Political Parties as Drivers of U.S. Polarization: 1927-2018 (Appendix)

Nathan Canen, Chad Kendall, and Francesco Trebbi

January 2026

Appendix A: Identification

This Appendix proves the Identification of our model in two dimensions under the following assumptions.

Assumptions ID:

1. The set of ideal points, $\{(\theta_1^i, \theta_2^i)\}_{i=1}^N$, is not perfectly collinear within at least one party.
2. (i) There exists a politician 0 such that $\bar{\theta}^0 = (0, 0)$. (ii) There exists a politician k whose first dimension ideology, θ_1^k, is known.
3. (i) There exists a bill 0 such that $m_0 = 0$. (ii) There exists a bill, s, for which $m_s \neq 0$.
4. The two parties apply pressure in the same direction on at least one bill, and opposite directions on at least one other.

For $\mathcal{I}_t = 1$, we can rewrite (5) as:

$$\Pr\left(Y_{it} = 1 | \bar{q}_t \in Q_p^1, \bar{x}_t; \theta^i, y_p^{max}, m_t\right) = \Phi\left(\sqrt{\frac{1}{1+m_t^2}}\left(\theta_2^i - m_t\theta_1^i - b_t\right) + W_{p,t} \times y_p^{max}\right)$$

Let us use the simplified notation, $\Pr(Y_{it} = 1) = \Pr\left(Y_{it} = 1 | \bar{q}_t \in Q_p^1, \bar{x}_t; \theta^i, y_p^{max}, m_t\right)$. This term is the likelihood component of politician i voting Yes on a bill t if $\mathcal{I}_t = 1$. It is more convenient for us to work with the standardized likelihood:

$$\Phi^{-1}\left(\Pr\left(Y_{it} = 1\right)\right) = \sqrt{\frac{1}{1+m_t^2}}\left(\theta_2^i - m_t\theta_1^i - b_t\right) + W_{p,t} \times y_p^{max}, \tag{A-1}$$

which makes explicit the unique correspondence between data (on the left hand side) and model parameters (on the right hand side).

Using Assumption ID3(i), we begin by comparing the probability of voting Yes on the normalizing bill 0 between any two politicians, i and j, belonging to the same party:

$$\Phi^{-1}\left(\Pr\left(Y_{i0}=1\right)\right)-\Phi^{-1}\left(\Pr\left(Y_{j0}=1\right)\right)=\theta_2^i-\theta_2^j$$

It is immediate that with $j=0$ (the normalized member in Assumption ID2(i)), we obtain identification of θ_2^i for all members of the party containing member 0, which, in correspondence with our empirical application, we assume is party D (without loss).

For $\mathcal{I}_t=0$, we have instead

$$\Phi^{-1}\left(1-\Pr\left(Y_{it}=1\right)\right)=\sqrt{\frac{1}{1+m_t^2}}\left(\theta_2^i-m_t\theta_1^i-b_t\right)+W_{p,t}\times y_p^{max}. \tag{A-2}$$

One can see immediately that the difference in standardized likelihoods, using (A-2), for bill 0 will again identify the second dimension ideologies, $\{\theta_2^i\}_{i=1}^N$ for members of party D.

We next show that the cutlines for each party, and directions, $\mathcal{I}_t$, are unique for each bill. Consider the vote decisions of politician 0 and another member of party D, j, on an arbitrary bill, t. The standardized likelihoods are given by:

$$\begin{aligned}\Phi^{-1}\left(\Pr\left(Y_{0t}=1\right)\right)&=\pm\sqrt{\tfrac{1}{1+m_t^2}}\left(\theta_2^0-m_t\theta_1^0-b_t\right)\pm W_{D,t}\times y_D^{max}\\ \Phi^{-1}\left(\Pr\left(Y_{jt}=1\right)\right)&=\pm\sqrt{\tfrac{1}{1+m_t^2}}\left(\theta_2^j-m_t\theta_1^j-b_t\right)\pm W_{D,t}\times y_D^{max},\end{aligned} \tag{A-3}$$

where the sign of the RHS depends upon $\mathcal{I}_t$.

The set of points in the (θ_1,θ_2) space that are at distance $\Phi^{-1}\left(\Pr\left(Y_{it}=1\right)\right)$ from i's ideal point define a circle centered at $\bar{\theta}^i$. Allowing for both $\mathcal{I}_t=0$ and $\mathcal{I}_t=1$, the equations for members 0 and j in (A-3) define the tangents to each of the two circles for members 0 and j. At most four $(m_t,\hat{b}_t,y_D^{max})$ triplets define cutlines that are tangent to both circles: at most two outer tangents that place members 0 and j on the same side of a cutline, and at most two inner tangents that place the members 0 and j on opposite sides of a cutline. Figure A1 illustrates the possible cutlines.

For an outer tangent for which both members lie on the same side, we have $\theta_2^i<m_t\theta_1^i+\hat{b}_t\mp W_{D,t}\times y_{D,t}$ for $i\in\{0,j\}$, or $\theta_2^i>m_t\theta_1^i+\hat{b}_t\mp W_{D,t}\times y_{D,t}$ for $i\in\{0,j\}$. These inequalities imply $\Pr\left(Y_{it}=1\right)<\frac{1}{2}$ for both members or $\Pr\left(Y_{it}=1\right)>\frac{1}{2}$ for both members, depending on $\mathcal{I}_t$.

For an inner tangent for which one member lies on each side, we instead have either $\Pr\left(Y_{0t}=1\right)<\frac{1}{2}$ and $\Pr\left(Y_{jt}=1\right)>\frac{1}{2}$, or $\Pr\left(Y_{0t}=1\right)>\frac{1}{2}$ and $\Pr\left(Y_{jt}=1\right)<\frac{1}{2}$, again depending on $\mathcal{I}_t$.

Therefore, given knowledge of the voting probabilities, at most two of the four possible cutlines (with an appropriate $\mathcal{I}_t$ associated with that cutline) can simultaneously satisfy the equations for the standardized likelihood of 0 and j: either two cutlines that form outer tangents, or two cutlines that form inner tangents.[1]

Assumption ID1 allows us to show that the cutline and direction of each bill is uniquely determined from the two remaining possibilities by means of contradiction. Suppose, to the contrary, that two cutline/direction pairs satisfy the two standardized likelihood equations for 0 and j. Recall that each associated cutline must be tangent to both of the circles centered on each member's ideal point.

Now consider the possible locations of the other members, i, of party D. To ensure the two cutlines are indistinguishable, the circle centered on $\bar{\theta}^i$ with radius $\Phi^{-1}\left(\Pr\left(Y_{it}=1\right)\right)$ for each member must also be tangent to both potential cutlines. Following the Locus theorem, a generic D member i must then lie on the line, A, passing through $\bar{\theta}^0$ and the intersection of the two potential cutlines, or on the line orthogonal to A, that also passes through the intersection, A'. Points on these two lines are the only points that ensure i is equidistant from both cutlines, so that the circle associated with i is tangent to both.

We can rule out points on the line A'. If the two potential cutlines are outer tangents to the circles of 0 and j, then if a member i is located on A', he lies on the same side as 0 and j for one cutline and on the opposite side for the other. But, we know how each of the three probabilities, $\Pr\left(Y_{0t}=1\right)$, $\Pr\left(Y_{jt}=1\right)$, and $\Pr\left(Y_{it}=1\right)$, compares to one-half. If all are on the same side, all must be greater than one-half or all must be less. If i is on the opposite side, then his probability must be greater than one-half if the other two are less than one-half, or vice versa. Thus, if i lies on A', we can distinguish between the two pairs of solutions, a contradiction. Similarly, if the two potential cutlines are inner tangents to the circles of 0 and j then for one of the cutlines, i is on the same side as 0 (and opposite to j) and for the other i is on the same side as j (and opposite to 0). Knowing which voting probabilities are greater or less than one-half again allows us to tell the solutions apart.

We have then shown that if we have two potential solutions, all members of party D must lie on the line A. But, the same argument applies to party R: taking any two members for party R, we can show that for there to be two potential cutlines for party R (with associated directions), all members of party R must also be collinear. But, if the members of each party are collinear, we violate Assumption ID1. Thus, the cutline for each party, as well as the direction, $\mathcal{I}_t$, is unique for all bills.

Uniqueness of the cutlines immediately guarantees m_t is unique for each bill (but not necessarily b_t or y_D^{max}, because, for each bill, only their sum or difference enters the vote probabilities). Furthermore, given uniqueness of the cutlines and direction of each bill, if the ideological position

---

[1] In the two limiting cases in which a cutline passes exactly through a member's ideal point, the two possible cutlines are such that they pass on opposite sides of the other member's ideal point. The appropriate cutline is then immediately identified by knowing whether this second member's voting probability is greater or less than one-half.

of a member of either party is known in one dimension, the ideological position in the other dimension is generically unique, because only one possible ideological position for the member at the distance, $\Phi^{-1}\left(\Pr\left(Y_{it}=1\right)\right)$, from the cutline exists (the vote probabilities in (A-3) are linear in each dimension). The two exceptions are: (i) the first dimensional ideology is known and the cutline is vertical, or (ii) the second dimensional ideology is known and the cutline is horizontal. But, given that θ_1^k is known for member k (Assumption ID2(ii)), θ_2^k is unique because we have at least one bill that doesn't have a vertical cutline (the normalizing bill). And, given that θ_2^j is known for all members of party D, each θ_1^j is unique because we have at least one cutline that is not horizontal (Assumption ID3(ii)).

We next establish uniqueness of each of b_t, y_D^{max}, and y_R^{max} using only uniqueness of the cutlines, directions, and positions of members 0 and k. In our empirical application, the normalizing member, k, of Assumption ID2(ii) belongs to party R.[2] The difference in the normalized likelihoods of members 0 and k is given by

$$\begin{aligned}
&\Phi^{-1}\left(\Pr\left(Y_{0t}=1\right)\right)-\Phi^{-1}\left(\Pr\left(Y_{kt}=1\right)\right) \\
= \;&\pm\sqrt{\frac{1}{1+m_t^2}}\left(\theta_2^0-\theta_2^k-m_t(\theta_1^0-\theta_1^k)\right)\pm W_{D,t}\times y_D^{max}\mp W_{R,t}\times y_R^{max} \qquad \text{(A-4)}
\end{aligned}$$

The party pressure directions are known from the data on leadership votes up to the indeterminacy of $\mathcal{I}_t$. From Assumption ID4, we can write the equations corresponding to (A-4) for two bills, t and r, one in which the two parties exert pressure in the same direction (t) and one in which they exert pressure in opposite directions (r)[3]:

$$\begin{aligned}
\Phi^{-1}\left(\Pr\left(Y_{0t}=1\right)\right)-\Phi^{-1}\left(\Pr\left(Y_{kt}=1\right)\right) &= \pm\sqrt{\frac{1}{1+m_t^2}}\left(-\theta_2^k+m_t\theta_1^k\right)\pm y_D^{max}\mp y_R^{max} \\
\Phi^{-1}\left(\Pr\left(Y_{0r}=1\right)\right)-\Phi^{-1}\left(\Pr\left(Y_{kr}=1\right)\right) &= \pm\sqrt{\frac{1}{1+m_r^2}}\left(-\theta_2^k+m_r\theta_1^k\right)\pm y_D^{max}\pm y_R^{max} \qquad \text{(A-5)}
\end{aligned}$$

Regardless of the directions, $\mathcal{I}_t$, for each bill, the two equations of (A-5) are linearly independent, because the first equation has the difference of the party pressure parameters on the right-hand side and the second equation has the sum. Thus, given uniqueness of the other parameters in the equations, the pressure parameters are also unique.[4]

---

[2] We do not require the two normalizing members of Assumption ID2 to belong to different parties. In fact, the proof is somewhat simpler if they are in the same party.

[3] We take $W_{D,t}=W_{R,t}=1$, $W_{D,r}=1$, and $W_{R,r}=-1$, but the same argument holds for the other possibilities.

[4] In the version of the model in which parties only exert pressure (in opposite directions) when the party leaderships disagree, we cannot separately identify the party pressure parameters. In this case, we have only the

Given uniqueness of all of the cutlines, directions, and y_D^{max}, the unique value of each b_t is determined by the equation corresponding to (A-1) for member 0. Then, to establish uniqueness of members $i \neq k$ of party R, we can take the difference in normalized likelihoods between member i and member 0 on the normalizing bill:

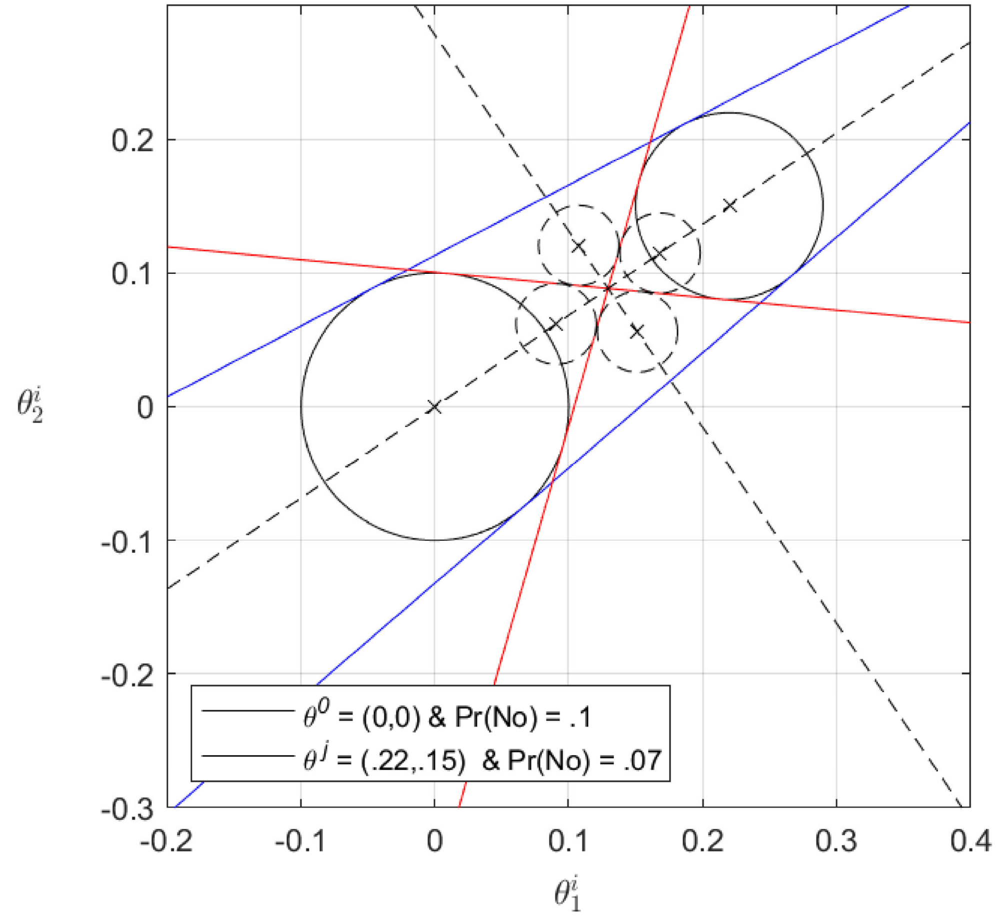
$$\begin{aligned} & \Phi^{-1}\left(\Pr\left(Y_{00}=1\right)\right)-\Phi^{-1}\left(\Pr\left(Y_{it}=1\right)\right) \\ = \; & \mp\theta_2^i \pm W_{D,t} \times y_D^{max} \mp W_{R,t} \times y_R^{max} \end{aligned}$$

which establishes that each θ_2^i for a member of party R is unique. Finally, given each θ_2^i is unique, each θ_1^i of a member of party R must be unique because we have at least one cutline that is not horizontal (Assumption ID3(ii)).□

---

second of the two equations in (A-5) so that only the sum of the pressure parameters, $y_D^{max} + y_R^{max}$, is identified. We can make use of the bills without party pressure to establish uniqueness of the other parameters using similar arguments to those for the main case.

