Notes on Strategic Voting

Francesco Trebbi

February 12, 2018

Idea

Kawai and Watanabe (AER 2013): Inferring Strategic Voting.

- They structurally estimate a model of strategic voting and quantify the impact it has on election outcomes.
- Use Japanese general-election data.
- $D = 300$ plurality rule elected members of parliament. Elected in single-member districts, each district has 9 municipalities on average.
- Plurality rule.
- District is the unit of observation and voting games are played in each district independently of each other.

Strategic Voting

Kawai and Watanabe make an important distinction:

- **Strategic voters:** Voters that make voting decisions conditioning on the event that their votes are pivotal. Their share is a primitive of the environment.
- **Misaligned voters:** Subset of strategic voters that actually vote for a candidate other than the one they most prefer. Their share is an equilibrium object.

- Note: Voters can be strategic, but you will not know it if in a specific election their unconditionally preferred candidate is also the one they would vote for, conditionally on being pivotal.

Identification Idea

Identification Idea:(Partial) Identification of share of strategic voters.

The idea:

- Multiple districts D each with multiple m.
- Consider two liberal municipalities: one in a generally conservative electoral district and the other in a generally liberal district.
- Suppose that there are three candidates ($K = 3$), a liberal, a centrist, and a conservative candidate in both districts.

Identification Idea

Identification Idea:(Partial) Identification of share of strategic voters.

The idea:

- If there are no strategic voters, we would not expect the voting outcome to differ across the two municipalities.
- In the presence of strategic voters, the voting outcome in these two municipalities could differ.
- If the strategic voters of the municipality in the conservative district believe that the liberal candidate has little chance of winning, those voters would vote for the centrist candidate, while strategic voters in the other municipality (in the liberal district) would vote for the liberal candidate according to their preferences (if they believe that the liberal candidate has a high chance of winning).

Results

Results:

- Based on structural parameters –preferences and demographics – the authors can predict what would happen under sincere voting and use the difference from what actually happens.
- They take this discrepancy as a measure of the extent of strategic voting.
- Find a large fraction $[63.4\%, 84.9\%]$ of strategic voters, only a small fraction $[1.2\%, 2.7\%]$ of whom voted for a candidate other than the one they most preferred (misaligned voting).

- Existing empirical literature has not distinguished between the two, in fact estimating misaligned voting instead of strategic voting.
- Note: you should always be wary of papers that assign the residual between an empirical model and data to their preferred hypothesis. Lack of fit gets interpreted as quantitative validation of hypothesis.

Setup

300 district/races, $\simeq$ 3000 municipalities. Voter i in municipality m has utility from having candidate k elected in office:

$$U_{ik} = u\left(X_i, Z_k; \theta^{PREF}\right) + r_{km} + e_{ik}$$

θ^{PREF} preference parameters
X_i voter characteristics
Z_k candidate characteristics
r_{km} candidate-municipality match shock ($\sim$Normal$(0, \theta_r)$)
e_{ik} candidate-voter preference shock ($\sim$Type-I extreme value)

Setup

- In an election a single vote is rarely decisive (at a tie or a candiate down one vote). However, if that happens the vote is said to be pivotal.
- Conditionally on voter i being pivotal, the gain having candidate $k \in \{1, ..., K\}$ in office instead of, say, candidate l is given by:

$$U_{ik} - U_{il}$$

- This pivotal event is defined by a belief T_{kl} and may or may not be common to every voter i in district d
- Some of these beliefs must be nonzero, but does not matter how small (just rescale by $1/\sum_{l \in \{1,...,K\}} T_{kl}$).
- The expected utility of voting for k is:

$$\mathrm{E}\left[U_{ik}\right] = \sum_{l \in \{1,...,K\}} T_{kl}\left(U_{ik} - U_{il}\right)$$

Setup

- A strategic voter i in municipality m will support candidate k if for any other candidate l:

$$\mathrm{E}\left[U_{ik}\right] \geq \mathrm{E}\left[U_{il}\right]$$

- A sincere voter i in municipality m will support candidate k if for any other candidate l:

$$U_{ik} \geq U_{il}$$

- In municipality m there is a share of strategic voters $\alpha_m(\sim\text{Beta}(\theta_\alpha))$ and the rest are sincere voters.

Setup

- Note that there is no $T_{i,kl}$. The set of pivotality beliefs $\mathbf{T}$ is common across all voters in d.
- Equilibrium candidate vote shares $v_{km}\left(\theta^{PREF};\alpha_m;\mathbf{T}\right)=$ the share of the population that is sincere and votes for k + share of the population that is strategic and votes for k.

