Parties & Legislatures

Prof. Francesco Trebbi

UC Berkeley – Haas School of Business

Goals of the Lecture

1. Empirical Organizational Economics traditionally focuses on firms. Detailed organizational/personnel data at the firm level is occasionally available, but not easy to access (confidentiality), nor necessarily randomly sampled (external validity)
2. But **data** is crucial to push forward, inform, test future Applied Theory work
3. In this lecture I will try to push forward an idea: focus on Political Organizations
4. **Political parties** offer plenty of data, relevant mechanism (agency, hierarchy, non-market features), and external validity
5. Obviously the study of government bureaucracy & public administration is relevant to this discussion, but I won't focus on these formal structures here

Goals of the Lecture

- First of all, let us **understand the context**. We begin by providing an application of structural methods (i.e. empirical approaches based on recovering primitives of the problem, such as preferences, beliefs, or fundamental elasticities)
- In order to represent our agents (politicians), we will focus on a formal analysis of voting behavior in **legislatures** applying what we call a "spatial voting" approach in Political Economy
- We focus on "sincere voting" & derive the structure at the basis of the most popular preference estimation approach followed in political science & political economy: DW-Nominate (Poole and Rosenthal, 1984; Poole and Rosenthal, 1997; McCarty, Poole and Rosenthal, 2006). This is the linchpin of modern discussion of political polarization (i.e. not a trivial application - we are not estimating preferences for bike tires...)
- Only then we will move to more realistic & complex environments where **political organizations** play a role (Canen, Kendall and Trebbi, ECMA 2020; Canen, Kendall and Trebbi, 2021)

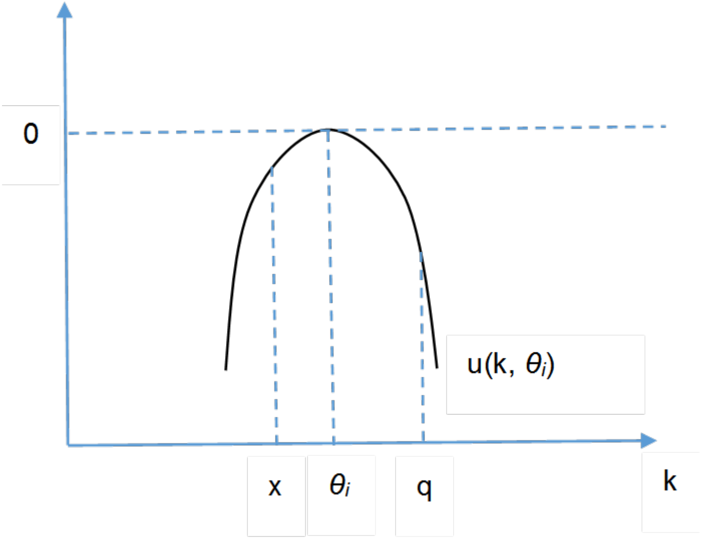
The Nominate Approach

- One of the few example of structural econometrics in political science
- What “spatial voting” means is simple: every politician will be endowed with certain policy preferences defined as a point in a multidimensional policy space & they will chose between two policy options based on which one is closer to their ideal point
- It starts from thinking about the policy space as some low-dimensional space. Say, the real line or the 2D Cartesian space. Make it oriented so that higher values mean more conservative (right direction) & lower values more liberal (left direction)

Preferences of the Politician

- Suppose policy space is the real line (1D) & define a distance d. A metric space $(\mathbb{R}, d)$.
- Policy space does not need to be uni-dimensional, but easier for now.
- Each politician has ideology $\theta_i \in \mathbb{R}$ & evaluates her utility from policy alternative $k \in \mathbb{R}$
- Deterministic part of utility function $u(k, \theta_i) = - \parallel k - \theta_i \parallel$ where $\parallel . \parallel$ is distance function.
- For example, a simple quadratic loss function $u(k, \theta_i) = -(k - \theta_i)^2$ or something fancier like $-|k - \theta_i|^\xi$

Preference for alternative x relative to status quo q



Preferences (Cont.)

- Note that we can add shocks to get a full random utility setting: $U_i(k) = -(k - \theta_i)^2 + \varepsilon_{ik}$ where the shock ε_{ik} is an i.i.d. random variable that hits politician i when she picks that specific k as her choice
- Think of it as a taste element that is random - not deterministic - & known only at the moment of choosing
- This means that i has a probability of choosing x over the status quo q (voting Yes to an alternative on the House floor) with probability:

$$\begin{aligned} Pr(i \text{ votes for } x) &= Pr\left(-(x-\theta_i)^2 + \varepsilon_{ix} \geq -(q-\theta_i)^2 + \varepsilon_{iq}\right) \\ &= Pr\left((q-\theta_i)^2 - (x-\theta_i)^2 \geq \varepsilon_{iq} - \varepsilon_{ix}\right) \\ &= CDF\left((q-\theta_i)^2 - (x-\theta_i)^2\right) \end{aligned}$$

& with CDF indicating the cumulative distribution function of the random variable $\varepsilon_{iq} - \varepsilon_{ix}$.

Interpreting Probabilities

- The interesting thing is that if you pick appropriate distributions for the shock ε_{ik} a likelihood of the vote choice made by i can be formulated in closed form
- ...and because the N choices are made independently (ε are iid) you can express the likelihood of observing your entire sample (sample likelihood) as a product of marginal probabilities (likelihood contribution of each i multiplied together)
- Then you maximize such sample likelihood with respect to the parameters (i.e. MLE)
- ...and essentially from that you can estimate the whole set of ideal points $\Theta = \{\theta_i\}$

Voting Likelihood

- For clarity, suppose the CDF is Standard normal $\Phi(.)$
- Say a dummy $Yes^i = 1$ vote by i indicates choice of x over q; and $Yes^i = 0$ choice of q over x
- You observe whether politician i voted Yes or No on the bill (i.e. $Yes^i = 0$)

Voting Likelihood (Cont.)

- Putting together all choices made by the politicians in the sample is possible because they are independent r.v.'s (shocks are iid)
- You just multiply the marginals to obtain the joint likelihood of observing that sample
- This is going to produce a sample likelihood function like this:

$$
\begin{aligned}
&\mathcal{L}(\Theta; Yes^i) = \\
&\quad \prod_{i=1}^{N} \Phi((q-\theta_i)^2 - (x-\theta_i)^2)^{Yes^i} \\
&\quad\quad \times \left(1 - \Phi\left((q-\theta_i)^2 - (x-\theta_i)^2\right)\right)^{1-Yes^i}
\end{aligned}
$$

Voting Likelihood (Cont.)

- There is a feature to clarify. How can I get one preference parameter θ_i for each i if I only see politician i vote once?
- You cannot. But fortunately politicians vote many times: $t = 1, ..., T$ during a (two year) congressional cycle. Thousands of choices for each i are observed.
- T is the number of roll call votes cast by each politician per congressional cycle
- You have a vote for each i on each bill t. Panel structure.

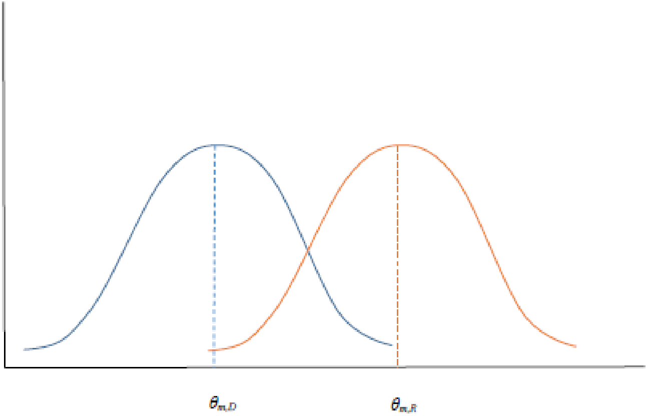
Voting Likelihood (Cont.)