Figure A1: Identification Assumptions in a Numerical Example



Appendix B: DW-Nominate's Lack of Identification in Two Dimensions (or higher)

In this section, we provide new insights as to the lack of identification of DW-Nominate (Dynamically Weighted Nominal Three-Step Estimation) in two dimensions. In Section B.1, we formally prove (building on, but correcting the proof in Potthoff (2018)), that W-Nominate is not identified. This result immediately extends to DW-Nominate, as it is a generalization of W-Nominate with dynamically changing ideal points (i.e. preferences linearly changing in time).[5] In Section B.2, we show that, even if the utility weight in W-Nominate were constrained to 1, the Gaussian utility function assumed in Nominate makes it very difficult to determine the number of normalizations necessary for it to be identified. This section builds on the work of Rivers (2003), which is, to date, the best formal discussion of identification of multidimensional spatial models. Finally, in Section B.3, we consider the effect of normalizing members' ideologies to lie within a unit circle: the only clearly specified normalization that Nominate imposes.

As background, the current version of DW-Nominate, updates active members' ideologies and estimates the cutline parameters for new bills as they become available (Boche et al., 2018). To do so, it holds constant inactive members' ideologies and the cutlines of previous bills (no "back-propagation"). New ideology and cutline estimates all rely on previous runs of DW-Nominate for identification. To quote Boche et al. (2018), p.24, "*...By effectively locking in place the locations that Poole last estimated for past members, we guarantee that our scores maintain compatibility with the widely used DW-Nominate scores with which scholars are familiar.*" Thus, unfortunately, beyond the unit circle normalization that DW-Nominate imposes, we do not know what other normalizations were initially imposed. As we show, however, no matter what these normalizations were, DW-Nominate is not identified.

B.1: Lack of Identification of W-Nominate

In W-Nominate, the 'W' stands for 'weighted'. It normalizes the utility weight in the first dimension to be one and allows the weight in second dimension, w_2, to be estimated. Here, we prove that this model is not identified by providing a transformation that can change the rank ordering of members in either (or both) dimensions. Importantly, the transformation we provide is not a combination of a rotation, scale, and translation and thus poses a problem even if the rotation,

---

[5] In fact, the parameters that govern the changes in ideology over time are also easily shown to not be identifiable. As the cutline parameters of each Congress are arbitrary, one can simultaneously change both the cutline parameters and the parameter that governs the change in ideology without changing the vote probabilities. To identify changes in ideology, one would either need to assume some reference ideology remains unchanged across Congresses or assume that some bill is identical (has the same cutline parameters) in each Congress.

scale, and location of the estimates are constrained via suitable normalization (as in our work).

Consider the likelihood argument in Carroll et al. (2009):

$$\begin{aligned} Pr(Y_{i,t}=1) &= \Phi\left[u(\overline{\theta}_i,\mathbf{x}_t)-u(\overline{\theta}_i,\mathbf{q}_t)\right]= \\ & \Phi\left[\beta e^{-\frac{1}{2}\left(\theta_1^i-x_{1,t}\right)^2-\frac{w_2}{2}\left(\theta_2^i-x_{2,t}\right)^2}-\beta e^{-\frac{1}{2}\left(\theta_1^i-q_{1,t}\right)^2-\frac{w_2}{2}\left(\theta_2^i-q_{2,t}\right)^2}\right] \end{aligned}$$

where $\Phi(\cdot)$ is the CDF of the standard normal distribution. The vector of parameters of interest is $\Theta=\{\theta_1^i,x_{1,t},q_{1,t},\theta_2^i,x_{2,t},q_{2,t},w_2\}$.

Consider $s>0$ and $0<r<1$ and define the following candidate (nonlinear) transformation of the parameter vector, which can be proven to not be a rotation (other than in the special case $w_2=s=1$):

$$\begin{aligned} \tilde{\theta}_1^i &= \theta_1^i\sqrt{r}-\theta_2^i\sqrt{w_2\left(1-r\right)} \\ \tilde{x}_{1,t} &= x_{1,t}\sqrt{r}-x_{2,t}\sqrt{w_2\left(1-r\right)} \\ \tilde{q}_{1,t} &= q_{1,t}\sqrt{r}-q_{2,t}\sqrt{w_2\left(1-r\right)} \\ \tilde{\theta}_2^i &= s\times\left(\theta_1^i\sqrt{(1-r)}+\theta_2^i\sqrt{w_2 r}\right) \\ \tilde{x}_{2,t} &= s\times\left(x_{1,t}\sqrt{(1-r)}+x_{2,t}\sqrt{w_2 r}\right) \\ \tilde{q}_{2,t} &= s\times\left(q_{1,t}\sqrt{(1-r)}+q_{2,t}\sqrt{w_2 r}\right) \\ \tilde{w}_2 &= \frac{1}{s^2} \end{aligned}$$

To check that within this class of transformations one obtains the same likelihood of the vote data:

$$\begin{aligned} &\Phi\left[\beta e^{-\frac{1}{2}\left(\tilde{\theta}_1^i-\tilde{x}_{1,t}\right)^2-\frac{\tilde{w}_2}{2}\left(\tilde{\theta}_2^i-\tilde{x}_{2,t}\right)^2}-\beta e^{-\frac{1}{2}\left(\tilde{\theta}_1^i-\tilde{q}_{1,t}\right)^2-\frac{\tilde{w}_2}{2}\left(\tilde{\theta}_2^i-\tilde{q}_{2,t}\right)^2}\right]= \\ &\quad \Phi\left[\beta e^{-\frac{1}{2}\left(\theta_1^i-x_{1,t}\right)^2-\frac{w_2}{2}\left(\theta_2^i-x_{2,t}\right)^2}-\beta e^{-\frac{1}{2}\left(\theta_1^i-q_{1,t}\right)^2-\frac{w_2}{2}\left(\theta_2^i-q_{2,t}\right)^2}\right] \end{aligned}$$

it suffices to show that:

$$\left(\tilde{\theta}_1^i - \tilde{x}_{1,t}\right)^2 + \tilde{w}_2\left(\tilde{\theta}_2^i - \tilde{x}_{2,t}\right)^2 =$$

$$\begin{aligned}&\left(\theta_1^i\sqrt{r} - \theta_2^i\sqrt{w_2\left(1-r\right)} - x_{1,t}\sqrt{r} + x_{2,t}\sqrt{w_2\left(1-r\right)}\right)^2 \\ &+\frac{1}{s^2}\left(s\times\left(\theta_1^i\sqrt{(1-r)}+\theta_2^i\sqrt{w_2 r}\right) - s\times\left(x_{1,t}\sqrt{(1-r)}+x_{2,t}\sqrt{w_2 r}\right)\right)^2 =\end{aligned}$$

$$\left(\left(\theta_1^i - x_{1,t}\right)\sqrt{r} - \left(\theta_2^i - x_{2,t}\right)\sqrt{w_2\left(1-r\right)}\right)^2 + \left(\left(\theta_1^i - x_{1,t}\right)\sqrt{(1-r)} + \left(\theta_2^i - x_{2,t}\right)\sqrt{w_2 r}\right)^2 =$$

$$\begin{aligned}&\left(\theta_1^i - x_{1,t}\right)^2 r + \left(\theta_2^i - x_{2,t}\right)^2 w_2\left(1-r\right) - 2\left(\theta_1^i - x_{1,t}\right)\sqrt{r}\left(\theta_2^i - x_{2,t}\right)\sqrt{w_2\left(1-r\right)} \\ &+\left(\theta_1^i - x_{1,t}\right)^2\left(1-r\right) + \left(\theta_2^i - x_{2,t}\right)^2 w_2 r + 2\left(\theta_1^i - x_{1,t}\right)\sqrt{(1-r)}\left(\theta_2^i - x_{2,t}\right)\sqrt{w_2 r} =\end{aligned}$$

$$\left(\theta_1^i - x_{1,t}\right)^2 + w_2\left(\theta_2^i - x_{2,t}\right)^2$$

This proves that W-Nominate in two dimensions is not identified up to this class of transformations, which is broader than than the class of transformation that only rotate, scale, and/or change the location of the ideal points.

To show how this class of transformations is particularly damaging, consider the three individuals, $i = a, b, c$, located at points $\bar{\theta}^a = (-.3, -1), \bar{\theta}^b = (.1, -.3)$, and $\bar{\theta}^c = (.25, -1.2)$ in Figure B1.

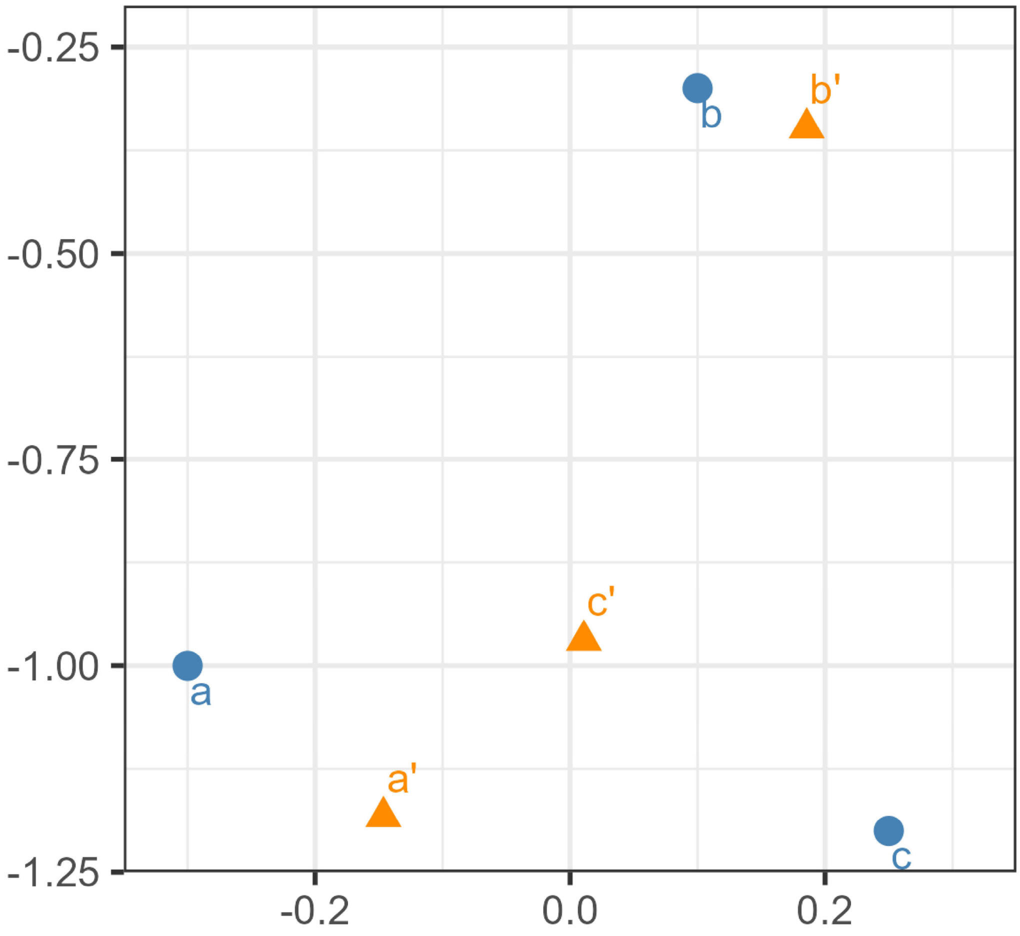
Consider the proposed transformation:

$$\begin{aligned}\tilde{\theta}_1^i &= \theta_1^i\sqrt{r} - \theta_2^i\sqrt{w_2\left(1-r\right)} \\ \tilde{\theta}_2^i &= s\times\left(\theta_1^i\sqrt{(1-r)}+\theta_2^i\sqrt{w_2 r}\right)\end{aligned}$$

for the case of $r = .8; s = 12.5; w_2 = .3$. After applying this transformation (which, with the corresponding transformations for x_t, q_t, and w_2 will not change the vote probabilities) results in individuals a, b, c being located at $(-.0234, -7.8008), (.1629, -1.2781)$, and $(.5175, -5.9509)$, respectively. That is, applying this transformation to each i, rearranges the data cardinally and, more significantly, ordinally. However, so do other types of transformations, including linear transformations such as rotations, and therefore this may appear of no particular concern.

What is damaging is that once the transformation is applied to the original data, it is no longer possible to recover the original ordinal ranking of the true positions. We illustrate this fact by optimally rotating the transformed data back into the original data space.[6] In this example, even after optimally rotating the transformed positions back to the original data space, we observe ordinal changes with respect to the true positions along both dimensions. The new locations are at the points, $\bar{\theta}_{a'} = (-.1464, -1.1828), \bar{\theta}_{b'} = (.1856, -.3481)$, and $\bar{\theta}_{c'} = (.0108, -.9691)$ as illustrated in Figure B1. As can be seen, a', b', c' are now misordered along both dimensions relative to the original ideal points.

Figure B1: Problematic Example for DW-Nominate



---

[6] In particular, we apply the Procrustes rotation to the transformed data employing optimal shift, scale, and rotation, so as to bring the transformed data back to the original data space.