Identification

Preference parameters and fraction of strategic voters in the municipality:

1. The paper posits to identify preference parameters θ^{PREF}
2. and the distribution of fractions of strategic voters in each municipality m in district d, $\{\alpha_m\}_{m=1}^{M_d}$
3. To do so the authors use information on vote shares for each candidate in each municipality m, candidate characteristics, and municipality economic and social characteristics (aggregate of i).
4. **We will now show how the observed vote shares (data) will discipline the parametric spaces of both preference and strategic voters shares, i.e. that we have identification.**

Inequalities providing partial (set identification)

Voters:

1. Suppose there are three candidates $k = A, B, C$ in district d
2. There are six possible groups of voters in municipality m of district d based on their preference orderings:
   $A \succ B \succ C$
   $A \succ C \succ B$
   $B \succ A \succ C$
   $B \succ C \succ A$
   $C \succ A \succ B$
   $C \succ B \succ A$
3. Assume each are $1/6$ of the voting population.

Inequalities providing partial (set identification)

Voters:

Restriction I: Consider voters with ordering $B \succ C \succ A$ and $C \succ B \succ A$

Call the pivotal event for these candidates T_{AB} T_{BC} T_{BC}

1. If $T_{AB} \simeq 1$ and T_{BC} $T_{BC} \simeq 0$ then both types of voters will vote for B.
2. If $T_{AC} \simeq 1$ and T_{AB} $T_{BC} \simeq 0$ then both types of voters will vote for C.
3. If $T_{BC} \simeq 1$ and T_{AB} $T_{AC} \simeq 0$ then $1/2$ types of voters will vote for C and the rest for B. But none for A

This means that voters with ordering $B \succ C \succ A$ and $C \succ B \succ A$ (i.e. that like A the worst) **never vote** for A.

It is a weakly dominated strategy for voters (strategic or sincere) to vote for their least preferred candidate. This imposes restrictions on the relations between vote shares and voter preferences.

This bounds the vote share for A in m between $[0, 2/3]$
($1/3$ of the voters will never pick A).

Inequalities providing partial (set identification)

Voters:

Restriction II: Consider voters with ordering $B \succ A \succ C$ and $C \succ A \succ B$

The pivotal events for these candidates T_{AB} T_{BC} T_{BC} are common.
Common beliefs.

1. If $T_{AB} \simeq 1$ and T_{BC} $T_{BC} \simeq 0$ then the first type of voters will vote for B, the second type for A
2. If $T_{AC} \simeq 1$ and T_{AB} $T_{BC} \simeq 0$ then the first type of voters will vote for A, the second type for C
3. If $T_{BC} \simeq 1$ and T_{AB} $T_{AC} \simeq 0$ then the first type of voters will vote for B, the second type for C

This means that voters with ordering $B \succ A \succ C$ and $C \succ A \succ B$ (i.e. that like A so-so) never vote for A under a common set of beliefs (by points 2 and 3)

This further bounds the vote share for A between $[0, 1/2]$
($1/3$ of the voters will never pick A and you never get another $1/6$ under common beliefs).

Inequalities providing partial (set identification)

Voters:

Restriction III: The authors are not assuming just common beliefs within a municipality m, but across municipalities in district d.

1. For example municipality m_1 and m_2 in district d (see Figure 4 of the paper).
2. This adds extra restrictions and further tightens the set of preference ordering allowed once the data give you the actual vote shares for A, B, C in each municipality m in district d.
3. That is the arguments states that if there is no pivotality beliefs $\mathbf{T}$ that can rationalize the data under a specific set of preferences (i.e. whose parameters we are trying to estimate), then that set of parameters for the preferences has to be rejected - it's not in the identified set.

Inequalities providing partial (set identification)

Consistency Requirement C1:

Restriction IV: The authors also assume that if one sees higher vote shares for a candidate A than B, $V_A > V_B$, then

1. For any other candidate $A, B, k \in \{1, ..., K\}$ in the set of possible candidates $T_{Ak} \geq T_{Bk}$
2. Pivot probabilities for candidates with higher vote shares are larger than those with low vote shares.

Identification (cont.d)

Preference parameters:

1. Restrictions I-IV allow to identify preference parameters θ^{PREF} by only leaving the identified space where beliefs obey such restrictions for any $\{\alpha_m\}_{m=1}^{M_d}$ and without bothering to pin down the pivotality beliefs $\mathbf{T}$

2. How do we identify the fractions of strategic voters in each municipality m in district d, $\{\alpha_m\}_{m=1}^{M_d}$?

Identification of Strategic Voters

Fraction of Strategic Voters:

1. Now take as given parameters θ^{PREF}
2. Intuition for identification of $\{\alpha_m\}_{m=1}^{M_d}$: use the variation in voting outcomes among municipalities in different districts with identical characteristics vis-à-vis the variation in the vote shares of candidates (and characteristics of other municipalities in the same district).
3. Again, partial identification only.