- Then, the likelihood that one maximizes looks more like:

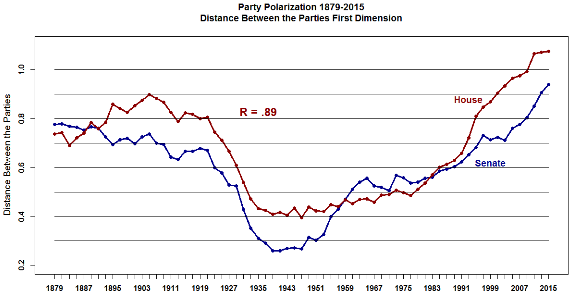
$$\mathcal{L}(\Theta; Q; X; Yes_t^i) = \prod_{t=1}^{T}\prod_{i=1}^{N}\Phi((q_t-\theta_i)^2-(x_t-\theta_i)^2)^{Yes_t^i} \times \left(1-\Phi\left((q_t-\theta_i)^2-(x_t-\theta_i)^2\right)\right)^{1-Yes_t^i}$$

- $\mathcal{L}(\Theta; Q; X; Yes_t^i)$ is identified & feasible with large N & large T (no nuisance parameter issue, see Fernandez-Val and Weidner, 2016)
- This is the clear representation of where Nominate scores come from (Θ). See Rivers (2003)

Elite Polarization: Distance of Party Medians



Elite Polarization in DW-Nominate



[1]

"Political polarization has reached levels not seen in decades, with nearly one-third of people in each party describing the other party as a threat to the nation's well-being. Trust in all institutions, including media, government, and business has fallen considerably." - T. R. Heath (2018)

[1] www.voteview.com

Actual DW-Nominate Likelihood

- Now I am going to show a problem with the actual DW-Nominate (not the simplified version so far) in 2D
- Consider the likelihood argument in Poole and Rosenthal (1997):

$$
\begin{aligned}
Pr(Y_{i,t}=1) &= \Phi\left[u(\theta_i, \mathrm{x}_t) - u(\theta_i, \mathrm{q}_t)\right] = \\
&\Phi\left[\beta e^{-\frac{1}{2}\left(\theta_i^1 - x_t^1\right)^2 - \frac{w_2}{2}\left(\theta_i^2 - x_t^2\right)^2} - \beta e^{-\frac{1}{2}\left(\theta_i^1 - q_t^1\right)^2 - \frac{w_2}{2}\left(\theta_i^2 - q_t^2\right)^2}\right]
\end{aligned}
$$

Not quadratic, but a Gaussian function of a quadratic loss in 2D. A normalization $\beta = 1$ is needed.

Actual DW-Nominate Likelihood (Cont.)

- This is a much more nonlinear function of the parameters than in our initial setup & this is not something you necessarily want in your structural estimation
- The problem? Identification
- Point identification is a property of an empirical model that requires a unique mapping between moments in the data and parameters to be estimated
- If multiple parameters correspond to the same empirical moment, then there is no point identification
- Sometimes you can work around this. E.g. MMI & set identification in Bombardini, Li & Trebbi (2020)

Identification Problems of DW-Nominate Structure

The vector of parameters of interest is $\Theta = \left\{\theta_i^1, x_t^1, q_t^1, \theta_i^2, x_t^2, q_t^2, w_2\right\}$. For $s > 0$ and $0 < r < 1$, define the following candidate (nonlinear) transformation, which can be proved is not a rotation (other than in the special case $w_2 = s = 1$):

$$\begin{aligned}
\tilde{\theta}_i^1 &= \theta_i^1 \sqrt{r} - \theta_i^2 \sqrt{w_2 (1-r)} \\
\tilde{x}_t^1 &= x_t^1 \sqrt{r} - x_t^2 \sqrt{w_2 (1-r)} \\
\tilde{q}_t^1 &= q_t^1 \sqrt{r} - q_t^2 \sqrt{w_2 (1-r)} \\
\tilde{\theta}_i^2 &= s \times \left(\theta_i^1 \sqrt{(1-r)} + \theta_i^2 \sqrt{w_2 r}\right) \\
\tilde{x}_t^2 &= s \times \left(x_t^1 \sqrt{(1-r)} + x_t^2 \sqrt{w_2 r}\right) \\
\tilde{q}_t^2 &= s \times \left(q_t^1 \sqrt{(1-r)} + q_t^2 \sqrt{w_2 r}\right) \\
\tilde{w}_2 &= \frac{1}{s^2}
\end{aligned}$$

Identification Fails

You can check that within this class of transformations one obtains the same likelihood:

$$\Phi\left[\beta e^{-\frac{1}{2}\left(\tilde{\theta}_i^1-\tilde{x}_t^1\right)^2-\frac{\tilde{w}_2}{2}\left(\tilde{\theta}_i^2-\tilde{x}_t^2\right)^2}-\beta e^{-\frac{1}{2}\left(\tilde{\theta}_i^1-\tilde{q}_t^1\right)^2-\frac{\tilde{w}_2}{2}\left(\tilde{\theta}_i^2-\tilde{q}_t^2\right)^2}\right]=$$
$$\Phi\left[\beta e^{-\frac{1}{2}\left(\theta_i^1-x_t^1\right)^2-\frac{w_2}{2}\left(\theta_i^2-x_t^2\right)^2}-\beta e^{-\frac{1}{2}\left(\theta_i^1-q_t^1\right)^2-\frac{w_2}{2}\left(\theta_i^2-q_t^2\right)^2}\right]$$

Proof

$$\left(\tilde{\theta}_i^1-\tilde{x}_t^1\right)^2+\tilde{w}_2\left(\tilde{\theta}_i^2-\tilde{x}_t^2\right)^2=$$

$$\left(\theta_i^1\sqrt{r}-\theta_i^2\sqrt{w_2\left(1-r\right)}-x_t^1\sqrt{r}+x_t^2\sqrt{w_2\left(1-r\right)}\right)^2$$

$$+\frac{1}{s^2}\left(s\times\left(\theta_i^1\sqrt{(1-r)}+\theta_i^2\sqrt{w_2r}\right)-s\times\left(x_t^1\sqrt{(1-r)}+x_t^2\sqrt{w_2r}\right)\right)^2=$$

$$\left(\left(\theta_i^1-x_t^1\right)\sqrt{r}-\left(\theta_i^2-x_t^2\right)\sqrt{w_2\left(1-r\right)}\right)^2$$

$$+\left(\left(\theta_i^1-x_t^1\right)\sqrt{(1-r)}+\left(\theta_i^2-x_t^2\right)\sqrt{w_2r}\right)^2=$$

$$\left(\theta_i^1-x_t^1\right)^2r+\left(\theta_i^2-x_t^2\right)^2w_2\left(1-r\right)$$

$$-2\left(\theta_i^1-x_t^1\right)\sqrt{r}\left(\theta_i^2-x_t^2\right)\sqrt{w_2\left(1-r\right)}+\left(\theta_i^1-x_t^1\right)^2\left(1-r\right)$$

$$+\left(\theta_i^2-x_t^2\right)^2w_2r+2\left(\theta_i^1-x_t^1\right)\sqrt{(1-r)}\left(\theta_i^2-x_t^2\right)\sqrt{w_2r}=$$

$$\left(\theta_i^1-x_t^1\right)^2+w_2\left(\theta_i^2-x_t^2\right)^2$$

Identification Problems (Cont.)

- This proves that W-Nominate in 2 dimensions is not identified up to this class of transformations
- This is broader than the class of transformations up to rotation, scale, and location change (a priori the problem is not identified unless one fixes rotation, scale and location, but this is trivial in $\mathbb{R}^2$)

Identification Problems (Cont.)

- This new class of transformations is particularly damaging. Consider three politicians $i = a, b, c$ one more conservative than the other with respect to the first dimension, i.e. suppose $0 < \theta_a^1 < \theta_b^1 < \theta_c^1$.
- Suppose further that the ideal points of politicians a and c are known.
- We can show that, for an infinite set of values of r, either we can achieve the wrong order $\tilde{\theta}_b^1 < \theta_a^1 < \theta_c^1$ or the wrong order $\theta_a^1 < \theta_c^1 < \tilde{\theta}_b^1$ along the first dimension.