B.2: Identification of Nominate

The previous section proves lack of identification for nonlinear transformations when, as in W-Nominate and DW-Nominate, the utility weight in the second dimension is estimated. Here, we discuss the identification of Nominate, which constrains all utility weights to be equal to one.[7]

In Section B.2.1, we consider the problem of identifying members' ideologies under the assumption that some of the cutline parameters, $\bar{x}_t$ and $\bar{q}_t$, are known. In Section B.2.2, we discuss the reverse problem: identifying the cutline parameters assuming some of the ideology parameters are known. Sections B.2.1 and B.2.2 are illustrative of the interim steps of the Nominate method (Nominal Three-Step Estimation), where either the cutlines or the ideal points are taken as given and the remaining set of parameters are estimated, iterating until convergence.

B.2.1: Known Bill Parameters

Making use of the Gaussian preferences employed in Nominate, let us start by highlighting that, for known roll call "0"

$$\begin{gathered}\Phi^{-1}\left[Pr(Y_{i,0}=1)\right] = u(\bar{\theta}^i,\bar{x}_0) - u(\bar{\theta}^i,\bar{q}_0)\\ = e^{-\frac{1}{2}\left[\left(\theta_1^i-x_{1,0}\right)^2+\left(\theta_2^i-x_{2,0}\right)^2\right]} - e^{-\frac{1}{2}\left[\left(\theta_1^i-q_{1,0}\right)^2+\left(\theta_2^i-q_{2,0}\right)^2\right]}\end{gathered}$$

is a highly-nonlinear equation in two unknowns (θ_1^i,θ_2^i). A generalized cubic equation in (θ_1^i,θ_2^i) follows from a second-order Taylor expansion of the difference in the deterministic utilities on the RHS for each vote:

$$\begin{gathered}\Phi^{-1}\left[Pr(Y_{i,0}=1)\right] =\\ e^{-\frac{1}{2}\left[\left(\theta_1^i-x_{1,0}\right)^2+\left(\theta_2^i-x_{2,0}\right)^2\right]} - e^{-\frac{1}{2}\left[\left(\theta_1^i-q_{1,0}\right)^2+\left(\theta_2^i-q_{2,0}\right)^2\right]} =\\ \textstyle\sum_{n=0}^{\infty}\frac{\left(-\frac{1}{2}\right)^n}{n!}\left[\left[\left(\theta_1^i-x_{1,0}\right)^2+\left(\theta_2^i-x_{2,0}\right)^2\right]^n - \left[\left(\theta_1^i-q_{1,0}\right)^2-\left(\theta_2^i-q_{2,0}\right)^2\right]^n\right] \approx\\ -\tfrac{1}{2}\left[\textstyle\sum_{j=1}^{2}\left(\theta_j^i-x_{j,0}\right)^2-\sum_{j=1}^{2}\left(\theta_j^i-q_{j,0}\right)^2\right]+\tfrac{1}{8}\left[\left[\textstyle\sum_{j=1}^{2}\left(\theta_j^i-x_{j,0}\right)^2\right]^2-\left[\textstyle\sum_{j=1}^{2}\left(\theta_j^i-q_{j,0}\right)^2\right]^2\right] =\\ -\tfrac{1}{2}\left[\textstyle\sum_{j=1}^{2}\left(x_{j,0}-q_{j,0}\right)\left(x_{j,0}+q_{j,0}-2\theta_j^i\right)\right]\times\left[1-\tfrac{1}{4}\textstyle\sum_{j=1}^{2}\left[\left(x_{j,0}\right)^2+\left(q_{j,0}\right)^2-2\theta_j^i\left(x_{j,0}+q_{j,0}-\theta_j^i\right)\right]\right]\end{gathered}$$

It is therefore possible to see that, even using approximations, a single normalization on a "0" bill is insufficient to uniquely pin down the (θ_1^i,θ_2^i) unknowns from the data $\Phi^{-1}\left[Pr(Y_{i,0}=1)\right]$.

Notice further that even for a quadratic loss function, instead of a Gaussian utility function, a

---

[7] We discuss the difficulties a Gaussian utility function creates even when $\beta=1$ is assumed (Nominate estimates the parameter β as well, creating a further burden for identification on top of the ones discussed here).

single roll call normalization would still be insufficient for an unique mapping:

$$\begin{gathered}
\Phi^{-1}\left[Pr(Y_{i,0}=1)\right] = \\
-\tfrac{1}{2}\left(\left(\theta_1^i - x_{1,0}\right)^2 + \left(\theta_2^i - x_{2,0}\right)^2\right) + \tfrac{1}{2}\left(\left(\theta_1^i - q_{1,0}\right)^2 + \left(\theta_2^i - q_{2,0}\right)^2\right) = \\
-\tfrac{1}{2}\left[\textstyle\sum_{j=1}^{2}\left(\theta_j^i - x_{j,0}\right)^2 - \sum_{j=1}^{2}\left(\theta_j^i - q_{j,0}\right)^2\right] = \\
-\tfrac{1}{2}\textstyle\sum_{j=1}^{2}\left(x_{j,0} - q_{j,0}\right)\left(x_{j,0} + q_{j,0} - 2\theta_j^i\right)
\end{gathered}$$

To see the extent of the normalizations needed for different classes of individual utility functions, consider full knowledge of all policy issues $\bar{x}_t, \bar{q}_t$ for the set of T bill upon which a politician i votes, which can be treated as data. Then we can write the system of polynomials in the unknown ideology parameters, (θ_i^1, θ_i^2):

$$\begin{cases}
\Phi^{-1}\left[Pr(Y_{i,0}=1)\right] = \delta_0^0 + \delta_1^0\theta_1^i + \delta_2^0\theta_2^i + \delta_3^0\left(\theta_1^i\right)^2 + \delta_4^0\left(\theta_2^i\right)^2 + \delta_5^0\theta_1^i\theta_2^i + ... \\
... \\
\Phi^{-1}\left[Pr(Y_{i,t}=1)\right] = \delta_0^t + \delta_1^t\theta_1^i + \delta_2^t\theta_2^i + \delta_3^t\left(\theta_1^i\right)^2 + \delta_4^t\left(\theta_2^i\right)^2 + \delta_5^t\theta_1^i\theta_2^i + ... \\
... \\
\Phi^{-1}\left[Pr(Y_{i,T}=1)\right] = \delta_0^T + \delta_1^T\theta_1^i + \delta_2^T\theta_2^i + \delta_3^T\left(\theta_1^i\right)^2 + \delta_4^T\left(\theta_2^i\right)^2 + \delta_5^T\theta_1^i\theta_2^i + ...
\end{cases} \tag{B-1}$$

Here, full knowledge of all $\bar{x}_t = (x_t^1, x_t^2), \bar{q}_t = (q_t^1, q_t^2)$ delivers what essentially amounts to bill-specific data $\{\delta_0^t, \delta_1^t, \delta_2^t, \delta_3^t, \delta_4^t, \delta_5^t, ...\}$, and (B-1) remains a system of T (typically nonlinear) equations in the two original unknowns (θ_1^i, θ_2^i). Generally, there cannot be any theoretical assurance of a unique exact mapping from the data on the LHS of the equations in the system to a unique $(\theta_1^i, \theta_2^i)^*$ for every i beyond the linear system case. However, operating under the hypothesis that the model is correctly specified the system in (B-1) will admit a unique solution for T large enough. In fact, (θ_1^i, θ_2^i) may be identifiable given knowledge of only the bill parameters for $\tau < T$ bills. We illustrate a few cases here, but emphasize that a general proof is not available (to the best of our knowledge).

For the quadratic utility case, the number of necessary normalizations is $\tau = 2$ bills (i.e. 8 parameter restrictions for $\bar{x}_0, \bar{x}_1, \bar{q}_0, \bar{q}_1$), given that the polynomials in (B-1) are of the first order. This implies that two roll calls can uniquely identify a solution (θ_1^i, θ_2^i) to (B-1), i.e. there is no observationally equivalent $\left(\tilde{\theta}_1^i, \tilde{\theta}_2^i\right) \neq (\theta_1^i, \theta_2^i)$ delivering the same set of values $\Phi^{-1}\left[Pr(Y_{i,t}=1)\right]$.

This result for quadratic utility is conceptually identical to the result in Rivers (2003), which proves that, for $d = 2$, the number of required restrictions is $d(d+1) = 6$. The difference here is that here we are considering as parameters the policy points, and not simply the policy cutlines (the

6 parameter restrictions on $\{\delta_0^0, \delta_1^0, \delta_2^0, \delta_0^1, \delta_1^1, \delta_2^1\}$). This difference does not affect the identification of the set of ideal points, but makes identification of the bill parameters more burdensome.

For utility functions that deliver conic functions in the system (B-1), the number of required normalizations $\tau = 5$ (i.e. 20 parameter restrictions). To see why, consider first that any system of two conic equations admits at most four solutions. Define these solutions as $\left\{\theta^A, \theta^B, \theta^C, \theta^D\right\}$. All of these solution are observationally equivalent in the sense of exactly satisfying both equations. This system defines the first two roll calls $\{\bar{x}_t, \bar{q}_t\}_{t=0,1}$ that are required for normalization. Let us now add an additional third bill $\bar{x}_2, \bar{q}_2$ introducing another conic equation and under the assumption that such conic equation is non-redundant in the sense of the direction of axes of the associated ellipse are not the same as those of any of the previously normalized conic equations. At most, three of the elements of the set $\left\{\theta^A, \theta^B, \theta^C, \theta^D\right\}$ will satisfy this third equation (if all the elements of $\left\{\theta^A, \theta^B, \theta^C, \theta^D\right\}$ satisfied this third restriction, than that would imply that the third conic equation is, in fact, redundant). Without loss, define the remaining set of candidate solutions as $\left\{\theta^A, \theta^B, \theta^C\right\}$. Adding a fourth bill to the normalization (again assuming non-redundancy), delivers a set of candidate solutions satisfying this fourth constraint of (at most) two elements $\left\{\theta^A, \theta^B\right\}$, and a fifth bill, pins down the ideology vector uniquely to, say, $\left\{\theta^A\right\}$. In summary, normalization of five bills is needed for theoretical identification of the ideology parameters (θ_1^i, θ_2^i) under the assumption that the model is correctly specified.

For utility functions that deliver cubic functions in (B-1), as in the case of a second-order approximation of the difference in Gaussian utilities used in Nominate, the number of normalizations is higher than $\tau = 5$ bills, as the number of conditions grows. This exercise illustrates that the number of normalizations required for Gaussian utility functions in Nominate is likely much higher than that required for quadratic utility functions, and that it is difficult to determine how many bills must be normalized to uniquely identify the ideal points for N members.

The discussion in this subsection illustrates the inherent difficulty in proving identification within each of Nominate's interim steps (i.e. the algorithm's iteration step where all of the cutline parameters are assumed given and the ideal points are estimated). It is not immediate that each iteration is guaranteed to deliver a unique vector of ideal point estimates.

B.2.2: Known Ideal Points

Concerning the policy choice parameters $\bar{x}_t, \bar{q}_t$, let us focus on the expression

$$Pr(Y_{i,t} = 1) =$$
$$\Phi\left[e^{-\frac{1}{2}\left(\theta_1^i - x_{1,t}\right)^2 - \frac{1}{2}\left(\theta_2^i - x_{2,t}\right)^2} - e^{-\frac{1}{2}\left(\theta_1^i - q_{1,t}\right)^2 - \frac{1}{2}\left(\theta_2^i - q_{2,t}\right)^2}\right]$$

for known ideology parameters. Specifically, under a normalization for $\theta^0 = (\theta_1^0, \theta_2^0)$, we can write:

$$\begin{gathered}
\Phi^{-1}\left[Pr(Y_{0,t}=1)\right] = \\
e^{-\frac{1}{2}\left[\left(\theta_1^0 - x_{1,t}\right)^2 + \left(\theta_2^0 - x_{2,t}\right)^2\right]} - e^{-\frac{1}{2}\left[\left(\theta_1^0 - q_{1,t}\right)^2 + \left(\theta_2^0 - q_{2,t}\right)^2\right]} = \\
\textstyle\sum_{n=0}^{\infty} \frac{\left(-\frac{1}{2}\right)^n}{n!} \left[\left[\left(\theta_1^0 - x_{1,t}\right)^2 + \left(\theta_2^0 - x_{2,t}\right)^2\right]^n - \left[\left(\theta_1^0 - q_{1,t}\right)^2 + \left(\theta_2^0 - q_{2,t}\right)^2\right]^n\right] \qquad \approx \\
-\tfrac{1}{2}\left[\textstyle\sum_{j=1}^{2} \left(x_{j,t} - q_{j,t}\right)\left(x_{j,t} + q_{j,t} - 2\theta_j^0\right)\right] \times \left[1 - \tfrac{1}{4}\textstyle\sum_{j=1}^{2} \left[\left(x_{j,t}\right)^2 + \left(q_{j,t}\right)^2 - 2\theta_j^0\left(x_{j,t} + q_{j,t} - \theta_j^0\right)\right]\right]
\end{gathered}$$

which, even in second-order approximate form, does not lend to an immediate analysis of the mapping from data to policy points and generally admits multiple solutions.

With a further normalization for $\theta^1 = (\theta_1^1, \theta_2^1)$ one can make more progress focusing on quadratic losses or first-order approximation of the (difference in) Gaussian utilities. In particular, note that with quadratic losses:

$$\begin{aligned}
&\Phi^{-1}\left[Pr(Y_{0,t}=1)\right] - \Phi^{-1}\left[Pr(Y_{1,t}=1)\right] = \\
-\frac{1}{2}\sum_{j=1}^{2}\left(x_{j,t} - q_{j,t}\right)\left(x_{j,t} + q_{j,t} - 2\theta_j^0\right) + &\frac{1}{2}\sum_{j=1}^{2}\left(x_{j,t} - q_{j,t}\right)\left(x_{j,t} + q_{j,t} - 2\theta_j^1\right) = \\
&\sum_{j=1}^{2}\left(x_{j,t} - q_{j,t}\right)\left(\theta_j^0 - \theta_j^1\right).
\end{aligned} \tag{B-2}$$

Following a similar approach to that laid out in the preceding section, we can observe that for every roll call t, four equations of the type (B-2) are necessary for the four unknown bill parameters. We require therefore four politicians to be normalized (i.e. 8 parameters) to uniquely identify all parameters $\bar{x}_t, \bar{q}_t$ from the data.

