Supplementary Appendix to
"Backward Induction in the Wild?
Evidence from Sequential Voting in the U.S. Senate"

Contents

List of Figures

List of Tables

Appendix C: Remaining Proofs

C.1. *Controlling for the Votes of Copartisans Introduces Bias*

In Section 3.2, we state that controlling for the votes of copartisans introduces bias into the OLS estimates. Here, we state this claim more formally and provide an analytic expression for the bias.

Consider the following econometric model:

$$d_{i,c} = \alpha + \beta o_{i,c} + \gamma \overline{d}_{-i,c} + \eta_{i,c},$$

where $\overline{d}_{-i,c}$ denotes the share of i's copartisans who deviated on roll call c, and all other symbols are as defined in the main text. To derive the probability limit of the OLS estimator, let $o_{i,c} = r_0 + r_1\overline{d}_{-i,c} + R$ be the linear projection of $o_{i,c}$ on the space spanned by $\overline{d}_{-i,c}$ and a constant, and let $d_{i,c} = q_0 + q_1\overline{d}_{-i,c} + Q$ be the linear projection of $d_{i,c}$ on the same space. By the Frisch-Waugh Theorem (Frisch and Waugh 1933)

$$\begin{aligned} \text{plim } \widehat{\beta}_{OLS} &= \frac{\text{Cov}(Q,R)}{\text{Var}\,(R)} \\ &= \beta + \frac{\text{Cov}(\eta_{i,c}, o_{i,c} - r_1\overline{d}_{-i,c})}{\text{Var}\,(R)}. \end{aligned}$$

Hence, $\widehat{\beta}_{OLS}$ is unbiased if only if $\text{Cov}(\eta_{i,c}, o_{i,c} - r_1\overline{d}_{-i,c}) = 0$.

By construction, $r_1 = \frac{\text{Cov}\left(\overline{d}_{-i,c}, o_{i,c}\right)}{\text{Var}\left(\overline{d}_{-i,c}\right)}$. Thus, when senators views on a particular bill do not depend on the vote order, i.e., when $\text{Cov}(\eta_{i,c}, o_{i,c}) = 0$, then $\widehat{\beta}_{OLS}$ will be an upper bound on the true vote order effect if and only if $\text{Cov}\left(\overline{d}_{-i,c}, o_{i,c}\right)\text{Cov}(\eta_{i,c}, \overline{d}_{-i,c}) < 0$.

To gain some intuition for when we would expect this condition to hold, note that negative vote order effects imply that senators who vote later in the order are more likely to vote with the party line. Thus, when the true vote order effect is negative, then $\overline{d}_{-i,c}$ is, on average, smaller when $o_{i,c}$ is small, and $\text{Cov}\left(\overline{d}_{-i,c}, o_{i,c}\right) > 0$.

The sign of $\text{Cov}\left(\overline{d}_{-i,c}, \eta_{i,c}\right)$ depends on whether one senator deviating from the party line crowds in or crowds out defections among her copartisans. If senators strategically preempt each other, so that, on average, defection by one senator results in fewer deviations among her copartisans then $\text{Cov}\left(\overline{d}_{-i,c}, \eta_{i,c}\right) < 0$. Coupled with negative vote order effects this would imply that plim $\widehat{\beta}_{OLS} > \beta$.

C.2. *Vote Order Effects with $S = 5$ and Overlapping Ideal Points*

To show that the assumption of nonoverlapping ideal points is not strictly necessary for Proposition 2(a), which predicts negative vote-order effects, we now consider a variant of our model of "parties as teams" with five senators, i.e., $S = 5$, and no such restriction on preferences.

Note, regardless of whether senators' ideal points do or do not overlap, every senator is either an unconflicted yea voter (Y), a unconflicted nay voter (N), a conflicted yea voter ($\widetilde{Y}$), or a conflicted nay voter ($\widetilde{N}$). Furthermore, conditional on knowing a senator's type, α_i and δ contain no relevant information for determining her choices. Thus, equilibrium behavior is uniquely determined the by the "preference configuration," i.e., the profile of senator's types and the order in which they vote.

We assume simple majority rule, i.e., $\underline{y} = \underline{n} = 3$, and restrict attention to all 252 configurations which correspond to the case considered in Proposition 2(a), i.e., $|Y| < \underline{y} < |Y| + \left|\widetilde{Y}\right|$ or $|N| < \underline{n} < |N| + \left|\widetilde{N}\right|$. Consistent with the language in the main text, we say that a senator "deviates" if she belongs to $\widetilde{Y}$ and votes "nay" or if she belongs to $\widetilde{N}$ and chooses "yea."

Appendix Table A.13 lists all preference configurations for which $|Y| < \underline{y} < |Y| + \left|\widetilde{Y}\right|$ (left-most colums), the equilibrium choices of agents (middle columns), as well as indicators that highlight deviations (right-most columns). (Limiting ourselves to $|Y| < \underline{y} < |Y| + \left|\widetilde{Y}\right|$ is without further loss of generality as results for the case of $|N| < \underline{n} < |N| + \left|\widetilde{N}\right|$ can be obtained by a simple relabeling.) Counting the number of deviations at each position in the vote order shows that they are more frequent early in the order than late. Thus, if one believes that all preference profiles are *a priori* equally likely, then one would expect to find negative vote-order effects on average.