Why Getting it Right is Important: Further Motivation

Percentage of Foreign Policy Opinion Leaders Seeing Issue as Critical Threat			
	Republicans	Democrats	Independents
Political polarization in US	**71**	**74**	**74**
North Korea's nuclear program	**54**	**53**	**50**
Iran's nuclear program	**64**	28	33
Development of China as a world power	**52**	37	42
Decline of democracy around the world	35	**57**	42
Russian influence in US elections	41	**78**	45
Trade war with China	25	34	36
Drug-related violence and instability in Mexico	25	18	23
Large numbers of immigrants and refugees entering US	23	2	14
Economic competition from low-wage countries	0	9	10

Chicago Council on Global Affairs-Texas National Security Network Survey of Foreign Policy Opinion Leaders, August 2 – October 16, 2018 [2]

- Survey responses of 588 foreign policy opinion leaders

[2] D.Smeltz, J. Busby, and J. Tama, The Hill, 2018

Canen, Kendall, Trebbi (ECMA 2020)

- CKT paper sets to:
  (i) point identify parameters
  (ii) quantify the *sources* of political polarization
  (iii) determine how polarization affects policy *outcomes*
  (iv) clarify the role of agenda setting & selection on votes

Sources of Polarization

- Two main sources:
  - members' **ideological positions** themselves (McCarty, Poole, Rosenthal, 2006)
  - **party discipline** (Snyder and Groseclose, 2000; Cox and McCubbins, 2005)
- Difficulty separating the two is a well known problem (Krehbiel, 1993; 1999; 2000)
  - cohesion/party unity may reflect self-selection into parties
  - parties may only pursue **agendas**/bills on which they agree (Cox and McCubbins, 2005)
- Source is important:
  - party discipline/organization may be more amenable to change
  - differential effects on outcomes

What CKT does

- Provide a model of the legislative process from policy selection to roll-call votes
  - where votes on policy are the result of:
    1. heterogeneous ideologies
    2. party discipline
    3. agenda-setting
- Use new internal party records - whip counts - to identify key sources of party control:
  - whip counts provide information on ideology before discipline
  - presence of a whip count indicates the ‘value’ of a bill
- Structurally estimate model & perform counterfactual exercises to illustrate how polarization affects outcomes

Preliminary: Whip Counts

- Informal polls of party members typically taken a day or two before the roll call vote (Evans, 2018)
- e.g. Whip counts show that repeal of ACA won't have enough votes:

  With Democrats united in opposition, House Republicans are currently short of the 216 votes they need to pass the bill before the Senate could take it up. They can afford only 22 defections, and the latest whip counts put Republican "no" votes at about 20, with a dozen more undecided. - BBC

- e.g. On the Tax Bill, after roll call (it passed with 227 votes vs. 205, with 13 Republicans breaking rank):
  Ryan and House GOP leaders were confident throughout the week that they'd have the 218 votes needed for passage, even with unified Democratic opposition. In fact, they've felt so good about their whip count they barely called on the White House to twist arms. - Politico

Preliminary: Whip Counts

- Key assumption: Whip counts are *on average* truthful (Evans, 2013):

  "One common question about whip counts is whether the responses of members can be trusted...Four points are worth mentioning in response. First, the whip process is a "repeated game" and members develop reputations. There are incentives for them to be truthful. Second, congressional leaders generally know a lot about the constituencies of rank-and-file members and can be very difficult to fool. Third, in a sense it does not matter. If a member claims that she will oppose a bill or amendment unless she receives some concession, then that essentially becomes her position and the polled question and the concession are for all practical purposes inseparable. Fourth, and most important, participants in the whip process believe that whip poll responses are accurate, which is precisely why they base strategic decisions on the results."

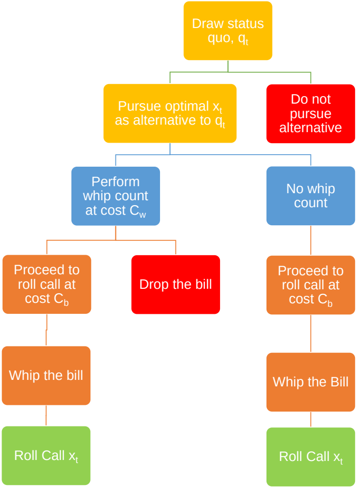
Literature

- Very large and important literature on estimating ideal points (Poole and Rosenthal, 1984;...)
- More closely related to that which attempts to separate out party effects (Jenkins, 2000; Snyder and Groseclose, 2000; Nokken, 2000; Clinton, 2004)
    - now we incorporate new data (whip counts) via a new theoretical, estimable framework
- Much smaller literature on the effects of polarization on policy (Binder, 2003; Mian et al., 2014)
    - here we provide theory & quantitative estimates

Setup

- Two parties, $p = D, R$, compete for votes over a series of bills
  - have preferences of their median members, $\theta_{m,D}$ and $\theta_{m,R}$
  - continuum of members i in each party
- One-dimensional ideological space w/ symmetric loss functions
  - bliss points, θ_i
  - $\|x_t - \omega_{i,t}\|$ where $\omega_{i,t} = \theta_i + \sum_{s=1}^{2} \varepsilon_{i,t}^s + \eta_t^s$
- Votes, and hence policy outcomes, are stochastic
  - idiosyncratic shocks, $\varepsilon_{i,t}$, & aggregate shocks, η_t (normally distributed)
    - with continuum of members, require aggregate shocks so that outcomes are uncertain & eq.m policy uniquely defined
    - aggregate shocks capture anything that affects overall perception of a bill (including changes to bill)

Timeline



Draw status quo, q_t
Pursue optimal x_t as alternative to q_t
Do not pursue alternative
Perform whip count at cost C_w
No whip count
Proceed to roll call at cost C_b
Drop the bill
Proceed to roll call at cost C_b
Whip the bill
Whip the Bill
Roll Call x_t
Roll Call x_t

Agenda-setting

- Random recognition model - each party is chosen to be the proposer with some probability γ
  - required to match empirical fact that a significant number of bills have majority leadership voting ‘no’ and minority leadership voting ‘yes’
- Proposing party:
  - observes a randomly drawn status quo policy issue, q_t
  - decides whether to pursue an alternative policy to q_t or drop it
  - if it does not drop the issue, party sets alternative, x_t
  - decides whether or not to conduct a whip count at cost, C_w
    - whip count allows proposer to learn about first aggregate shock & drop the bill if not looking promising
    - dropping the bill saves the cost of pursuing a bill at roll call, C_b
    - absent a whip count, goes straight to roll call vote

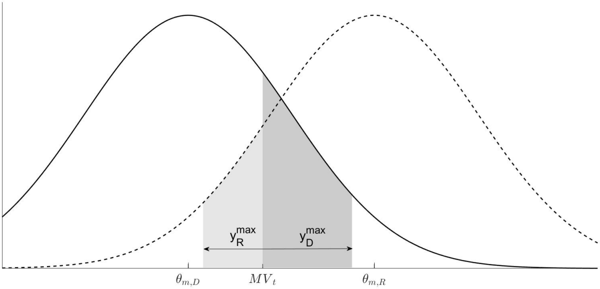
Voting Decisions

- Discrete-choice model as in DW-Nominate but with two key improvements:
  - shocks are on bliss points, θ_i, instead of utility
    - no need to specify utility function (other than concavity)
    - likelihood becomes a function of marginal voter, $MV_t = \frac{x_t + q_t}{2}$, rather than both q_t and x_t
  - bliss point is subject to influence from party through whipping, $y_{i,t}$, so $\|x_t - \omega_{i,t} - y_{i,t}\|$

Whips

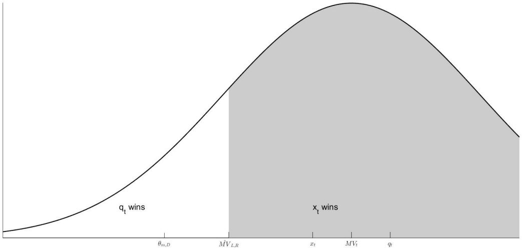
- Vote just as any other member
- Whips are assigned members for which they are responsible:
  - at roll call time, obtain information - know their members' (stochastic) bliss points
  - can exert influence at a personal cost, $c(y_{i,t})$, strictly increasing
  - obtain r_p any time a member votes as the party prefers
- Whips themselves are subject to being whipped

Party Discipline



- Key parameter of interest is maximum distance a whip is willing to influence members, $y_p^{max} = c^{-1}(r_p)$