For the case of Gaussian preferences such as those used in Nominate, however, the situation appears more complex. For the case of the second order Taylor expansion, we see that the system of equations of conditions for identification will be composed of generalized quartic equations and so that we know that we need at least 20 restrictions. Again, this fact illustrates that Nominate with Gaussian preferences requires a substantially higher number of identification restrictions than for the quadratic utility case of Rivers (2003). Mirroring the problem with estimating the ideal points holding the cutlines fixed, it is not immediate that the alternative iteration steps in which the ideal points are held fixed and the cutlines estimated will deliver unique cutline estimates.

B.3: A discussion of further normalizations in DW-Nominate

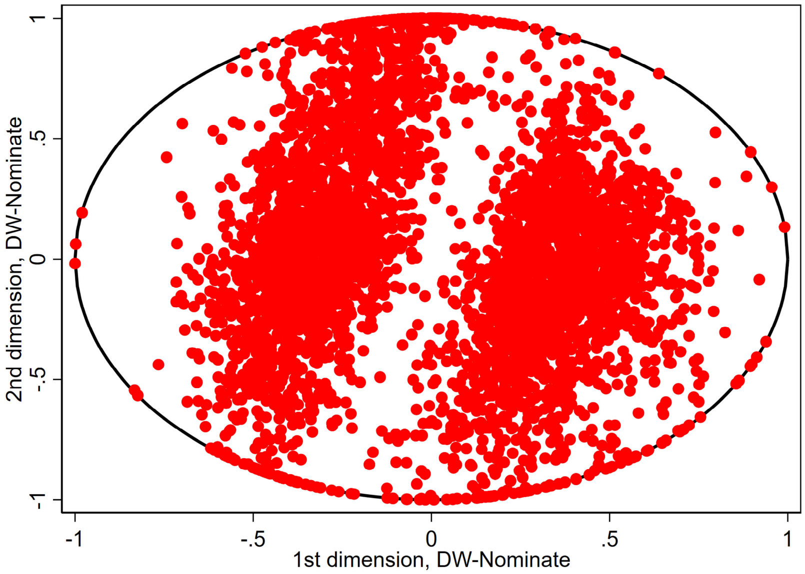
The only normalization that DW-Nominate imposes that is consistently specified (see p.268 of Armstrong et al. 2014) is that all of the ideologies must lie within a unit circle. This normalization may at first appear intuitive, but we point out two difficulties that it creates. Both of the difficulties arise because DW-Nominate does not re-estimate all ideologies and cutline parameters when new roll call data arrives (i.e. no back-propagation). If one were to estimate everything *without* restricting ideologies to the unit circle and then simply rescale them to lie within the unit circle, the normalization would pose no problem. For example, one could take our estimates and simply rescale them all to lie within the unit circle given that the scaling is arbitrary. But, because DW-Nominate imposes the restriction in the estimation process, two complications arise.

The first difficulty is that a unit circle restriction creates an artificial negative correlation between the two dimensions of members' ideological positions. To see this problem most clearly, consider a new member of Congress, i, that is very liberal in the first dimension. Locating this member at $\theta_1^i = -1$ forces him or her to be perfectly moderate in the second dimension (θ_2^i must be 0). In reality, the estimation procedure will be forced to make a compromise: to place a member at an extreme position along the first dimension, it must mechanically moderate the member in the second dimension (and similarly, for placing a member at an extreme position along the second dimension). We do not believe there is any ex ante reason to think that politicians cannot simultaneously hold extreme positions in both dimensions, but DW-Nominate rules out this possibility through the unit circle normalization.

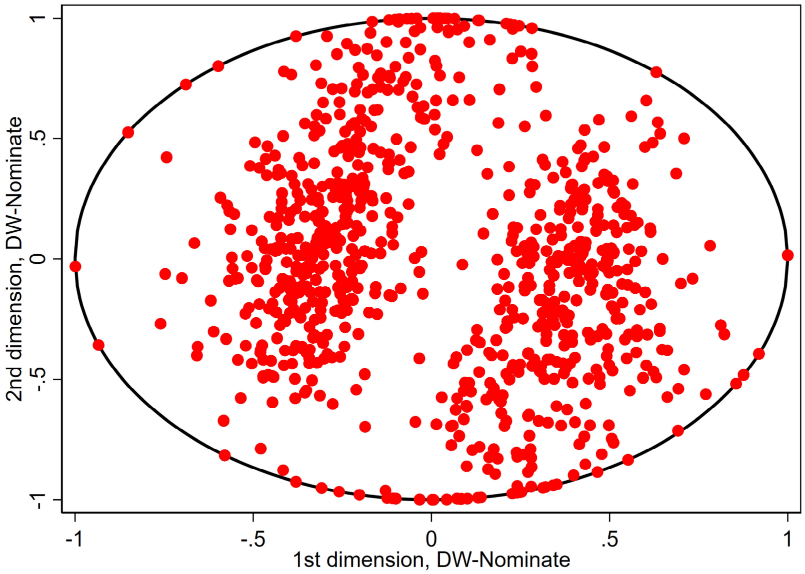
The second difficulty directly stems from the lack of back-propagation. At one point in time, prior to knowing all future members' ideological points, DW-Nominate was scaled such that all members at that time lied within the unit circle. But, unless the constraint was originally 'slack' (no members were located on the unit circle), this scaling implies that any future member that is more extreme than any of those in this initial set will lie on the unit circle boundary artificially. If progressively more extreme politicians are in fact replacing more moderate ones, this normalization starts to progressively become more problematic. To provide suggestive evidence that this artificial constraint is binding, in Figure B2, we plot the unit circle together with all DW-Nominate estimates for each ideology from Congress 70 to Congress 115, both for the House and for the Senate. Since Congress 70, approximately 7% of estimates in the House sit on the boundary of the unit circle, with 8% being on the boundary for the Senate. This evidence suggests that the unit circle boundary is directly and artificially constraining the estimated ideologies for a non-trivial number of legislators. Furthermore, note that this constraint also affects estimates of members away from the boundary, because their ideologies are estimated by incorporating information from those who sit on the boundary.

Figure B2: The Role of the Unit Circle Restriction in DW-Nominate

(a) House of Representatives



(b) Senate



Appendix C: Computational Details of the Estimation Procedure

We maximize the likelihood in (7) via an unconstrained optimization procedure, providing the analytic gradient to the algorithm to greatly improve estimation speed. Rather than using an off-the-shelf quasi-newton algorithm (such as Matlab's fminunc), which proved to perform very poorly given the non-convexity of our likelihood function, we instead use Adam, a version of the steepest descent algorithm. Adaptive Moment Estimation (Adam) is a stochastic optimization algorithm which is also ideal for problems with a large number of parameters like ours (Kingma and Ba, 2014).

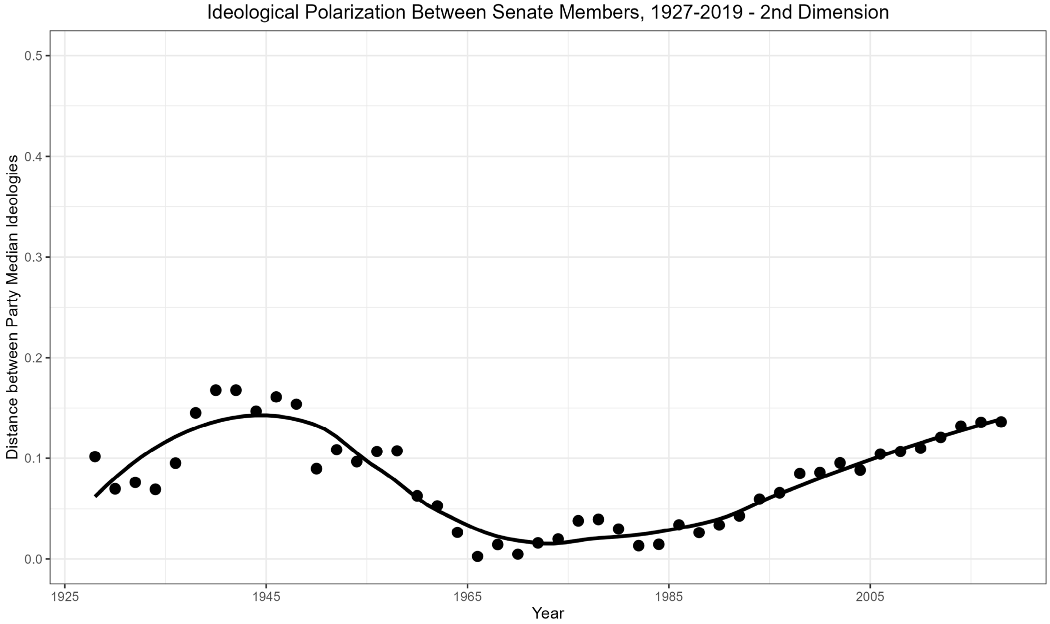
As is standard, we run the estimation procedure until either the stepsize or the gradient is small (for the 2D model, typically the estimation procedure terminated due to the stepsize being small, on the order of 1e-4).

Because for non-convex optimization problems, convergence to a global maximum cannot be guaranteed, we ran the estimation procedure for our main model (Senate 2D) with 60 starting points, with each batch of 12 taking roughly one day when each starting point runs in parallel. For the Senate 2D model, we use the first dimension ideological positions from the Senate 1D model as starting points. For the misspecified Senate 2D model (without party pressure), we use ideology estimates from the full Senate 2D model. Starting points were otherwise randomly chosen (i.e. for the cutlines, party pressure parameters, and ideologies for the 1D models).

We report the estimates for the estimation run that produced the largest likelihood across runs. But, we emphasize that the estimates of the main parameters of interest (namely, the party pressure parameters) were quantitatively very similar (although not identical) across runs.

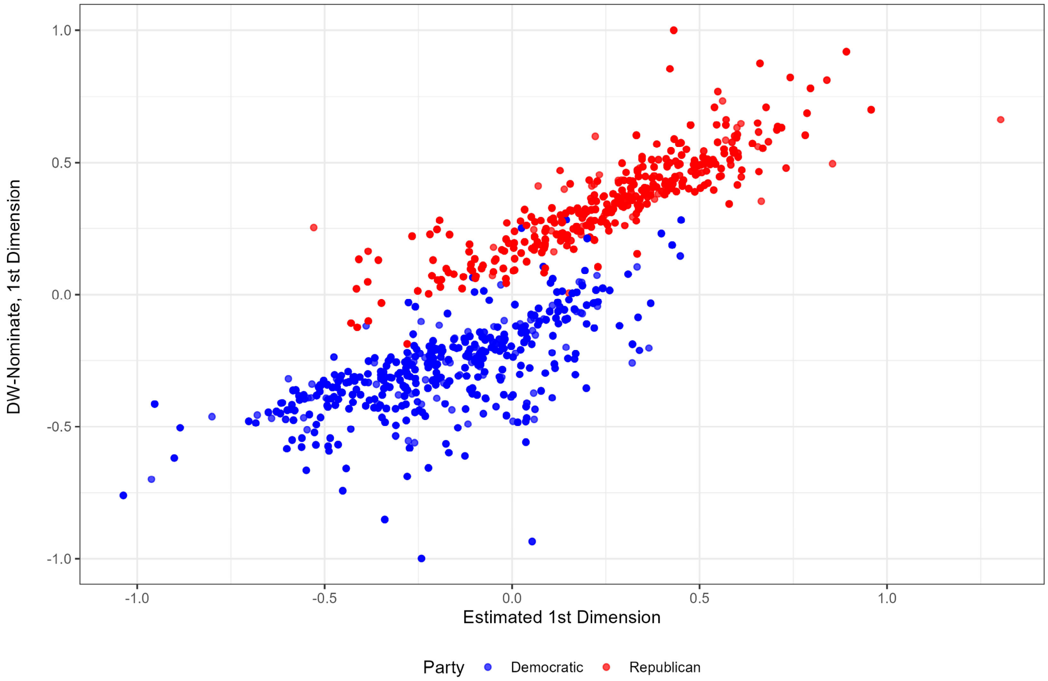
Appendix D: Additional Tables and Figures

Figure D1: Ideological Polarization Between Senate Members, 1927-2019 (2nd Dimension) - Senate 2D Model



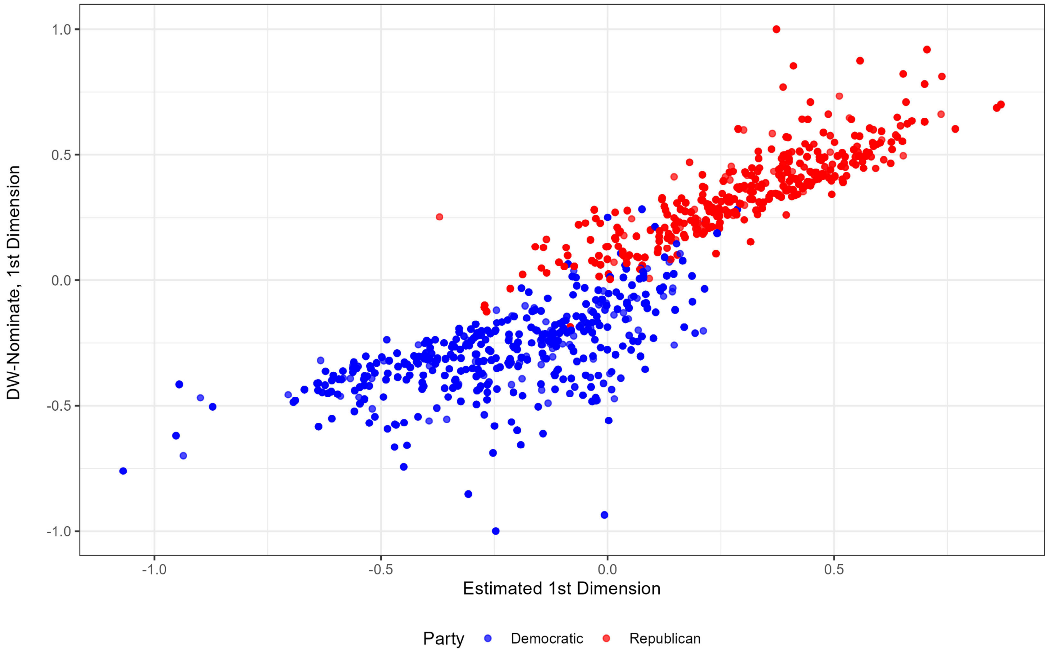
Notes: Estimates of the distance between party medians in the 2nd dimension for the Senate 2D Model are shown, together with a smoothed fit (Loess) curve.