Identification of Strategic Voters

Example: Consider two districts d_1 and d_2 with three identical candidates each, $k = A, B, C$, and consider two identical municipalities m_1 and m_2 set in each district, both characterized by a single type of voters with orderings $A \succ B \succ C$.

Suppose in d_1 $T_{AB} \simeq 1$ and T_{BC} $T_{BC} \simeq 0$ and in d_2 $T_{BC} \simeq 1$ and T_{AB} $T_{BC} \simeq 0$, then under 100% sincere voting (and no random shocks)

$$v_{Am_1}\left(\theta^{PREF}; \alpha_{m_1} = 0; \mathbf{T}\right) = v_{Am_2}\left(\theta^{PREF}; \alpha_{m_2} = 0; \mathbf{T}\right) = 1$$

but under 100% strategic voting:

$$v_{Am_1}\left(\theta^{PREF}; \alpha_{m_1} = 1; \mathbf{T}\right) = 1 > v_{Am_2}\left(\theta^{PREF}; \alpha_{m_2} = 1; \mathbf{T}\right) = 0$$

Identification of Strategic Voters

Because under sincere voting the vote shares in the two municipalities can be different only due to random shocks, the differences in vote share in excess to random shocks identify the strategic voters.

Example 2: Consider the same two districts d_1 and d_2 with three identical candidates each, $k = A, B, C$, and add to m_1 and m_2 a third identical one m_2' located in d_2. Again let m_2' be characterized by a single type of voters with orderings $A \succ B \succ C$, but now allow $\alpha_{m_2} \neq \alpha_{m_2'}$

Suppose in d_2 $T_{BC} \simeq 1$ and T_{AB} $T_{BC} \simeq 0$, then under $\alpha_{m_2} > \alpha_{m_2'}$ strategic voting (and no random shocks)

$$v_{Am_2}\left(\theta^{PREF}; \alpha_{m_2}; \mathbf{T}\right) > v_{Am_2'}\left(\theta^{PREF}; \alpha_{m_2'}; \mathbf{T}\right)$$

Hence, you have further constraints.

Estimation

Estimation

- Inequality based estimator proposed by Pakes et al. (2007). It's a set based estimator (Method of Moment Inequalities). No point identification.
- Uses only restrictions that do not depend on **T**, because **T** cannot be pinned down by the model. This is a very unfortunate conditions to be in (e.g. set identification). Kendall et al. (2015), Cruz et al. (2018) show how to properly address this issue of beliefs with micro data.

Estimation

Estimation Steps

1. Take district d.
2. Each municipality m in district d has some covariates X_m and we have $X_d = (X_1, ..., X_{M_d})'$. We also have the actual vote shares $v_{k,m}^{data}$ for each candidate $v_{k,d}^{data} = \left(v_{k,1}^{data}, ..., v_{k,M_d}^{data}\right)'$.
3. Regress $v_{k,m}^{data}$ on X_m obtaining $\beta_{k,d}^{data} = (X_d' \; X_d)^{-1} X_d' \; v_{k,d}^{data}$
4. Call $\theta = \left[\theta^{PREF}, \theta_\alpha, \theta_r\right]$.

Estimation

Estimation Steps (cont.d)

5 Fix $\theta, \mathbf{T}^d, \alpha_d, r_d$ and generate through the model $v_{k,d}^{\text{model}}\left(\theta, \mathbf{T}^d, \alpha_d, r_d\right) = \left(v_{k,1}^{\text{model}}\left(\theta, \mathbf{T}^d, \alpha_1, r_1\right), ..., v_{k,M_d}^{\text{model}}\left(\theta, \mathbf{T}^d, \alpha_{M_d}, r_{M_d}\right)\right)'$

6 Regress $v_{k,m}^{\text{model}}$ on X_m obtaining $\beta_{k,d}^{\text{model}}\left(\theta, \mathbf{T}^d, \alpha_d, r_d\right) = \left(X_d' \; X_d\right)^{-1} X_d' \; v_{k,d}^{\text{model}}\left(\theta, \mathbf{T}^d, \alpha_d, r_d\right)$

7 Search across all possible **C1** consistent $\mathbf{T}^d$ to get $\beta_{k,d}^{\max}\left(\theta, \alpha_d, r_d\right) = \max\left(\beta_{k,d}^{\text{model}}\left(\theta, \mathbf{T}^d, \alpha_d, r_d\right)\right)$ and $\beta_{k,d}^{\min}\left(\theta, \alpha_d, r_d\right) = \min\left(\beta_{k,d}^{\text{model}}\left(\theta, \mathbf{T}^d, \alpha_d, r_d\right)\right)$

Estimation

Estimation Steps (cont.d)