Optimal Policy Alternatives

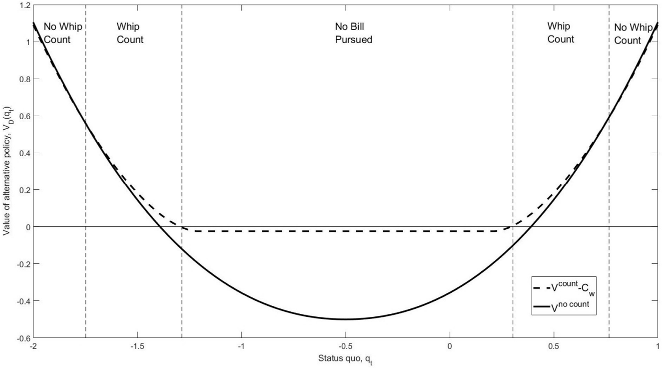


- If a policy alternative, x_t, is pursued, want to choose it close to the bliss point of the median member
  ...but, the closer it is, the less likely is the bill to pass
- Trade-off results in a unique optimal policy
  - always lies between status quo and party's bliss point

Whip Counts as Options

- On observing q_t, the proposing party can:
  1. do nothing
  2. pursue an alternative bill with a whip count
  3. pursue an alternative bill without a whip count
- Absent a whip count, bill goes straight to roll call & majority party pays C_b
- With a whip count (at cost C_w), bill can be dropped avoiding C_b
  - provides option value

Which Bills are Pursued



Proposition 1

There exists a strictly positive cutoff cost of pursuing a bill, $\hat{C}_b > 0$*, such that for all* $C_b < \hat{C}_b$*, the optimal alternative policies,* x_t^{count} and $x_t^{no\ count}$, *are unique and contained in* (q_t, θ_D^m) *for* $q_t < \theta_D^m$*, contained in* (θ_D^m, q_t) *for* $q_t > \theta_D^m$*, and equal to* θ_D^m *for* $q_t = \theta_D^m$*.*

Proposition 2

Fix $C_b < \hat{C}_b$ *such that the optimal alternative policies,* x_t^{count} and $x_t^{no\ count}$*, are unique and fix the cost of a whip count,* $C_w > 0$*. Then, we can define a set of cutoff status quo policies,* $\underline{q}_l, \overline{q}_l, \underline{q}_r$*, and* $\overline{q}_r$*, with* $\underline{q}_l \leq \overline{q}_l < \theta_D^m < \underline{q}_r \leq \overline{q}_r$ *such that:*

1. *for* $q_t \in [-\infty, \underline{q}_l] \cup [\overline{q}_r, \infty]$*, the optimal alternative policy,* $x_t^{no\ count}$*, is pursued without conducting a whip count.*
2. *for* $q_t \in (\underline{q}_l, \overline{q}_l] \cup [\underline{q}_r, \overline{q}_r)$*, the optimal alternative policy,* x_t^{count}*, is pursued and a whip count is conducted.*
3. *for* $q_t \in (\overline{q}_l, \underline{q}_r)$*, no alternative policy is pursued.*

Data

- U.S. House roll call voting data comes from the standard source, VoteView
- Whip count data covering 1977-1986 as compiled by Evans (2012)
  - Corresponds to time when polarization starts to rise
  - Democrats are majority over time period, but both parties conduct whip counts
  - Republican (1977-1980) data from Robert H. Michel Collection
  - Democratic (1977-1986) data from Congressional Papers of Thomas S. Foley
- We merge the data following Evans (2012)
  - $5,424$ roll called bills
  - 340 bills with whip counts
  - $238/340$ bills have subsequent roll calls

Identification (1)

- Key assumption is that whip counts reveal true ideological positions on average (i.e. cannot fool the party all the time)
  - if not revealing, whip counts would be uninformative... but parties do rely on them
  - reputation prevents lying
  - there is a reason why such accurate records were kept
  - deputy whips have detailed knowledge about members' positions (little info asymmetry)

Identification (2)

- Ideological positions come from repeated whip count polls (individual fixed effects)
- Marginal voters at time of whip count & time of roll call come from multiple reports/votes on same bill (bill fixed effects/cutoffs)
- Maximum whipping distance, y_p^{max}, comes from distance between marginal voter at time of whip count and *per party* marginal voter at roll call
  - identify direction of whipping from leadership votes
- Distributions of policies (q_t and thresholds) come from distributional assumptions + whip counts dropped

Estimation

- Two-step process (maximum likelihood in each step):
  1. estimate marginal voters, $\tilde{MV}_t$, party discipline parameters, y_p^{max}, and ideological bliss points, θ_i
     - we use *all* bills
  2. estimate flexible status quo distribution to fit estimated marginal voters
     - status quo drawn from truncated normal
     - impose model restrictions:
       leadership votes determine where status quo originated
       whip counts closer to party median
     - first-order conditions relate q_t to $\tilde{MV}_t$ (bills with roll calls only)
     - extensive Monte Carlo simulation to demonstrate truncations are recoverable

Deriving Likelihood

Under Assumption 2, the probability i from D votes Yes at the whip count:

$$\begin{aligned} P(Yes_t^{i,wc} = 1) &= P(\varepsilon_{1,t}^i + \theta^i \leq MV_t - \eta_{1,t}) \\ &= P(\varepsilon_{1,t}^i \leq \tilde{MV}_{1,t} - \theta^i) \\ &= \Phi(\tilde{MV}_{1,t} - \theta^i). \end{aligned}$$

At the roll call stage:

$$\begin{aligned} P(Yes_t^{i,rc} = 1) &= P(\varepsilon_{1,t}^i + \varepsilon_{2,t}^i \leq MV_t - \eta_{1,t} - \eta_{2,t} - \theta^i \pm y_D^{max}) \\ &= P(\varepsilon_{1,t}^i + \varepsilon_{2,t}^i \leq \tilde{MV}_{2,t} - \theta^i \pm y_D^{max}) \\ &= \Phi\left(\frac{\tilde{MV}_{2,t} - \theta^i \pm y_D^{max}}{\sqrt{2}}\right). \end{aligned}$$

Likelihood (1)

First step:

$$\mathcal{L}_D(\Theta_{\mathbf{1}}; Yes_{t,p}^{i,wc}, Yes_{t,p}^{i,rc}) =$$

$$\prod_{t=\mathbf{1}}^{T} \prod_{n=\mathbf{1}}^{N_D} \Phi(\tilde{M}V_{\mathbf{1},t} - \theta^i)^{Yes_{t,p}^{i,wc}} \left(\mathbf{1} - \Phi(\tilde{M}V_{\mathbf{1},t} - \theta^i)\right)^{\mathbf{1} - Yes_{t,p}^{i,wc}}$$

$$\times \Phi\left(\frac{\tilde{M}V_{\mathbf{2},t} - \theta^i \pm y_D^{max}}{\sqrt{2}}\right)^{Yes_{t,p}^{i,rc}} \left(\mathbf{1} - \Phi\left(\frac{\tilde{M}V_{\mathbf{2},t} - \theta^i \pm y_D^{max}}{\sqrt{2}}\right)\right)^{\mathbf{1} - Yes_{t,p}^{i,rc}}$$

Likelihood (2)

$$\mathcal{L}^{second\ \ step}(\Theta_1; \tilde{WC}_t, \tilde{MV}_{2,t}) = \prod_{t=1}^{T} P(WC_t)^{WC_t} P(\tilde{MV}_{2,t})^{RC_t}$$

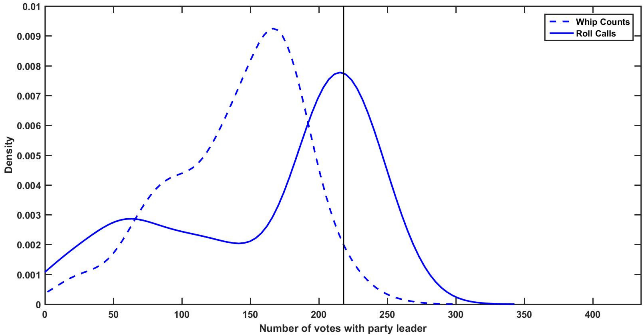
For example, for a whip count for a status quo to the right of a party's median, we have, using Proposition 2:

$$P(WC_t) = \frac{\Phi(\frac{\overline{q}_{r,p}-\mu_q}{\sigma_q}) - \Phi(\frac{\underline{q}_{r,p}-\mu_q}{\sigma_q})}{P(WC_t \cup RC_t)}$$