Figure D2: Estimated (Senate 2D) Model vs. DW-Nominate, 1st Dimension



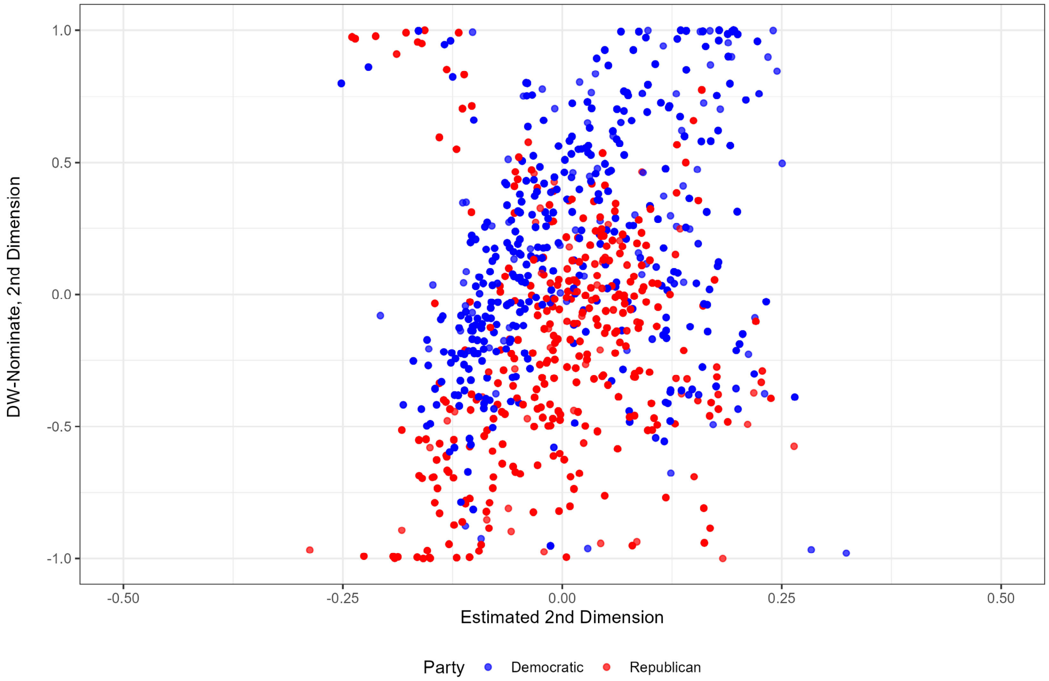
Notes: Scatter plot of first dimension estimated ideologies versus those from DW-Nominate, pooled across all Congresses. Democrats are shown in blue, Republicans are shown in red. The correlation is 0.866. The correlation within Republicans is 0.905, while the one within Democrats is 0.723.

Figure D3: Misspecified (Senate 2D) Model vs. DW-Nominate, 1st Dimension



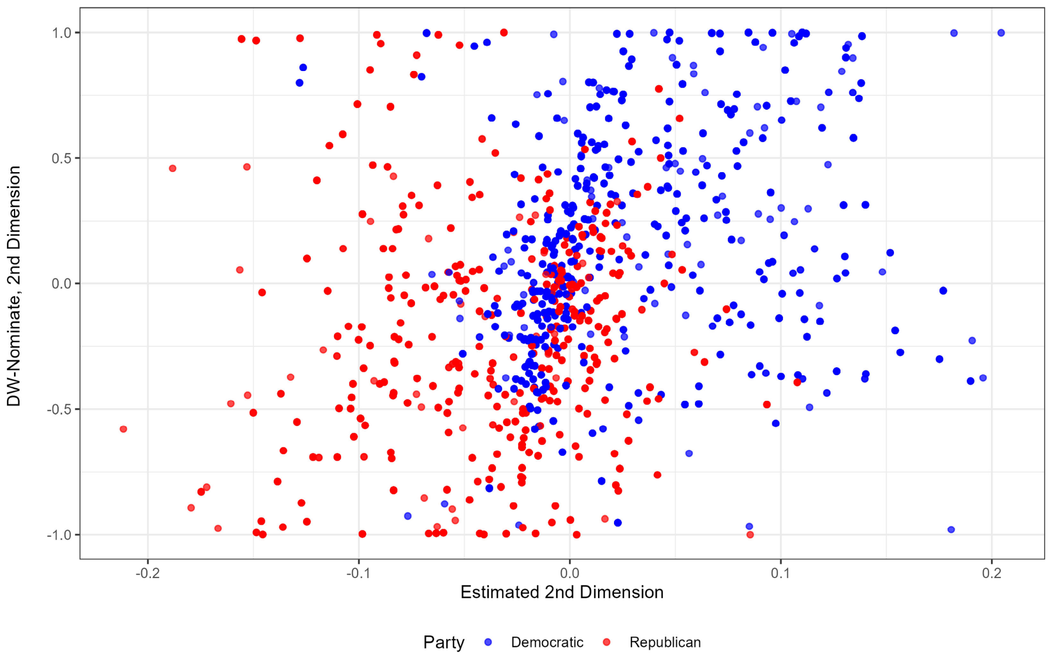
Notes: Scatter plot of the first dimension estimated ideologies of the misspecified model (no party pressure) versus those from DW-Nominate, pooled across all Congresses. Democrats are shown in blue, Republicans are shown in red. The correlation is 0.909.

Figure D4: Estimated (Senate 2D) Model vs. DW-Nominate, 2nd Dimension



Notes: Scatter plot of the second dimension estimated ideologies versus those from DW-Nominate, pooled across all Congresses. Democrats are shown in blue, Republicans are shown in red. The correlation is 0.412. The correlation within Republicans is 0.271, while the one within Democrats is 0.516.

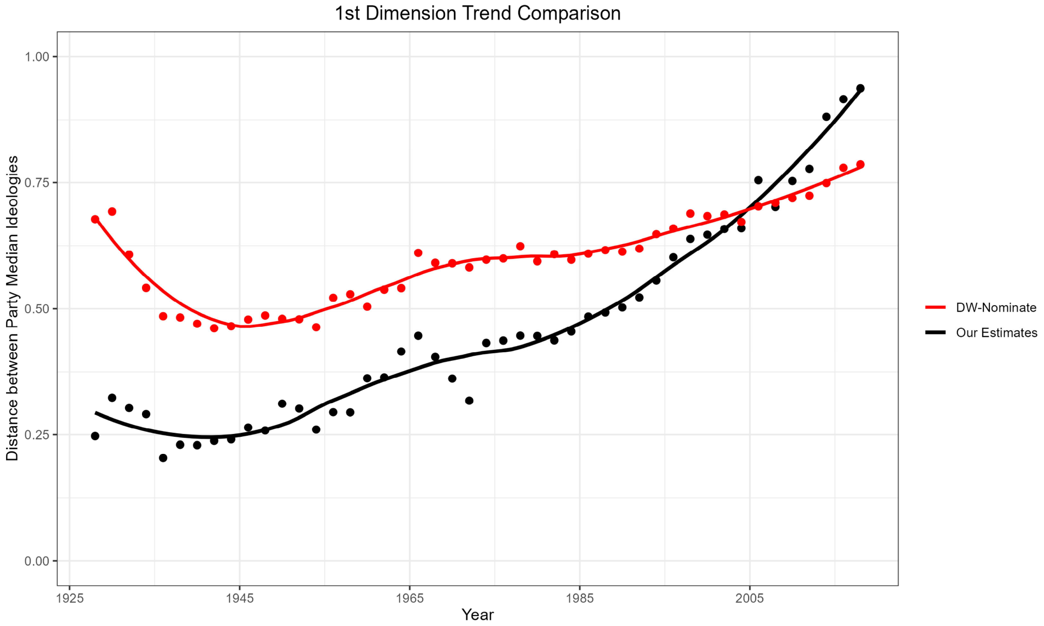
Figure D5: Misspecified (Senate 2D) Model vs. DW-Nominate, 2nd Dimension



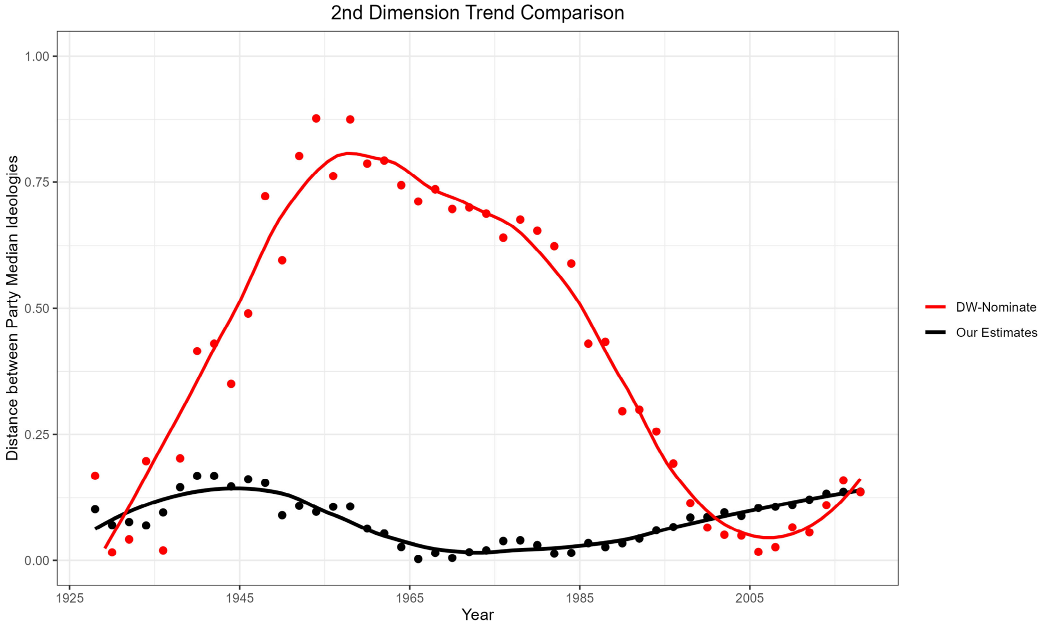
Notes: Scatter plot of the second dimension estimated ideologies of the misspecified model (no party pressure) versus those from DW-Nominate, pooled across all Congresses. Democrats are shown in blue, Republicans are shown in red. The correlation is 0.410.

Figure D6: Trends in Ideological Polarization: Senate 2D Model vs. DW-Nominate

(a) First Dimension



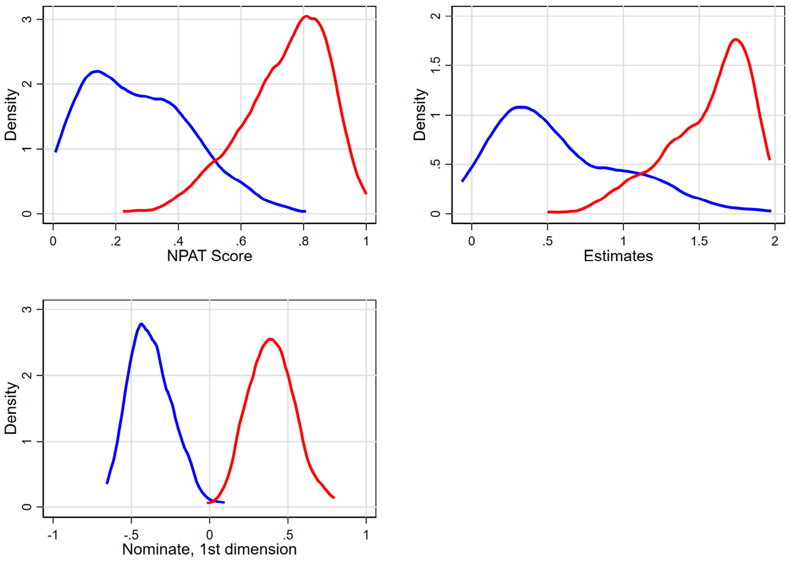
(b) Second Dimension



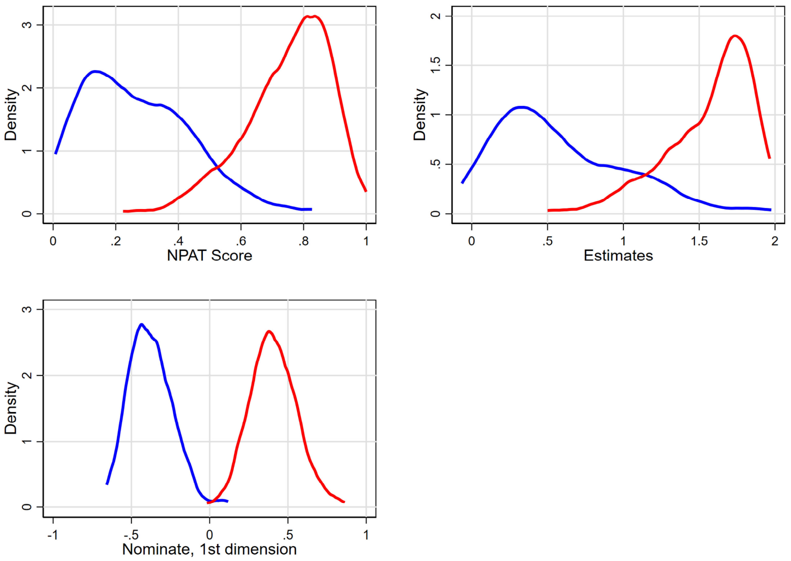
Notes: The two graphs compare the ideological polarization (difference between estimated party medians) across time for the Senate 2D model and DW-Nominate.

Figure D7: Trends in Ideological Polarization: Our Estimates, NPAT and DW-Nominate

(a) Congress 104, House

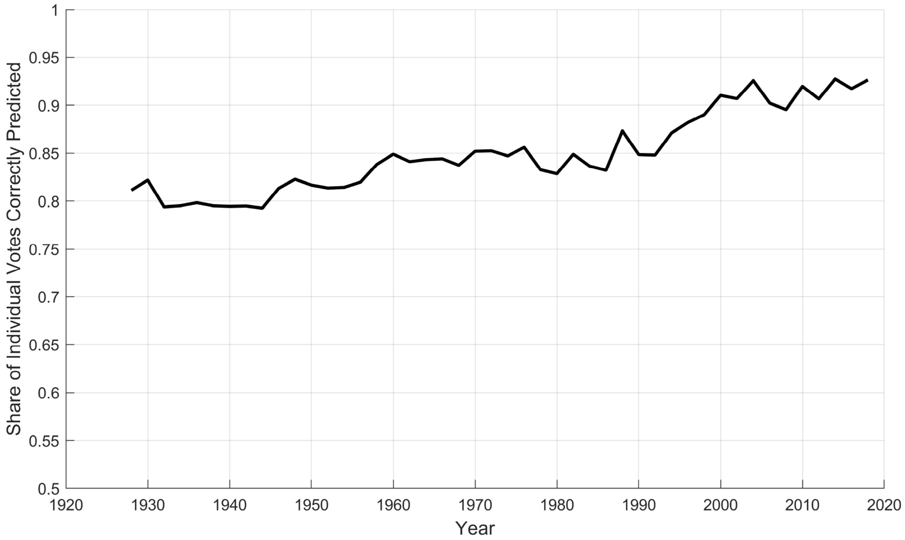


(b) Congress 105, House



Notes: The two graphs compare the distribution of Democrat and Republican ideologies across three different measures: NPAT (derived from House politicians' responses to policy surveys in 1996 and 1998), our estimates and DW-Nominate. We present the results for Congresses 104 and 105, the ones analyzed by Ansolabehere et al. (2001a,b). The figures show the distributions for the same politicians: while NPAT has less estimates than roll-call based methods, we only show the results for the set of politicians with NPAT scores for comparability.

