$'$: the lower bound.
- Intuitively, having "liberal municipalities" in conservative districts make L' and $\bar{L}$ close: necessary to make tight prediction on lower bound of the distribution of α.

Identification: Extent of Strategic Voting

1. In the actual data, the vote shares would not lie on the same line segment

- So far, our identification argument has ignored randomness in ξ_m
  - ξ_m is the municipality level shock that would account for dispersion.
  - If the distribution of ξ_m were completely unrestricted $\longrightarrow$ may be hard to separately identify the distribution of α_m and ξ_m.
- But as long as ξ_m is nicely behaved: uni-modal and mean zero, the same intuition carries through.
  - We assume ξ follows Normal $\longrightarrow$ can parametrically account for the dispersion around the line segment.

2. Preference parameters are only partially identified

- For each preference parameter in the identified set, we recover the distribution of α. Take the union.

Estimation: Overview

- If we knew T, the model predicts unique outcome given parameters $\implies$ we can estimate with MLE or GMM.
- Unobserved T , which results in partial identification of parameters, makes inequality based estimation appropriate.
- Operationalize our identification argument
  - Need to exploit variation in demographic characteristics and votes.
  - Need to exploit the structure (district-municipality)
- Use Pakes, Porter, Ho, Ishii (2007)

Estimation: Specification of Utility Function

- Recall utility of voter n in municipality m, electing candidate k:

$$u_{nmk} = u(\mathbf{x}_n, \mathbf{z}_{km}; \theta^{PREF}) + \xi_{km} + \varepsilon_{nk}$$

- We allow both ideological ("spacial" component) and non-ideological component in utility

$$u(\mathbf{x}_n, \mathbf{z}_{km}; \theta^{PREF}) = -(\theta^{ID}\mathbf{x}_n - \theta^{POS}\mathbf{z}_k^{POS})^2 + \theta^{QLTY}\mathbf{z}_{km}^{QLTY}$$

  - $\theta^{ID}\mathbf{x}_n$: ideological position of voter n
  - $\mathbf{x}_n$: income, education, age (over 65 or not)
  - $\theta^{POS}\mathbf{z}_k^{POS}$: ideological position of candidate k
  - $\mathbf{z}_k^{POS}$: political party affiliation of candidate k
  - $\theta^{QLTY}\mathbf{z}_{km}^{QLTY}$: quality of candidate k
  - $\mathbf{z}_{km}^{QLTY}$: hometown, political experience (e.g., incumbent) of candidate k

Estimation: Constructing Moment Inequalities

Construct moment inequalities based on our identification argument. Idea similar to Indirect Inference

1. Fix some θ and T
2. For any random shocks $\boldsymbol{\xi}$ and $\boldsymbol{\alpha}$, model predicts outcome $v^{PRED}(T, \theta)$
3. In each district d, regress $v^{PRED}(T, \theta)$ on demographic and candidate characteristics and obtain $\beta_d(T, \theta)$ for each district.
4. Vary T to find $\bar{\beta}_d(\theta) = \sup_T \beta(T, \theta)$ and $\underline{\beta}_d(\theta) = \inf_T \beta(T, \theta)$ and integrate them over distribution of shocks ξ and α.
5. Regress actual vote share v^{DATA} on demographic and candidate characteristics and obtain β_d^{DATA}.
6. Construct moments as

$$E[\underline{\beta}_{k,d}(\theta_0) - \beta_{k,d}^{DATA}] \leq 0, \text{ and } E[\overline{\beta}_{k,d}(\theta_0) - \beta_{k,d}^{DATA}] \geq 0.$$

Extensions:

1. The fraction of strategic voting (α) may depend on T.
   - people become aware of the possibility to vote strategically when T is close.
     - one way to test this is compare
       - variance of α within districts: $var^1(\alpha)$
       - overall variance of α across districts: $var^2(\alpha)$
     - If $var^1(\alpha) < var^2(\alpha)$, then evidence of dependence of α on T.
2. Include voter turn-out.
   - Include a cost (or consumption value) of voting: outside option.
     - Key difference is that the absolute value of T_n matter for turnout
       - Scale matters: cannot normalize $\sum_k \sum_{l>k} T_{n,kl} = 1$
     - Use level of turnout to pin down $\sum_k \sum_{l>k} T_{n,kl}$.
       - Standard pivotal-voting models are sensitive to very small changes in T.
       - Found computational burden to be quite high.

Results: Parameter Estimates

The utility function we estimate is

$$u(\mathbf{x}_n, \mathbf{z}_{km}; \theta^{PREF}) =$$
$$-\left\{[\theta^{const}, \theta^{income}, \theta^{educ}, \theta^{>65}]\mathbf{x}_n - [\theta^{LDP}, \theta^{JCP}, \theta^{DPJ}, \theta^{YUS}]\mathbf{z}_k^{POS}\right\}^2$$
$$+[\theta^{incumb}, \theta^{prev}, \theta^{no\_exp}, \theta^{htown1}, \theta^{htown2}, \theta^{htown3}, \theta^{htown4}]\mathbf{z}_{km}^{QLTY}$$
$$+\xi_{km} + \varepsilon_{kn},$$

	C.I.		C.I.
θ^{const}	$[-1.420, -1.418]$	$\theta^{incumbent}$	0
θ^{income}	$[-0.164, -0.162]$	$\theta^{previous}$	$[-0.204, -0.199]$
$\theta^{education}$	$[0.177, 0.179]$	$\theta^{no\_experiecne}$	$[0.080, 0.083]$
$\theta^{above65}$	$[-0.003, -0.001]$	$\theta^{hometown1}$	$[0.437, 0.444]$
θ^{JCP}	$[-3.467, -3.448]$	$\theta^{hometown2}$	$[0.180, 0.187]$
θ^{DPJ}	$[-2.998, -2.990]$	$\theta^{hometown3}$	$[0.038, 0.041]$
θ^{YUS}	$[-0.068, -0.065]$	$\theta^{hometown4}$	0
θ^{LDP}	0	θ_{ξ}	$[0.373, 0.385]$

Results

- Two parameters of the beta distribution
  - θ_{α_1}:$[5.210, 6.005]$ and θ_{α_2}:$[1.473, 1.706]$
- Average extent of strategic voting is [75.3%, 80.3%]
- Using the estimated model, we can recover the extent of misaligned voting.
  - Average extent of misaligned voting is [2.4%, 5.5%]
  - This number is comparable to the estimate of "strategic voting" in the existing studies, which varies from 3% to 15%.

Counterfactual Experiment: Sincere Voting

	JCP	DPJ	LDP	YUS
Actual				
Vote Share (%)	7.7	38.4	49.7	35.0
Number of Seats	0	35	131	9
Counterfactual				
Vote Share (%)	[8.4, 10.2]	[40.6, 43.8]	[42.6, 45.7]	[33.9, 38.8]
Number of Seats	[0, 0]	[52, 75]	[86, 111]	[11, 18]

- DPJ adds [17, 40] seats and LDP would lose [20, 45] seats.
  - The impact is large because the winning margin is often small.
- JCP and DPJ vote shares increase, and LDP vote shares decrease.
  - LDP candidates are strong while DPJ and JCP candidates were not.
- Sincere-voting effect for YUS is small.

Conclusion

This paper

- proposes an estimable model of strategic voting
- study identification in discrete choice setting
- estimation using inequality based estimator (Pakes et. al. (2006))
- results show a large fraction [75.3%, 80.3%] of *strategic voters*, and a small fraction [2.4%, 5.5%] of *misaligned voting.*
- counterfactual experiment: sincere voting under plurality