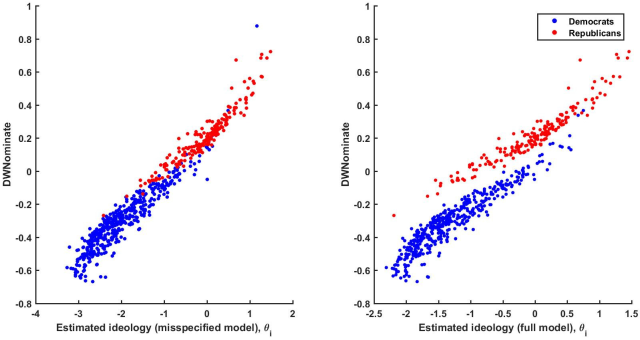
For example, the probability of observing a particular realized marginal voter for a status quo drawn from the right of the Democrats median (conditional on observing either a whip count or roll call) is:

$$P(\tilde{MV}_{2,t}) \quad = \quad \int_{\overline{q}_{r,D}}^{\infty} \phi\left(\frac{\tilde{MV}_{2,t} - MV(q_t)}{\sigma}\right) \frac{\phi\left(\frac{q_t-\mu_q}{\sigma_q}\right)}{P(WC_t \cup RC_t)} dq_t$$

Party Discipline - Reduced Form

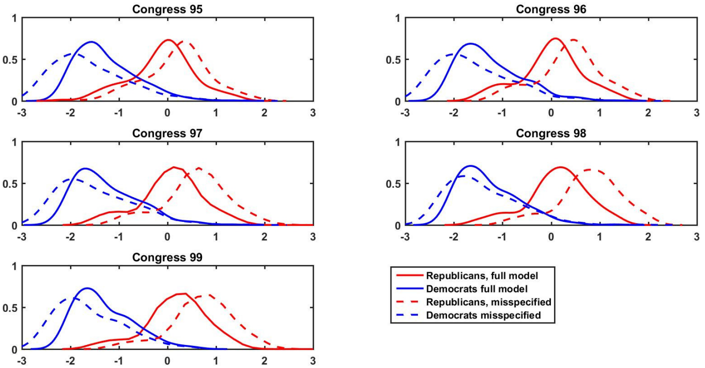


Ideologies (1)



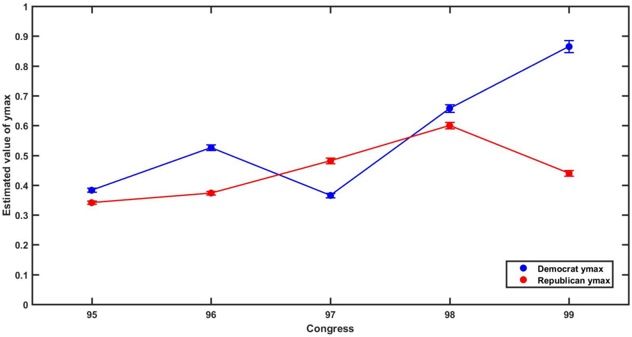
- Correlation between our estimates and DW-Nominate
  - strong, but not perfect, correlation
  - noticeable 'gap' introduced by party discipline (right graph)

Ideologies (2)



- 34% to 43% of *perceived* polarization is due to party discipline

Party Discipline Estimates



First Step Estimates

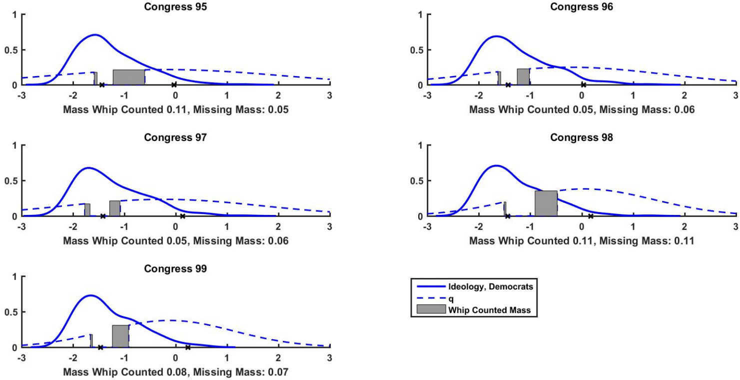
Parameter	Congress				
	95	96	97	98	99
y^{max}, Democrats	0.383	0.526	0.366	0.658	0.865
	(0.002)	(0.003)	(0.003)	(0.005)	(0.007)
y^{max}, Republicans	0.342	0.373	0.482	0.600	0.440
	(0.003)	(0.003)	(0.004)	(0.005)	(0.004)
Aggregate Shock, σ_η			0.859		
			(0.230)		
Party Median - Democrats, θ_D^m	-1.431	-1.431	-1.420	-1.435	-1.462
	(0.038)	(0.038)	(0.042)	(0.040)	(0.095)
Party Median - Republicans, θ_R^m	-0.036	0.042	0.134	0.181	0.236
	(0.049)	(0.138)	(0.139)	(0.034)	(0.049)
N: 711, T: 315 Whip Counted bills, 5424 Roll Called bills					

Which Bills are Whip Counted?

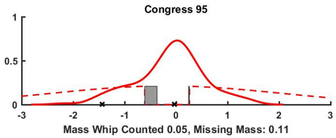
Distance from Marginal Voter to Party Median			
	Whip count	Roll call	p-value
Democrats	0.479	1.234	(0.000)
Republicans	0.910	1.163	(0.010)

- Model predicts whip counts are conducted for policies closer to the party's median (more difficult to pass)

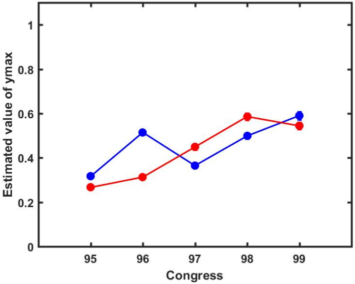
Agenda-Setting (Democrats)



Agenda-Setting (Republicans)



Robustness: e.g. No Whipping on Lopsided Bills



Salient Bills

- How would the outcomes of votes on important bills have changed if parties exercised no discipline?
  - hold the policies themselves fixed

Salient Bills - Economic Policies

Bill	Data	Model	No Whipp
Increase of Temporary Debt Limit, (H.R.9290, Congress 95)	221	242	185
Increase of Temporary Debt Limit, (H.R.13385, Congress 95)	210	235	201
Increase of Temporary Debt Limit, (H.R.2534, Congress 96)	220	239	208
Depository Inst. Dereg. and Monetary Ctrl. Act of 1980, (H.R. 4986, Congress 96)	369	404	391
Inc. of Public Debt Limit, Make it part of Budget Process (H.R. 5369, Congress 96)	225	244	217
Economic Recovery Tax Act of 1981 (H.R. 4242, Congress 97)	284	329	276
Garn-St. Germain Depository Institutions Act of 1982 (H.R.6267, Congress 97)	263	279	327
Social Security Amendments of 1983 (H.R.1900, Congress 98)	282	299	230
Tax Reform Act of 1984 (H.R. 4170, Congress 98)	319	370	292

Salient Bills - Other

Bill	Data	Model	No Whipping
Aid to Turkey/Lifting of Arms Embargo (H.R. 12514, Congress 95)	212	193	147
Foreign Intelligence Surveillance Act of 1978 (H.R. 7308, Congress 95)	261	283	280
National Energy Act, 1978 (H.R. 8444, Congress 95)	247	271	258
Panama Canal Treaty, 1979 (H.R. 111, Congress 96)	224	243	180
Contra Aid, 1984 (H.R. 5399, Congress 98)	294	279	343

Policies Pursued

- Absent party discipline, the optimal policies pursued x_t would have been different
- Look at two counterfactuals, accounting for change in policies themselves:
  1. No party discipline
  2. Increase in ideological polarization (to DW-Nominate levels)
- Look at average effects because we don't know status quo or alternative for any particular bill

Bill Approval

	Congress				
	95	96	97	98	99
Average Change in the Probability of Bill Approval					
Democrats					
Baseline Probability (Main Model)	0.357	0.467	0.421	0.431	0.544
Main Model - No Whipping	0.032	0.060	0.009	0.054	0.011
Main Model - Polarized Ideology	-0.005	-0.011	0.010	-0.013	-0.024
Republicans					
Baseline Probability (Main Model)	0.240	0.220	-	-	-
Main Model - No Whipping	-0.034	-0.042	-	-	-
Main Model - Polarized Ideology	0.028	0.032	-	-	-

- Absent whipping, majority party is less likely to pass a bill, minority party more likely