Figure D8: Model Fit: Share of Votes Correctly Predicted in the Senate (2D Model)



Notes: Average share of votes that are correctly predicted in each Congress. A vote is considered to be correctly predicted if, under our estimated parameters, the probability of a congress member voting as observed in the data is larger than 0.5.

Figure D9: Ideological Polarization Over Time (2nd dimension), 1927-2019 - Senate 2D Model

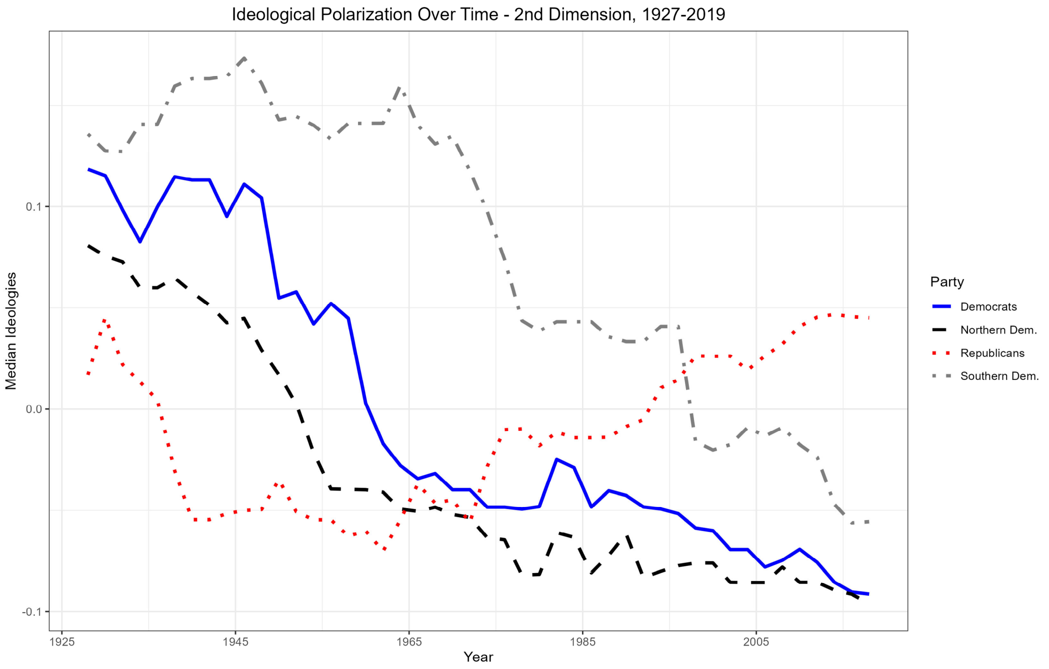
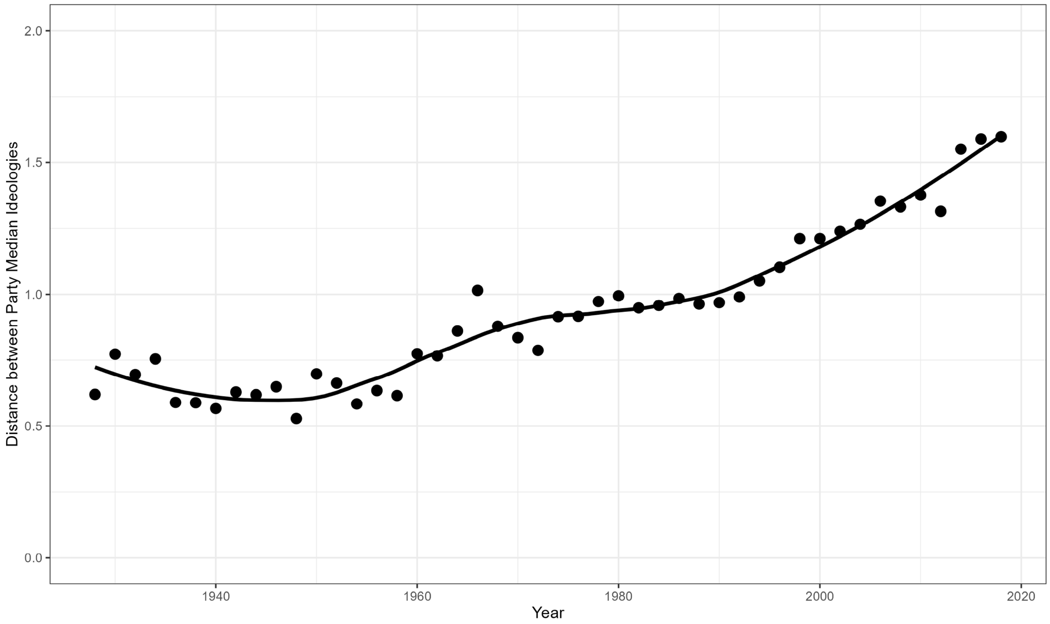


Figure D10: Ideological Polarization in the 1D Model

(a) Senate



(b) House

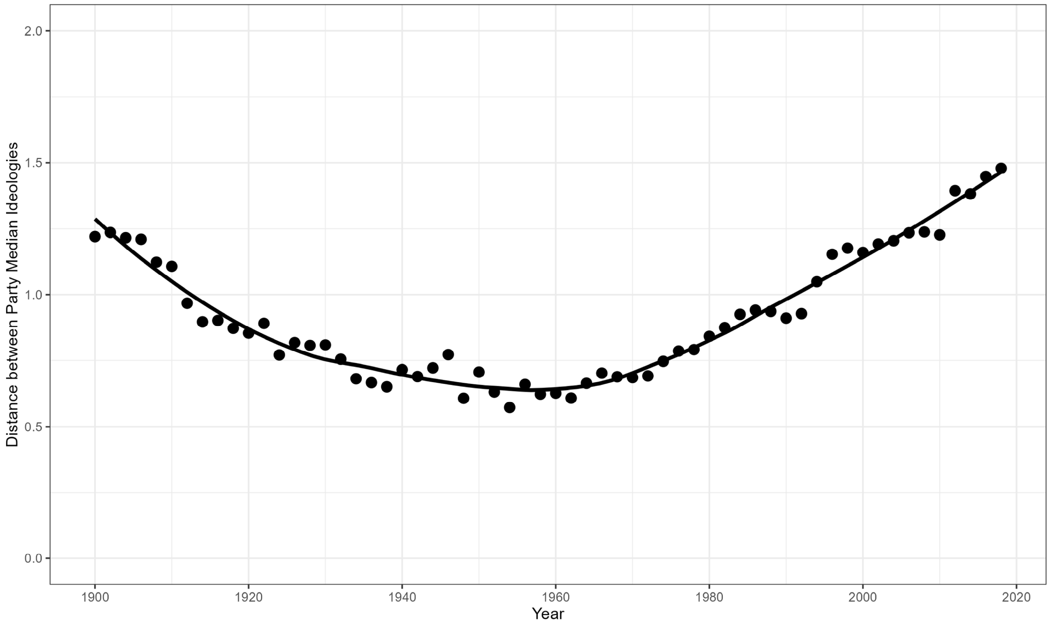
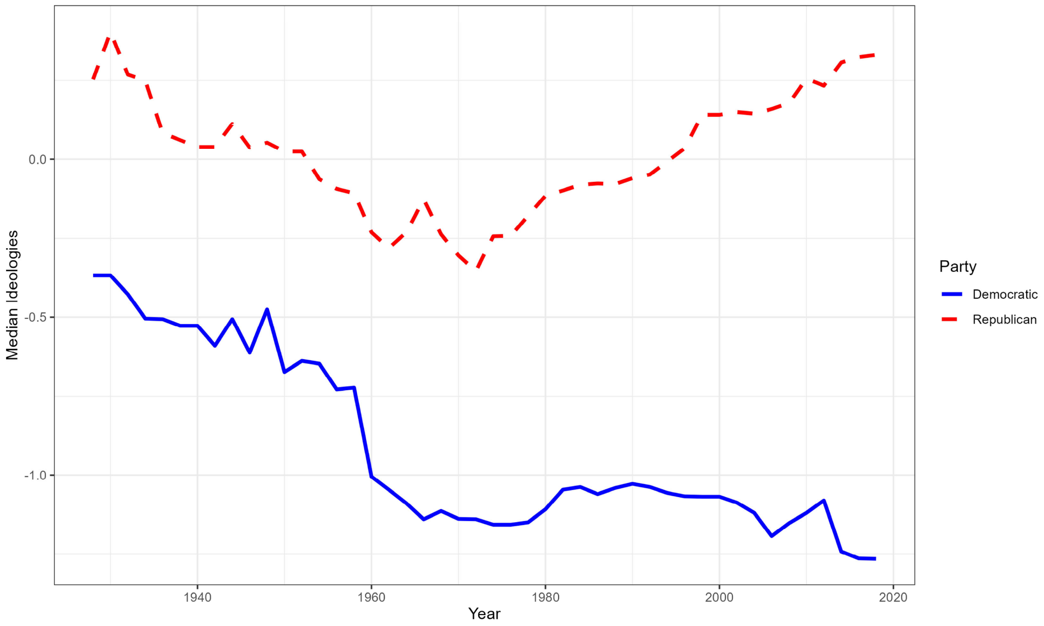


Figure D11: Ideological Polarization over Time, 1927-2019 - 1D Model

(a) Senate



(b) House

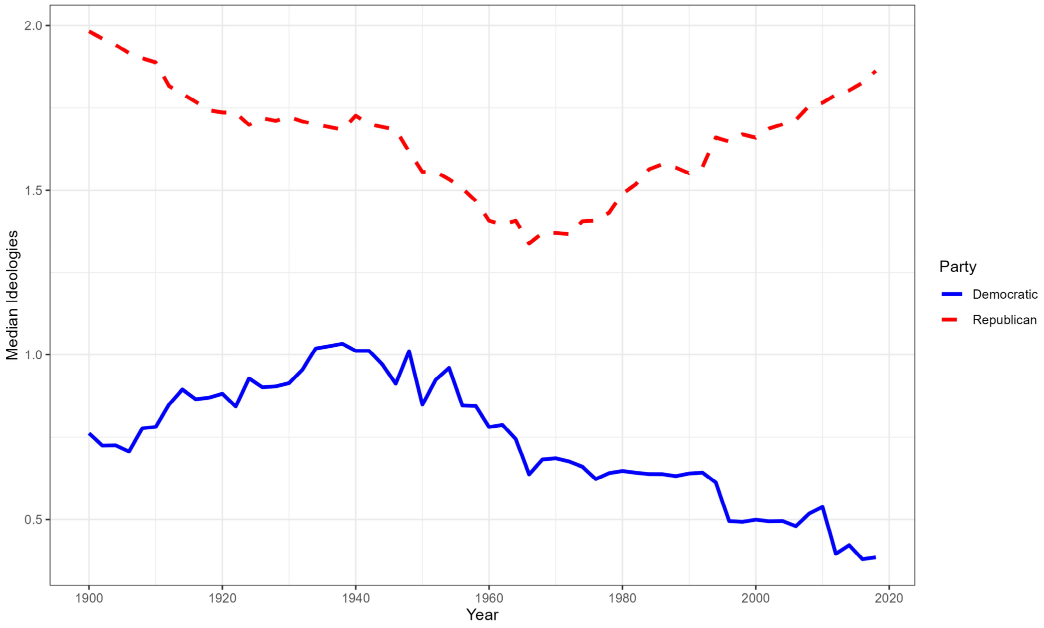
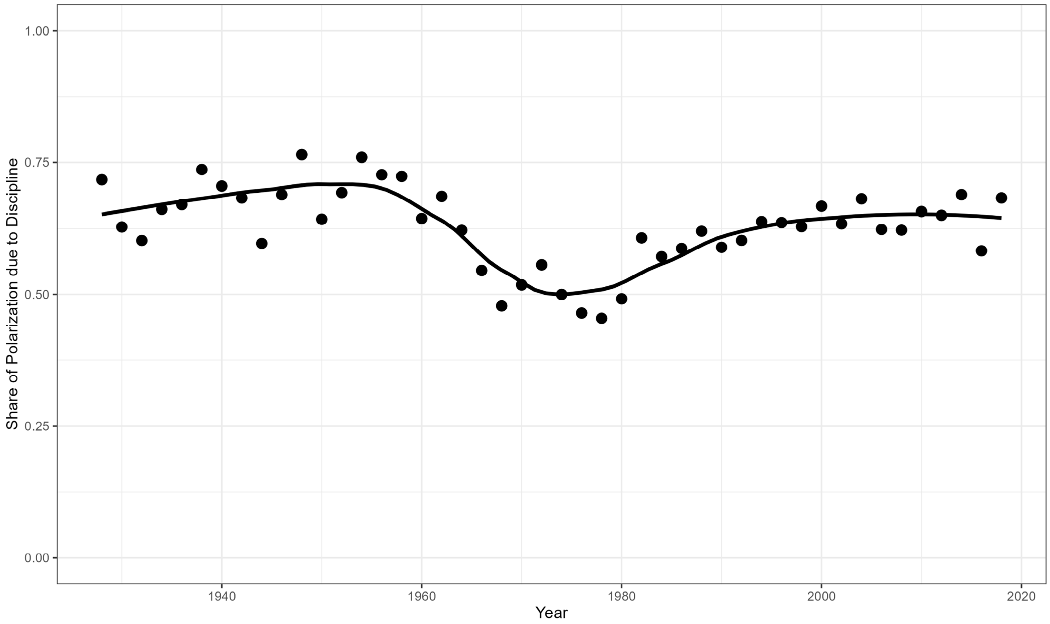


Figure D12: Share of Ideological Polarization Attributable to Party Pressure - 1D Model