8 Simulate α_d, r_d and integrate out creating
$\beta^{\max}_{k,d}(\theta) = \int\int \beta^{\max}_{k,d}(\theta, \alpha_d, r_d)\, dF_\alpha dF_r$ and
$\beta^{\min}_{k,d}(\theta) = \int\int \beta^{\min}_{k,d}(\theta, \alpha_d, r_d)\, dF_\alpha dF_r$

9 Then get θ by minimizing the following two criterion functions across all k and d:

$$
\begin{aligned}
Q^{\max}(\theta) &= \sum_k \min\left(0, \beta^{\max}_{k,d}(\theta) - \beta^{data}_{k,d}\right) \\
Q^{\min}(\theta) &= \sum_k \max\left(0, \beta^{data}_{k,d} - \beta^{\min}_{k,d}(\theta)\right)
\end{aligned}
$$

10 Confidence intervals by Pakes et al. (2007).

Counterfactual

Counterfactual

1. Assume all voters are sincere.
2. Then you can predict vote shares just based on $\left[\theta^{PREF}, \theta_r\right]$
3. And one can check ho different electoral outcomes can be under only sincere voting relative to what happens in reality (where strategic voters are about 66% of the populace). Substantial differences even if misaligned voters are only 4%. This is because voting margins are small in these FPTP elections.

Summary

Summary

1. I tried to give you a sense of the machinery here. To see whether we could open the black box of a non-trivial structural political economy model.
2. Not all structural models in our field have this intense IO feel to them though.

More on Strategic Voting: Duverger's Law

- Duverger's (1954): "Simple-majority single-ballot [Plurality or First-Past-The-Post rule] favors the "two party system" whereas "Simple Majority with a Second Ballot [dual-ballot or runoff] or Proportional Representation favors multipartyism."

More on Strategic Voting

- Fujiwara (QJPS 2011): A Regression Discontinuity Test of Strategic Voting and Duverger's Law.
- Regression Discontinuity Design in assignment of electoral rules in Brazilian municipalities' mayoral elections.

More on Strategic Voting: Party Abandonment

- Theoretical mechanism of the Duverger's Law: Strategic voting.
- Sincere voting: Voting own's preferences. Pick the candidate a voter likes the best in an electoral roster. In this case electoral rule does not matter.
- Strategic voting: Pick the candidates a voter likes the best weighted by their chance of electoral success. Electoral rule matters.
- Consider plurality rule and three candidates A, B, C. Suppose you are pivotal. You prefer candidate C to A & B, and A to B, but you are the only one who likes C in this system. A and B must be tied (we are considering the case you are pivotal & ties are solved by coin toss). Equilibrium: You will vote for A.
- Palfrey (1989); Myerson (2002); Myatt (2002); Bouton (2011). Under single ballot there exist an equilibrium where only the first two candidates receive all votes. But there are other equilibria, some with partial abandonment.

Natural Experiment

- The Brazilian constitution mandates that municipalities $<$ 200,000 registered voters use Single-Ballot plurality rule to elect their mayors, while Runoff rule if $>$ 200,000.
- Regression Discontinuity Design (Lee, 2008): Quasi-experiment. Balance on covariates. The data is dense enough around the treatment threshold (121 cities –with an 80-41 split observed repeatedly for the 1996-2008 period) to draw precise estimates of the causal impact of the electoral rule on party structure.

Evidence of Strategic Voting

A Regression Discontinuity Test of Strategic Voting and Duverger's Law 207

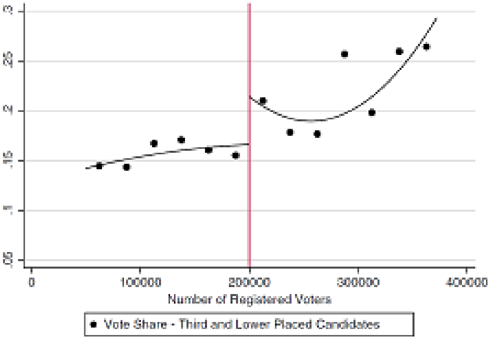

Figure 1. Vote share of third and lower placed candidates — local averages and parametric fit.

Evidence of Strategic Voting

- As predicted by Duverger's Law, a change from Single Ballot to runoff elections:
- Increases voting for the third placed (& lower-placed) candidates by 8.8 ppt (from 15 ppt under Single Ballot);
- Decreases the vote margins between second & third and the vote margins between first & third placed candidates, while does not affect the margin between first and second placed candidates;

Evidence of Strategic Voting

- Results are stronger in closely contested races, in which incentives to vote strategically in Single Ballot systems are higher.
- Mayoral elections contemporaneous to council elections, but no change in rule for council at 200,000. No change in skills of mayor or councilmen at threshold. Paper shows that these results are not likely driven by selection of different types of candidates across electoral systems.

- Typical RDD caveat: 120 municipalities out of ~5000 in total. External validity.