Policies Pursued

	Congress				
	95	96	97	98	99
Average Change in Pursued Policy Location, x_t					
Democrats					
Main Model - No Whipping	-0.011	-0.018	-0.003	-0.024	-0.042
Main Model - Polarized Ideology	0.085	0.161	0.107	0.163	0.285
Republicans					
Main Model - No Whipping	-0.011	-0.016	-	-	-
Main Model - Polarized Ideology	-0.057	-0.048	-	-	-

- Increase in ideological polarization results in more extreme policies: farther left for Democrats, farther right for Republicans

Conclusions on CKT

- CKT find that approximately 40% of polarization is due to party discipline
  - institutional changes may reduce party power
- The effects of polarization are complex due to the endogeneity of policies
  - a reduction in party discipline reduces the probability of bills passing
  - a reduction in ideological polarization results in less extreme bills being proposed

Issues with CKT

- Sample period constrained to 1977-1986. We need to generalize: Is party discipline the story for political polarization in the 20th century and now? History of political organizations in the US
- DW-Nominate score allows for *two* policy dimensions. 2D is important for the whole period between 1950-1980 with the Southern Democrats:
    1. Liberal-Conservative
    2. Civil rights - Voting Rights
- See Canen, Kendall and Trebbi (2021) a.k.a. CKT2

Relaxing Data Requirements

- The typical analysis of political polarization extends to the late 1800's
- We do not have internal party whip counts for the entire period (Evans 2018)
- We can lever on additional data though: The direction of voting by the organizational leadership
- It tells us which way the parties are whipping
- There are three main cases to consider: (i) D leadership votes Yes & R leadership votes No; (ii) D leadership votes No & R leadership votes Yes; (iii) D leadership votes Yes & R leadership votes Yes

1. New multidimensional model w/ party & agenda setting. Constructive proof of identification.
2. Discipline is statistically significant in every Congressional cycle since 1927
   - model selection tests (Vuong test, $p - value < 0.001$, each cycle)
   - discipline accounts for approx. 70% of voting polarization in 2018
3. Ideological polarization monotonically increases
4. No asymmetry
5. Robust to agenda setting/discipline specification
6. Offer comparison to DW-Nominate (no identification result)

Setup

Standard random utility model with unbounded shocks (e.g. DW-Nominate):

- Two parties, $p \in \{R, D\}$, compete for votes over a series of bills $t = 1, ..., T$
- Arrival of status quo $\bar{q}_t$ & decision of policy alternative to pursue (or not) $\bar{x}_t$
- Two-dimensional ideological space
    - bliss points, $\bar{\theta}^i = (\theta_1^i, \theta_2^i)$
- Preferences for policy $\bar{k}_t = \bar{q}_t, \bar{x}_t$:

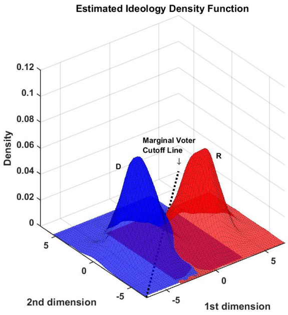
$$u(\bar{k}_t, \bar{\theta}^i) = u\left(\| \underbrace{\bar{\theta}^i + \bar{\varepsilon}_t^i}_{\equiv \omega_t^i} + \underbrace{\bar{y}_t^i}_{party} - \bar{k}_t \| \right), \ with\ u'(\cdot) < 0$$

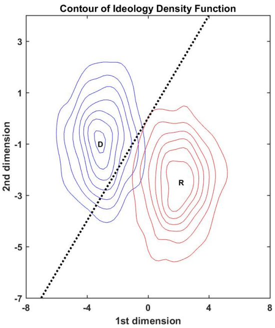
Absent party whipping, i votes for policy $\bar{x}_t$ over status-quo $\bar{q}_t$ if:

$$u(\|\bar{\theta}^i + \bar{\varepsilon}^i_t - \bar{x}_t\|) \geq u(\|\bar{\theta}^i + \bar{\varepsilon}^i_t - \bar{q}_t\|) \Leftrightarrow \|\bar{\theta}^i + \bar{\varepsilon}^i_t - \bar{x}_t\| \leq \|\bar{\theta}^i + \bar{\varepsilon}^i_t - \bar{q}_t\|.$$

The set of members who vote for $\bar{x}_t$ are those above the cutline:

$$\omega_{2,t} = m_t(\bar{x}_t, \bar{q}_t)\omega_{1,t} + b_t(\bar{x}_t, \bar{q}_t).$$





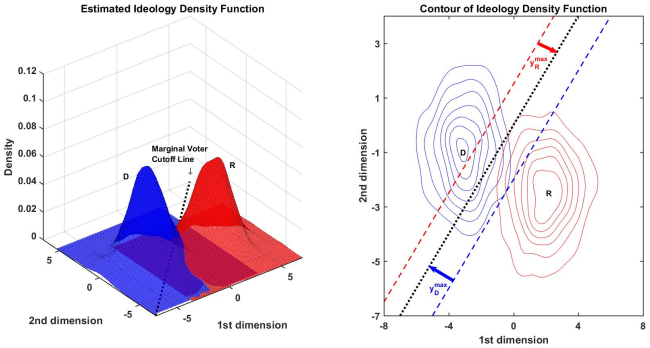
Key Model Assumptions

1. Shocks, $\bar{\varepsilon}^i_t$, are ideological
   $\Rightarrow$ *no need to specify utility function* $u(\cdot)$, *or to recover policy points* $\bar{x}_t, \bar{q}_t$ *(only a cutline & direction)*
2. Shocks move ideal points orthogonally to the separating cutline (projections of more general 2D shocks)
3. Shocks are unbounded & realize after agenda choice $\bar{x}_t(\bar{q}_t)$ by the proposer
   $\Rightarrow$ *Agenda-setting process needs not to be specified*
4. Discipline cost is borne by party whips
   $\Rightarrow$ *i whipped if distance to cutline is smaller than* y_p^{max}

Example. If party D whips for $\bar{x}_t$, and $x_{2,t} > q_{2,t}$, then those above:

$$\omega_{2,t} = m_t(\bar{x}_t, \bar{q}_t)\omega_{1,t} + b_t(\bar{x}_t, \bar{q}_t) - y_{p,t}.$$

vote for $\bar{x}_t$, where $y_{p,t} \equiv y_p^{max}\sqrt{1 + m_t^2}$.



Vote Probabilities for MLE

$$\Pr\left(Y_{it}=1|\bar{q}_t\in Q_p^1,\bar{x}_t;y_p^{max}\right)$$
$$=\begin{cases}\Phi\left(\sqrt{\frac{1}{1+m_t^2}}\left(\theta_2^i-m_t\theta_1^i-b_t\right)+W_{p,t}\times y_p^{max}\right) & \text{if } x_{2,t}>q_{2,t}\\ 1-\Phi\left(\sqrt{\frac{1}{1+m_t^2}}\left(\theta_2^i-m_t\theta_1^i-b_t\right)+W_{p,t}\times y_p^{max}\right) & \text{if } x_{2,t}<q_{2,t},\end{cases}$$

Goal is to recover all $\bar{\theta}^i$, b_t, m_t, vote directions, & party discipline parameters, y_p^{max}.

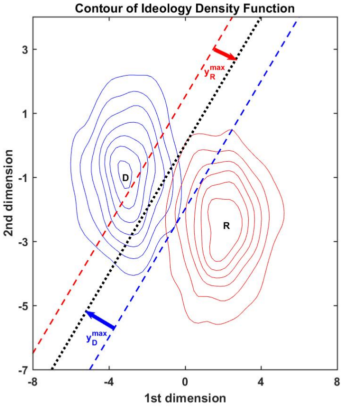
Identification

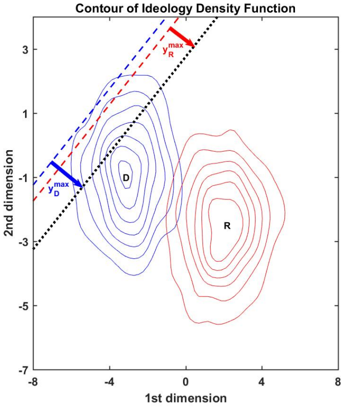
Assumptions ID:

1. The ideal points, $\left\{\left(\theta_1^i, \theta_2^i\right)\right\}$, are not perfectly collinear within at least one party.
2. (i) There exists a politician 0 s.t. $\bar{\theta}^0 = (0,0)$. (ii) There exists k with known first dimension, θ_1^k.
3. (i) There exists a bill 0 such that $m_0 = 0$. (ii) There exists a bill s with cutline slope $m_s \neq 0$.
   $\Rightarrow$ *with ID1, ensures two dimensions exist; with ID2, pins down location & rotation*
4. Parties D and R whip in the same direction on at least one bill, and in opposite directions on at least one other bill.
   $\Rightarrow$ *key variation to identify discipline*

Estimation via MLE. Large set of initial values. Monte Carlo validated.