(a) Senate



(b) House

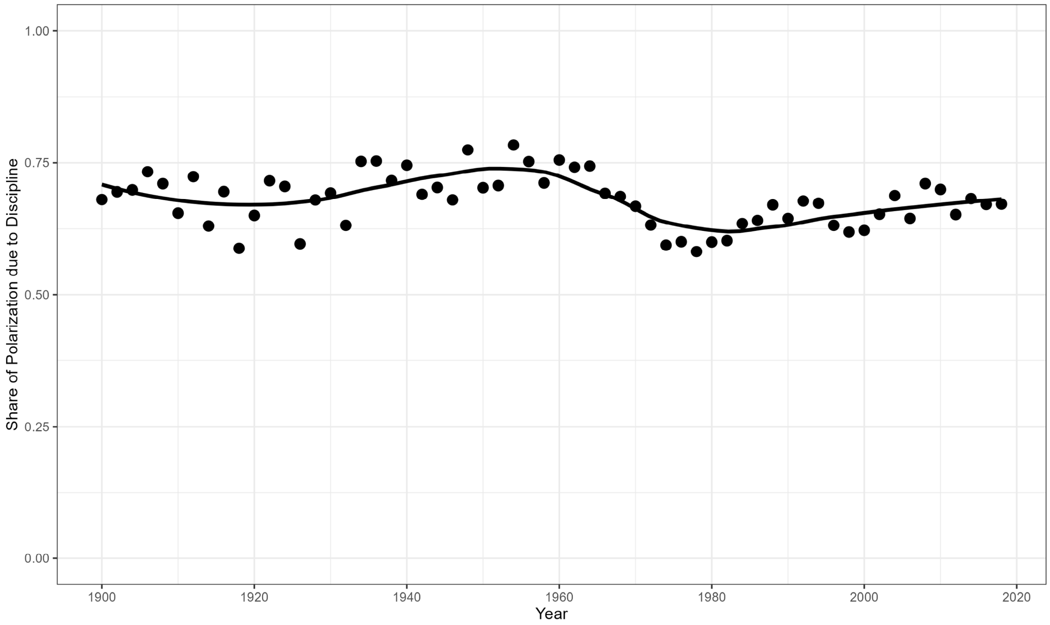
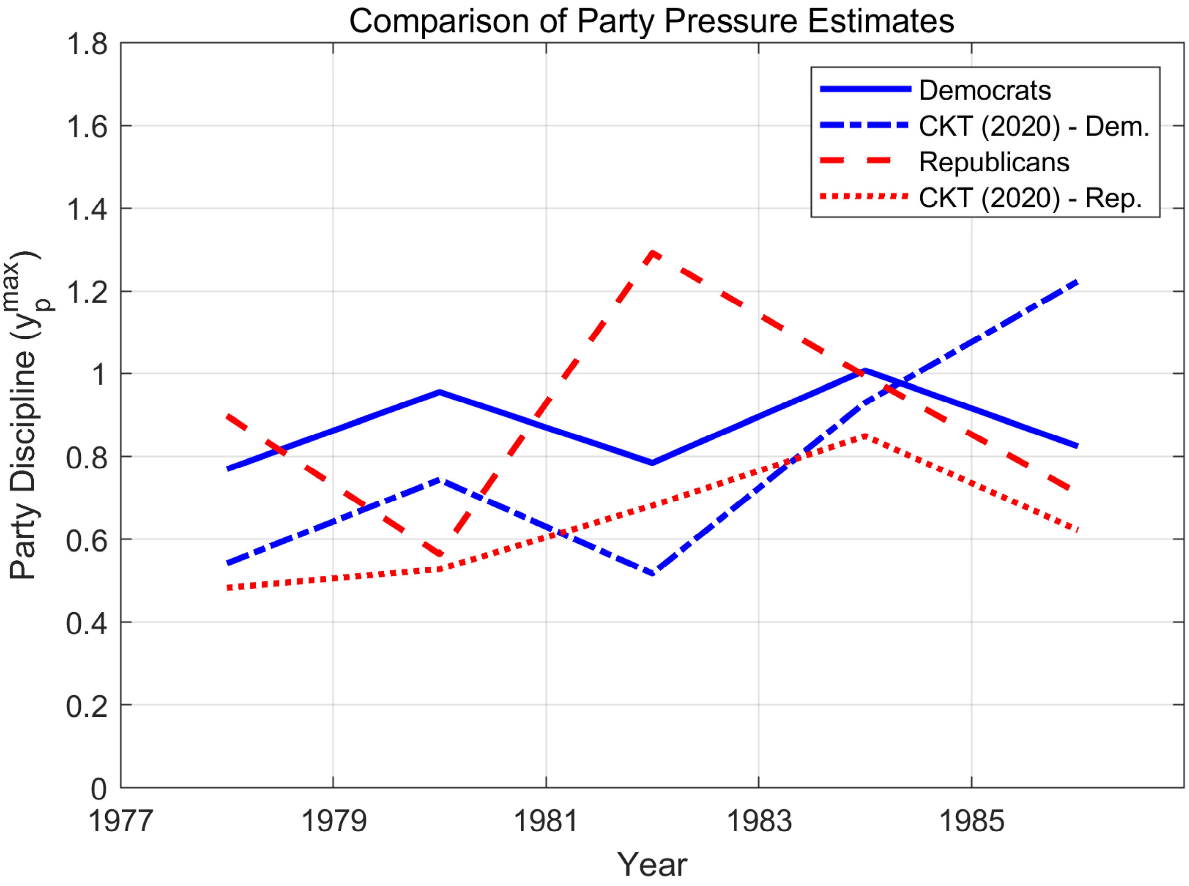
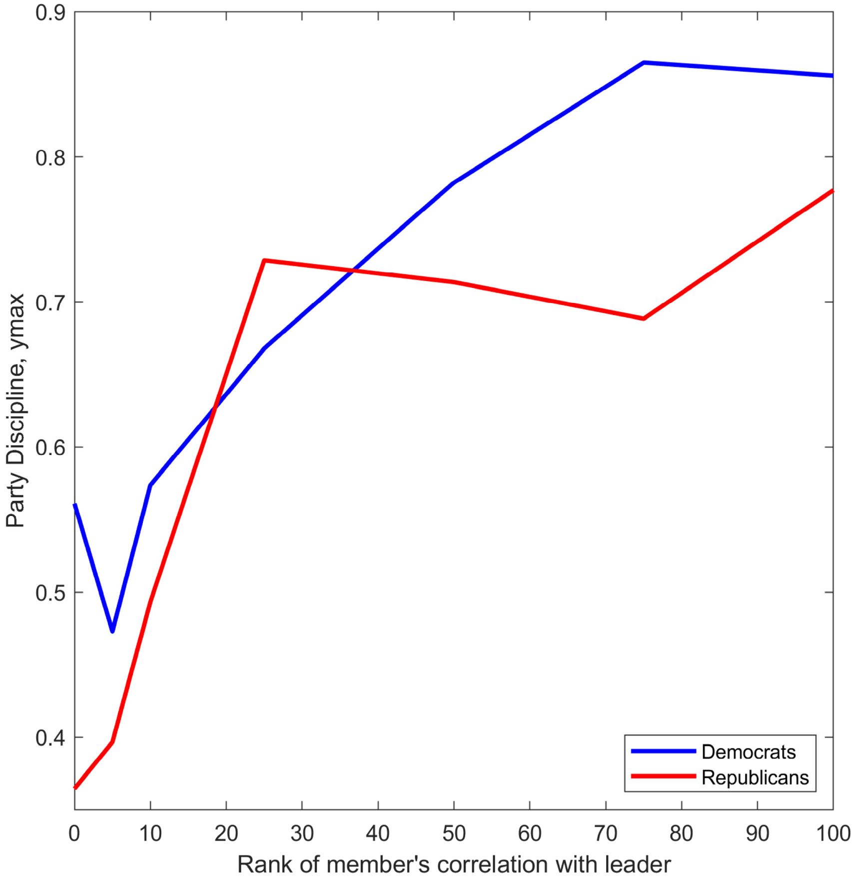


Figure D13: Comparison of Party Pressure Estimates



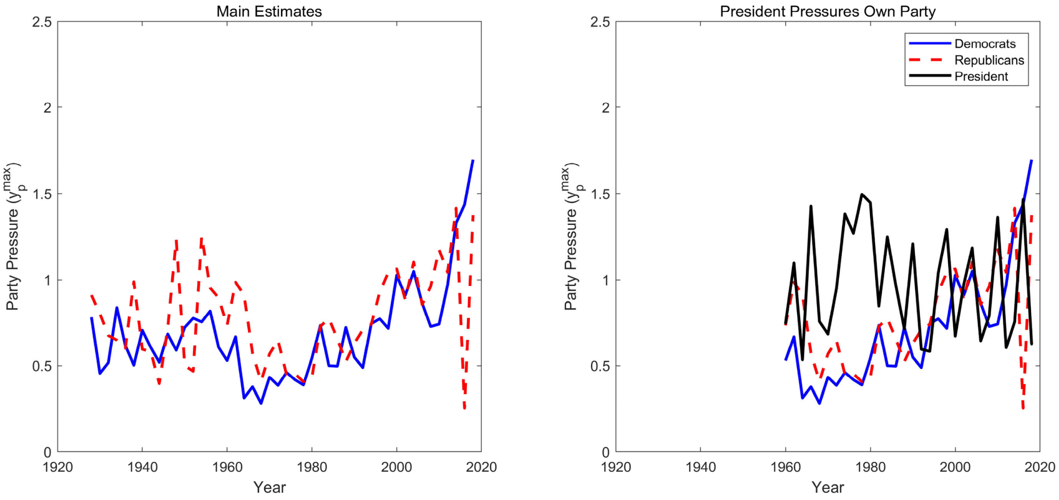
Notes: Estimates of y_p^{max} compared to those from Canen et al. (2020) for 1977-1986 (i.e. Congresses 95-99). Canen et al. (2020) assumed utility shocks have a variance equal to two (instead of one), so the prior estimates are rescaled by $\sqrt{2}$.

Figure D14: Party Pressure Estimates Using Members Other Than the Leader



Notes: Estimates of y_p^{max} when we use different party members to construct the discipline directions, $W_{p,t}$, in place of the true leader. The different party leaders are based on percentile ranks of the members whose votes are most correlated with the leader (100% being the leader). Estimates for Democrats are presented in blue, while estimates for Republicans are in red.

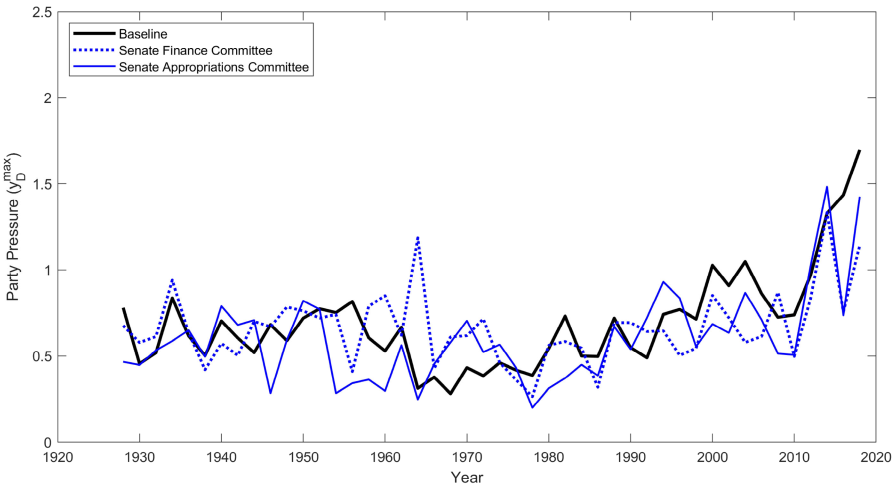
Figure D15: Party Pressure Estimates Allowing for Presidential Influence



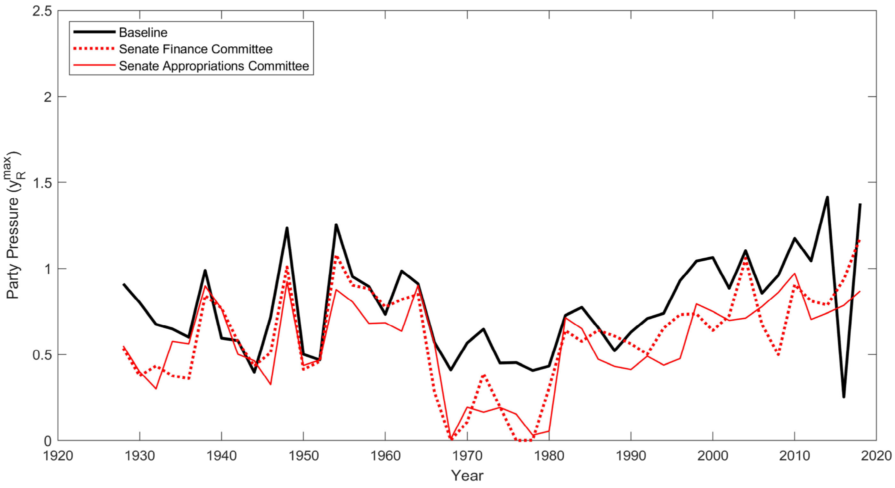
Notes: Estimates of y_p^{max} when we extend the model to allow the incumbent president to also pressure members of his own party. The left figure presents the baseline estimates for comparison purposes. The estimated parameter of the president's influence is shown in a black line in the right hand figure.

Figure D16: Party Pressure Estimates Allowing for Committee Influence

(a)
Democrats



(b) Republicans



Notes: Estimates of y_p^{max} when we replace the direction of party pressure obtained from party

Figure D17: Party Pressure Heterogeneity in the Presence of Re-Election Concerns

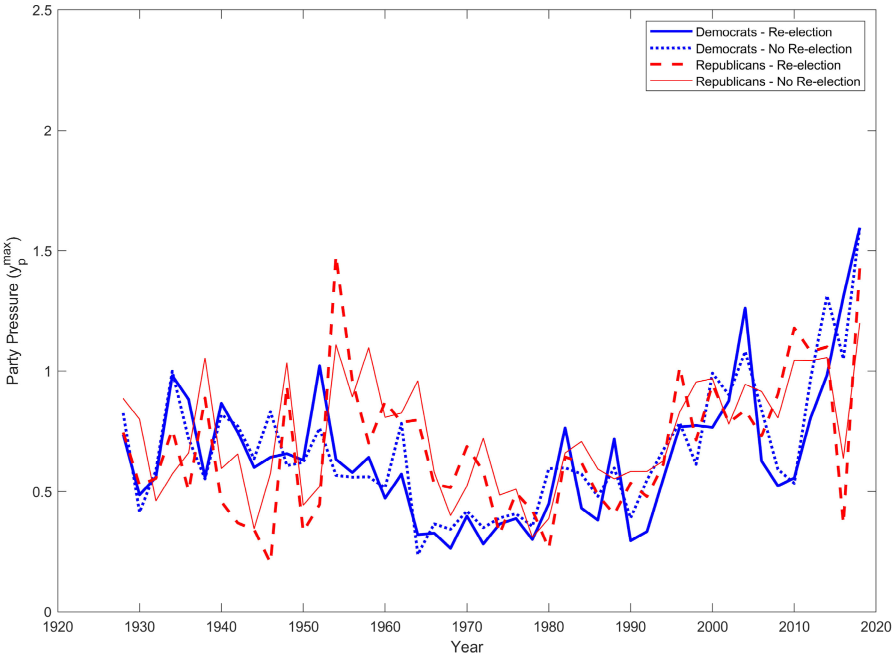
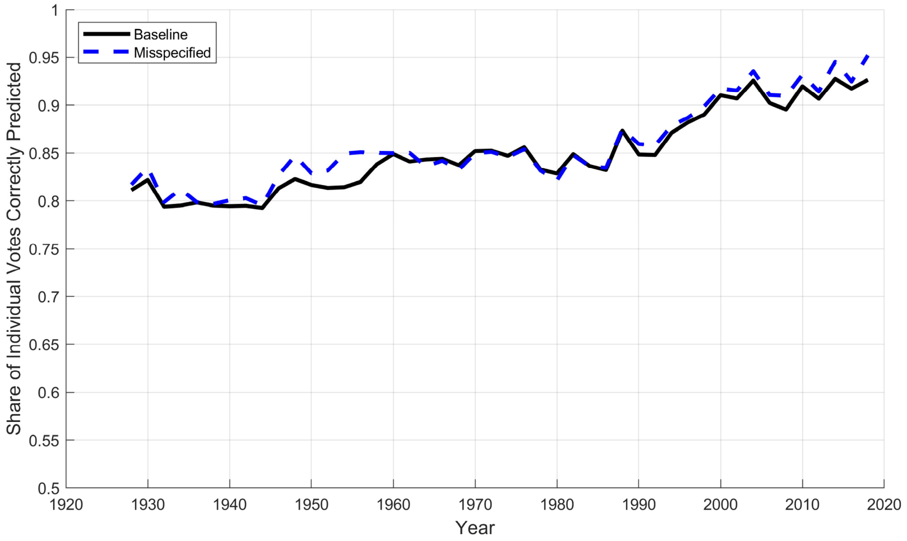


Figure D18: Comparison of Model Fit: Baseline Model and Misspecified Model



Notes: Average share of votes that are correctly predicted in each Congress for the baseline model (already shown in Figure D8) and for the misspecified model omitting party pressure. A vote is considered to be correctly predicted if, under our estimated parameters, the probability of a congress member voting as observed in the data is larger than 0.5.

Table D1: Summary Statistics

	Senate				House			
Congress	Bills introduced	Avg. bills per member	Bills passed	Fraction that pass	Bills introduced	Avg. bills per member	Bills passed	Fraction that pass
80th (1947-1948)	3,186	33.2	1,670	0.524	7,611	17.5	1,739	0.228
81st (1949-1950)	4,486	46.7	2,362	0.527	10,502	24.1	2,482	0.236
82nd (1951-1952)	3,665	38.2	1,849	0.505	9,065	20.8	2,008	0.222
83rd (1953-1954)	4,077	42.5	2,231	0.547	10,875	25.0	2,129	0.196
84th (1955-1956)	4,518	47.1	2,550	0.564	13,169	30.3	2,360	0.179
85th (1957-1958)	4,532	47.2	2,202	0.486	14,580	33.5	2,064	0.142
86th (1959-1960)	4,149	41.5	1,680	0.405	14,112	32.3	1,636	0.116
87th (1961-1962)	4,048	40.5	1,953	0.482	14,328	32.8	1,927	0.134
88th (1963-1964)	3,457	34.6	1,341	0.388	14,022	32.2	1,267	0.090
89th (1965-1966)	4,129	41.3	1,636	0.396	19,874	45.7	1,565	0.079
90th (1967-1968)	4,400	44.0	1,376	0.313	22,060	50.7	1,213	0.055
91st (1969-1971)	4,867	48.7	1,271	0.261	21,436	49.3	1,130	0.053
92nd (1971-1972)	4,408	44.1	1,035	0.235	18,561	42.7	970	0.052
93rd (1973-1974)	4,524	45.2	1,115	0.246	18,872	43.4	923	0.049
94th (1975-1976)	4,115	41.2	1,038	0.252	16,982	39.0	968	0.057
95th (1977-1978)	3,800	38.0	1,070	0.282	15,587	35.8	1,027	0.066
96th (1979-1980)	3,480	34.8	976	0.280	9,103	20.9	929	0.102
97th (1981-1982)	3,396	34.0	786	0.231	8,094	18.6	704	0.087
98th (1983-1984)	3,454	34.5	936	0.271	7,105	16.3	978	0.138
99th (1985-1986)	3,386	33.9	940	0.278	6,499	14.9	973	0.150
100th (1987-1988)	3,325	33.3	1,002	0.301	6,263	14.4	1,061	0.169
101st (1989-1990)	3,669	36.7	980	0.267	6,664	15.3	968	0.145
102nd (1991-1992)	3,738	37.4	947	0.253	6,775	15.6	932	0.138
103rd (1993-1994)	2,805	28.1	682	0.243	5,739	13.2	749	0.131
104th (1995-1996)	2,266	22.7	518	0.229	4,542	10.4	611	0.135
105th (1997-1998)	2,718	27.2	586	0.216	5,014	11.5	710	0.142
106th (1999-2000)	3,343	33.4	819	0.245	5,815	13.4	957	0.165
107th (2001-2002)	3,242	32.4	554	0.171	5,892	13.5	677	0.115
108th (2003-2004)	3,078	30.8	759	0.247	5,547	12.8	801	0.144
109th (2005-2006)	4,163	41.6	684	0.164	6,540	15.0	770	0.118
110th (2007-2008)	3,738	37.4	556	0.149	7,441	17.1	1101	0.148
111th (2009-2010)	4,101	41.0	176	0.043	6,677	15.3	861	0.129
112th (2011-2012)	3,767	37.7	364	0.097	6,845	15.7	561	0.082

Table D2: Regression Results - Sources of Party Pressure

	Estimates of y_p^{max}				
Party (Republican)	0.085 (0.059)	0.085 (0.070)	0.085 (0.059)	0.085 (0.071)	0.085 (0.051)
Majority Status		-0.000 (0.070)		-0.000 (0.071)	-0.000 (0.050)
Divided Government (1 if Divided)			0.040 (0.059)	0.040 (0.059)	0.077 (0.046)
Observations	92	92	92	92	92
Decade Fixed Effect					Yes
R^2	0.023	0.023	0.028	0.028	0.566

Notes: Regressions of the time series of estimates of $\{y_p^{max}\}_{p\in\{D,R\}}$ for the Senate 2D model on a Party level dummy variable (equal to 1 if p is Republican), dummy variable for Majority Status (which equals 1 if party p held the majority of seats in the Senate, and 0 otherwise) and dummy variable for divided government (which is equal to 0 if the president's party is the same as the majority party in the House and in the Senate and 1 otherwise). Robust standard errors in parentheses.

Appendix E: Monte Carlo Simulations

As shown in the main text, we can identify the parameters for individual ideologies, $\{\theta^i\}$, and for party pressure, (y_D^{max}, y_R^{max}), under general forms of agenda-setting within the random utility framework when preference shocks are realized after the agenda is set. Here, we report Monte Carlo experiments to verify that the parameters we recover are independent of the agenda.

We use an analogous set-up to that of (Clinton et al., 2014). First, we set the number of politicians to $n = 100$, with two parties: a majority party (D) with 55 members and a minority party (R) with 45 members. We draw ideologies i.i.d. from the following normal distributions: for party D, $\theta^i \sim N(-\frac{\alpha}{2},1)$ and for party R, $\theta^i \sim N(-\frac{\alpha}{2}, 1)$. Hence, α parameterizes ideological polarization: the distance between parties' ideological distributions. We vary α across simulations to illustrate unbiased estimates even with polarized ideologies (larger α). We set the party leaders at the median of their respectively drawn distributions.

Preferences follow equation (1) in the main text for the one-dimensional model, with voting decisions being made analogous to equation (4) in Section 2.2.3. Preference shocks are drawn i.i.d. from a standard normal distribution.[8]

Our goal is to illustrate that we can obtain unbiased estimates with general forms of agenda-setting. To do so, we draw cutpoints *at random*: i.i.d., from a Normal distribution with mean zero and standard deviation σ_v. The parameter σ_v parameterizes the partisanship in agenda-setting: low values imply most cutpoints pin parties against each other. The ratio $\frac{\alpha}{\sigma_v}$ parameterizes the share of cutpoints between the party medians (i.e., where leaders exert pressure in opposite directions). We draw $T = 750$ cutpoints, which is close to the median value in our sample.

We consider a total set of eight exercises. First, we consider two sets of party pressure parameters: $(y_{D,}^{max} y_R^{max}) = (0.6, 0.4)$ and $(y_{D,}^{max} y_R^{max}) = (0, 0)$. The former are similar to those estimated within our sample, while the latter are used to illustrate that we obtain estimates of no party pressure if it does not exist. Second, we set $\alpha = 0.5$ (parties are not very polarized) and $\alpha = 1$ (parties are polarized, as their medians are 1 standard deviation of preference shocks apart). Finally, we set $\sigma_v = 0.5$ (cutpoints are concentrated between party medians) and $\sigma_v = 1$ (cutpoints are often drawn from extreme agendas). Our exercises consist of all possible combinations of the two different sets of parameters for party pressure, ideological polarization, and agenda-setting.

The results are shown in Table 4 below for $R = 100$ simulations. As we see, our procedure estimates party pressure parameters accurately and consistently across specifications: whether party pressure is zero or positive, whether agenda-setting is extreme or not, and whether ideological polarization is large or small. These results illustrate the robustness of obtaining unbiased estimates of both polarization and party pressure, provided the sample sizes are as large as in our

---

[8] The standard normal assumption on preference shocks is necessary for identification - see the main text.

data. By contrast, in small samples we may only rarely observe switching for some politicians. In this case, estimates can be biased as Bateman et al. (2017) demonstrate.

Table E3: Monte Carlo Simulation Results for Estimation of $\bar{y}^{max}$

Specification: $(y^D_{max}, y^R_{max}) = (0.6, 0.4)$

	Low Polarization ($\alpha = 0.5$)	High Polarization ($\alpha = 1$)	Low Polarization ($\alpha = 0.5$)	High Polarization ($\alpha = 1$)
	Divisive Agenda ($\sigma_v = 0.5$)	Divisive Agenda ($\sigma_v = 0.5$)	Extreme Agenda ($\sigma_v = 1$)	Extreme Agenda ($\sigma_v = 1$)
y^{max}_D	0.609	0.613	0.603	0.603
	(0.018)	(0.015)	(0.017)	(0.017)
y^{max}_R	0.405	0.419	0.399	0.394
	(0.016)	(0.020)	(0.022)	(0.020)

Specification: $(y^D_{max}, y^R_{max}) = (0, 0)$

	Low Polarization ($\alpha = 0.5$)	High Polarization ($\alpha = 1$)	Low Polarization ($\alpha = 0.5$)	High Polarization ($\alpha = 1$)
	Divisive Agenda ($\sigma_v = 0.5$)	Divisive Agenda ($\sigma_v = 0.5$)	Extreme Agenda ($\sigma_v = 1$)	Extreme Agenda ($\sigma_v = 1$)
y^{max}_D	0.004	0.007	0.005	0.004
	(0.007)	(0.010)	(0.008)	(0.007)
y^{max}_R	0.005	0.017	0.005	0.003
	(0.008)	(0.013)	(0.007)	(0.006)

Notes: Low ideological polarization refers to a distance between the party medians in the ideological distributions of $\alpha = 0.5$, while high polarization refers to $\alpha = 1$. σ_v refers to the standard deviation in cutpoints, which are drawn at random. This can be low ($\sigma_v = 0.5$) or high ($\sigma_v = 1$), capturing the extent to which most observations are between the party medians or beyond them. The two panels refer to different values of the party pressure parameters: in the first, party pressure exists for both parties at values similar to estimated values. In the second, there is no party pressure. Entries present the average estimates across simulations, while parentheses refer to the standard deviation in estimates across $R = 100$ simulations.

References

Ansolabehere, S., J. M. Snyder Jr, and C. Stewart III (2001a). Candidate positioning in us house elections. *American Journal of Political Science*, 136–159.

Ansolabehere, S., J. M. Snyder Jr, and C. Stewart III (2001b). The effects of party and preferences on congressional roll-call voting. *Legislative Studies Quarterly*, 533–572.

Armstrong, D. A., R. Bakker, R. Carroll, C. Hare, K. T. Poole, H. Rosenthal, et al. (2014). *Analyzing spatial models of choice and judgment with R.* CRC Press.

Bateman, D. A., J. Clinton, and J. S. Lapinski (2017). A house divided? political conflict and polarization in the u.s. congress, 1877- 2011. *American Journal of Political Science 61*(3), 698–714.

Boche, A., J. B. Lewis, A. Rudkin, and L. Sonnet (2018). The new voteview.com: preserving and continuing keith poole's infrastructure for scholars, students and observers of congress. *Public Choice 176*(1-2), 17–32.

Canen, N., C. Kendall, and F. Trebbi (2020). Unbundling polarization. *Econometrica 88*(3), 1197–1233.

Carroll, R., J. B. Lewis, J. Lo, K. T. Poole, and H. Rosenthal (2009). Measuring bias and uncertainty in dw-nominate ideal point estimates via the parametric bootstrap. *Political Analysis 17*(3), 261–275.

Clinton, J., I. Katznelson, and J. Lapinski (2014). Where measures meet history: Party polarization during the new deal and fair deal. *Governing in a Polarized Age: Elections, Parties, and Representation in America*.

Kingma, D. P. and J. Ba (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Potthoff, R. F. (2018). Estimating ideal points from roll-call data: explore principal components analysis, especially for more than one dimension? *Social Sciences 7*(1), 12.

Rivers, D. (2003). Identification of multidimensional spatial voting models. *Typescript. Stanford University*.