Assumption ID(4) in 2D





Heuristic Intuition for the Estimator

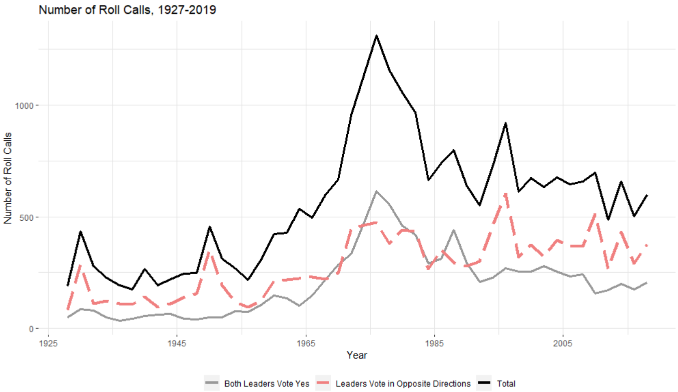
1. The direction of the leadership vote $W_{p,t}$ is known. Part of the roll call data.
2. Suppose that you could observe the "ideal point" of member i when she is whipped to the left & you could also observe the "ideal point" of member i when she is whipped to the right
3. Assuming whipping is symmetrical, if you take the midpoint, you recover information on the true ideal point
4. *Caveat*: Not an exact intuition (our problem is multidimensional & whipping is common across i), but gives a sense of how the estimator is going to interpret within-individual changes in voting subject to party pressure along certain directions

Assumptions on Whipping

- For baseline model, assume all bills are whipped
- Alternatives:
  1. no whipping when D & R party leaders agree (based on party leadership votes). Note: you lose information on party parameters, but still have info on Democrat and Republican with similar "ideal points" that whipped in opposite directions
  2. drop bills where leader & whip vote in opposite directions (Cox and McCubbins, 2005)
  3. no whipping on lopsided bills, i.e. 65-35 or 70-30 splits (Snyder and Groseclose, 2000)

Data

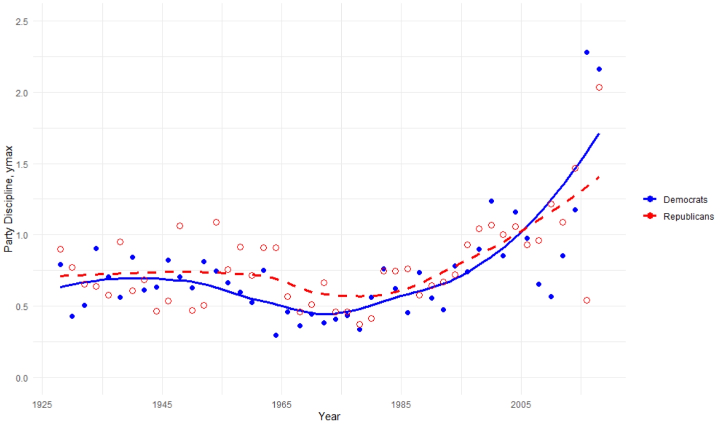
- Roll call data from Voteview
- Whipping directions from leadership votes (Cox and McCubbins, 1993)
- Today, focus on 2D Senate (see paper for 1D, House & Senate)



Details on Whipping Directions

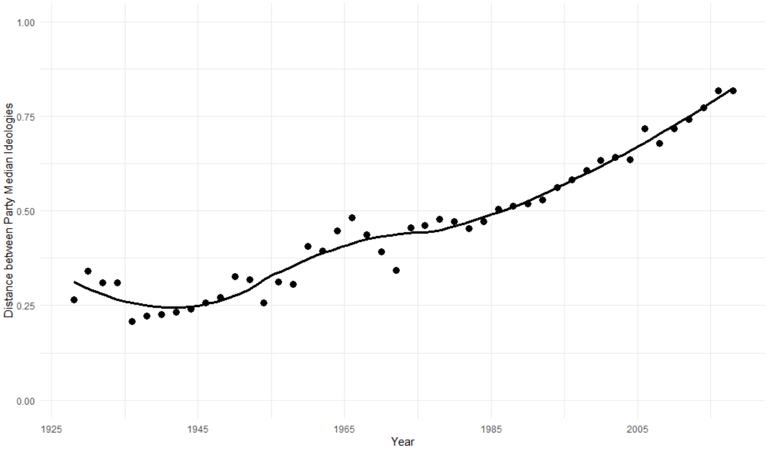
- To determine the whipping directions, $W_{p,t}$, we make use of leadership votes in the roll calls
- For each roll call vote, we code whether the party leadership voted Yes or No using the decisions by the Majority and the Minority Leader
- When such votes are unavailable, we use the Majority or Minority Whip's vote instead, and when that is also missing, the direction of the vote of the majority of the party

Party Discipline (y_p^{max})

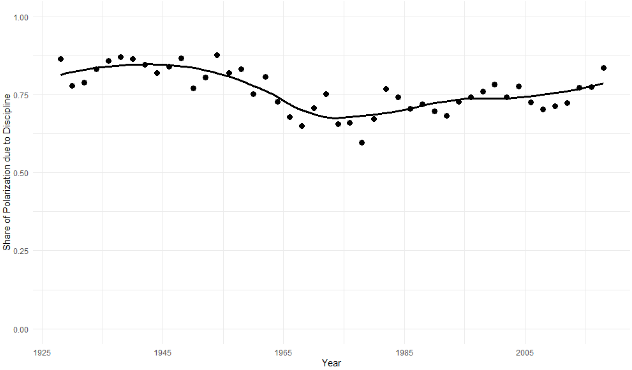


▸ Model Fit

Polarization (first dimension)

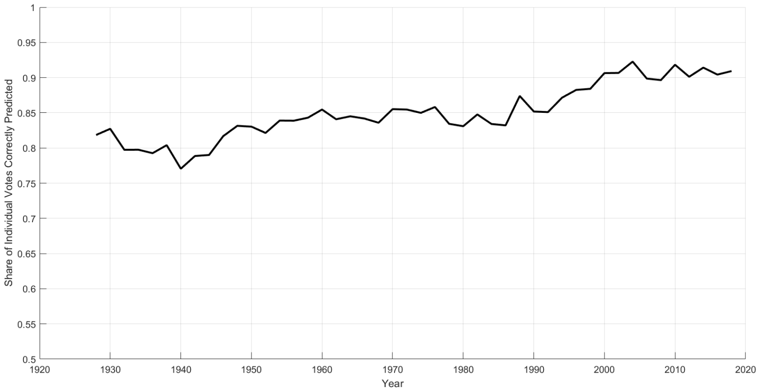


Fraction of Polarization Due to Discipline



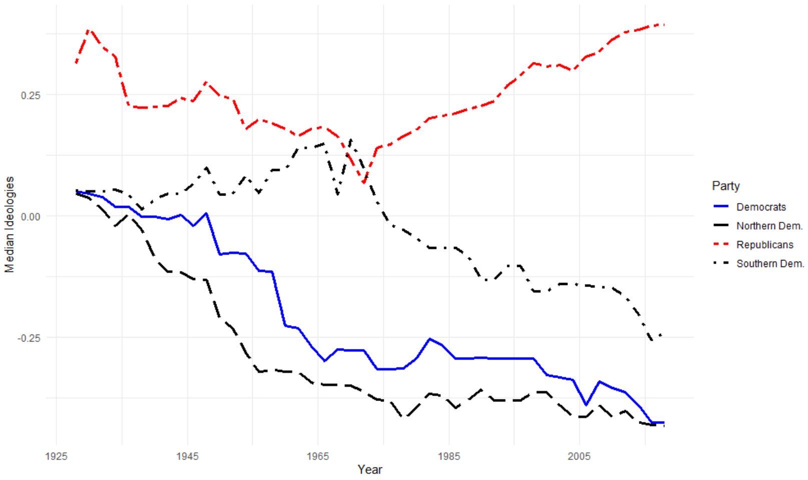
$$\text{share} = \frac{y_D^{max} + y_R^{max}}{y_D^{max} + y_R^{max} + (\theta_{m,R} - \theta_{m,D})}$$

Model Fit

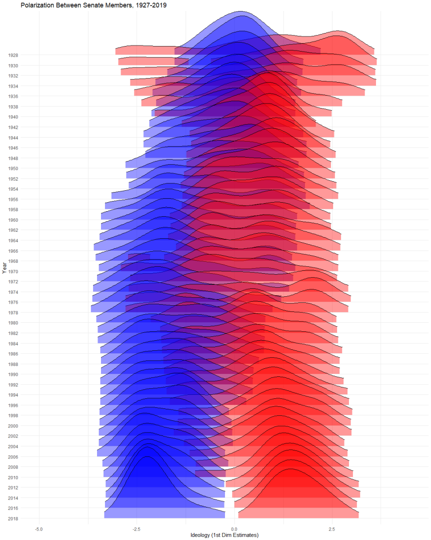


Share of Individual Votes Correctly Predicted
1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
1920
1930
1940
1950
1960
1970
1980
1990
2000
2010
2020
Year

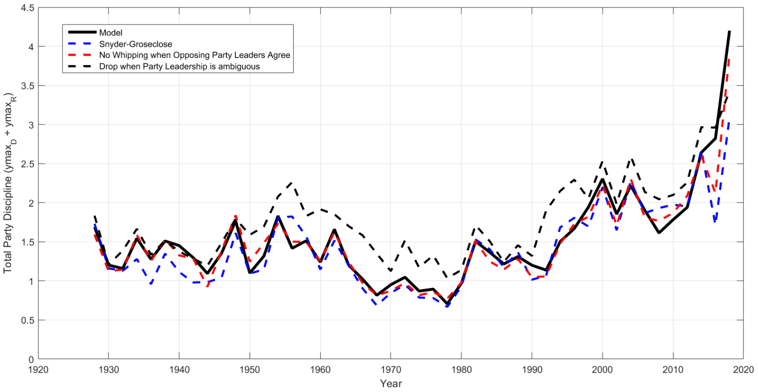
Asymmetric Polarization?



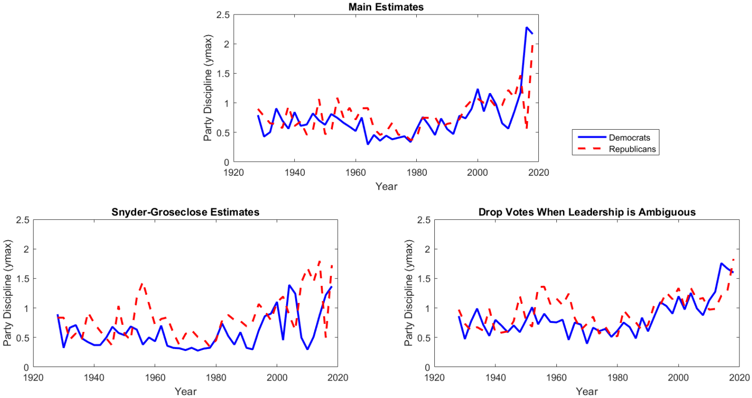
Asymmetric Polarization?



Other Whipping Assumptions (1)



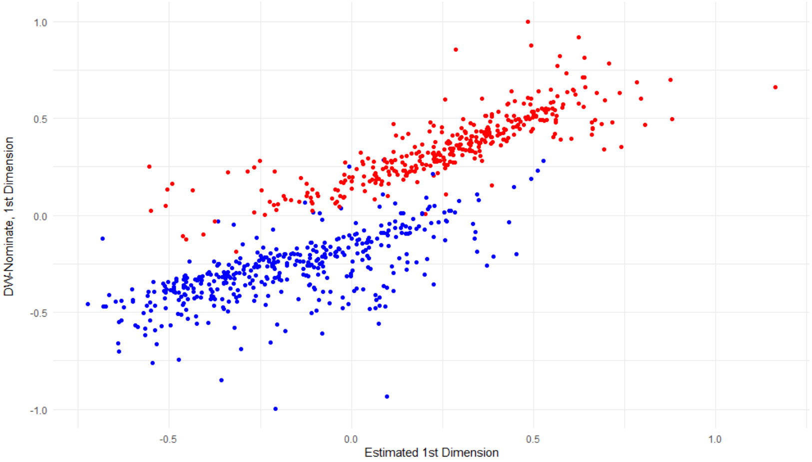
Other Whipping Assumptions (2)



Additional Robustness

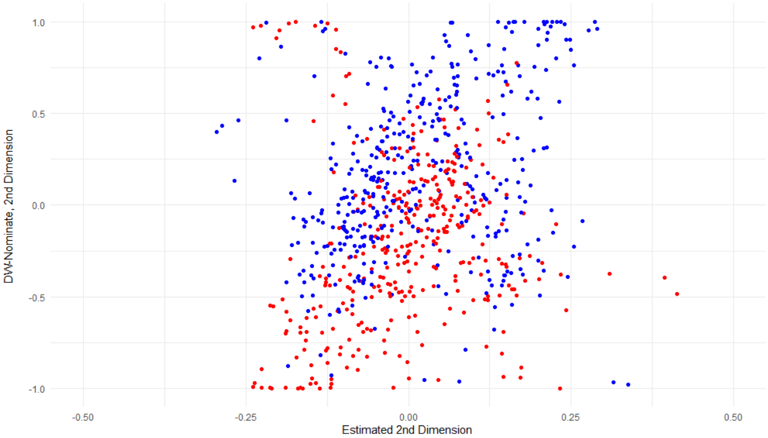
- Compare our estimates of party discipline with estimates from Canen et al., 2020 using alternative source of identification (whip counts) over sub-period 1977-1986:
    - estimates match qualitatively (picks up the uptick)
    - linear correlation between estimates is 88%
    - estimates match quantitatively for both parties (about 45% of polarization due to discipline in 1D model House)

Comparison to DW-Nominate (first dimension)



Note: We use Procrustes transformation to map our estimates onto DW-Nominate space

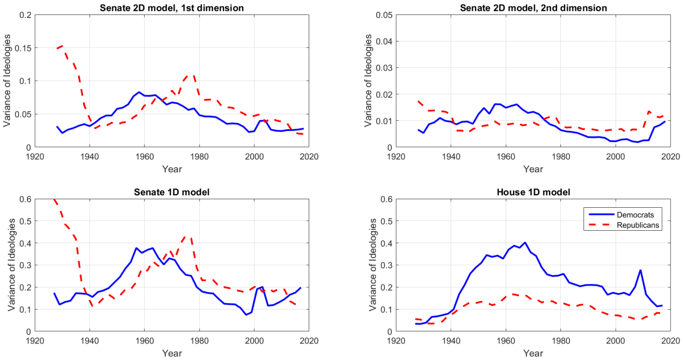
Comparison to DW-Nominate (second dimension)



Parties Matter

- Party leadership matters (statistically & quantitatively):
  - pre-war: parties of moderate strength, *but* ideologies not very polarized
  - 1960's: low party strength
  - today: unprecedented party strength *and* ideological polarization
- It's not simply voters electing extremists
  - timing of uptick suggests may even go other way
  - strategy of using divisive language drives affective polarization? (Iyengar and Westwood, 2015; Iyengar et al. (2019), Boxell et al. (2020)) After all, both Nancy Pelosi & Mitch McConnell were both Whips before becoming party leaders
- Time series also allows tests of theories of party behavior

Conditional Party Government (CPG)



- Negative correlation of within party *Var*(*ideologies*) & party discipline
- Consistent with Conditional Party Government (Aldrich, 1995; Rohde, 1991)
- Heterogeneous parties endogenously weaker

Other Ideas to Explore

- Majority/Minority party status
- Stability of discipline technology/whipping
- Effect of rule changes
- Executive

Conclusions on CKT2

- Still preliminary, but CKT2 shows promise as improvement over DW-Nominate
- Worked it out in 2D. This part of Canen, Kendall and Trebbi (2021) is available for you, if you want to dig deeper into these structural methods
- Historical record in Congress appears to support these findings (external validation is important for structural models - if the results are not reasonable, chances are the model can be called into question)