Put differently, a Bayesian with no prior information on which preference profiles are more prevelant among senators than others would aggregate over the set of possible configurations by taking simple means, which would result in a posterior with monotonically declining deviation rates.

Of course, there exist preference profiles for which deviations are nonmonotonic in the vote order, and even a (smaller) number of configurations in which only senators at the end of the order deviate. Thus, if one believes that the latter configurations are sufficiently more likely to occur than others, then negative vote-order effects need not necessarily obtain. Nonetheless, the point of our argument is that even a model that allows for the ideal points of senators to overlap has the potential to rationalize our main empirical finding. Nonoverlapping ideal points are sufficient to generate negative vote-order effects but not necessary.

Appendix D: Parties as Teams with Low Enough Uncertainty

This appendix illustrates by way of example how uncertainty may lead to situations in which senators early in the vote order may prefer to "play it safe" rather than deviate from the party line. At the same time, the appendix demonstrates that if uncertainty is low enough (relative to the gains from deviating), then conflicted senators' optimal strategies coincide with those in the common knowledge environment considered in Proposition 1.[1]

To see that there may be situations in which senators prefer to "play it safe" rather than preempt their colleagues, consider the game in Figure A.4. Here, the Democratic party requires one more "yea" for the bill to pass. Senator $D1$ is conflicted in the sense that she would like to vote "nay," but only if her vote was not pivotal. In one state of nature, Senator $D2$'s disdain for the bill is so strong that she will vote "nay" regardless of whether this causes the bill to fail. In the other

[1]This does not necessarily mean that the outcome of the game will be the same, as the outcome depends also on *realized* uncertainty. It does imply, however, that senators whose preferences make them "conflicted" (i.e., for whom $|\alpha_i| < |\delta_p|$ and $\alpha_i \delta_p < 0$) would choose the same action at any node in the game tree as in the game with common knowledge.

state of nature, $D2$ would support the bill if need be. Alternatively, one can think of the states of nature as determining whether $D2$ is conflicted and rational. In the first state, $D2$ is conflicted but irrational in the sense that she would always vote "nay." In the second state, $D2$ is rational and plays the equilibrium strategy specified by Proposition 1. Figure A.4 is, therefore, also an example of behavior when common knowledge of rationality fails.

Let q be the probability that $D2$ defects no matter what, i.e., $q \equiv \Pr(\alpha_2 < -\delta)$. Given the payoffs specified in the figure, $D1$ deviates from the party line if $q < \frac{1}{2}$. More generally, there will be a cutoff value $\overline{p} = \frac{\alpha}{\delta}$, such that $D1$ finds it worthwhile (in expectation) to defect if only if $q < \overline{p}$. As a consequence, there may be nodes in the game tree at which senators "play it safe."

To see that "low enough uncertainty" would lead to senators adopting the same strategies as in the common knowledge environment, focus on a conflicted Democrat, i, and fix her position-taking payoff from voting "yea" $\alpha_i = \alpha < 0$. Further, let $\delta \equiv \delta_D$ and assume a constant probability of "mistakes", i.e., that senators further down in the vote order deviate from their party's position with probability q even if that caused the roll call to be lost for sure.

We first show that if q is small enough such that $\alpha + \delta < (1-q)^{\left|\widetilde{Y} \cup Y\right|_{i'>i}} \delta$, then that conflicted Democrat would defect whenever $|Y|_{i'>i} + \left|\widetilde{Y}\right|_{i'>i} + 1 > \underline{y} - \overline{y}_i$. In words, with low enough uncertainty the senator would prefer to rely on the support of all $\left|\widetilde{Y} \cup Y\right|_{i'>i}$ remaining copartisans to still carry the bill to passage rather than to vote for the bill herself (even if it that guaranteed passage and the measure might otherwise fail). Since $\alpha < 0$, $\alpha + \delta > 0$ and $\left|\widetilde{Y} \cup Y\right|_{i'>i} < \infty$, there exists a $q_i^1 > 0$, such that for all $q < q_i^1$ the condition above is always satisfied.

To establish that the senator would deviate from the party line when $|Y|_{i'>i} + \left|\widetilde{Y}\right|_{i'>i} + 1 < \underline{y} - \overline{y}_i$ it suffices for q to be small enough such that $0 > \alpha + \left[1 - (1-q)^{\left|\widetilde{N} \cup N\right|_{i'>i}}\right] \delta$, i.e., the senator would rather defect and lose the roll for sure call than hope for the opposition to make at least one mistake. Again, given that $\alpha < 0$, $\alpha + \delta > 0$ and $\left|\widetilde{N} \cup N\right|_{i'>i} < \infty$, there exists a $q_i^2 > 0$ such that for all $q < q_i^2$ the condition holds.

Lastly, when $|Y|_{i'>i} + \left|\widetilde{Y}\right|_{i'>i} + 1 = \underline{y} - \overline{y}_i$ the conflicted senator would vote *with* the party line if q was small enough such that $\left[1 - (1-q)^{\left|\widetilde{N} \cup N\right|_{i'>i}}\right] \delta \leq \alpha + (1-q)^{\left|\widetilde{Y} \cup Y\right|_{i'>i} - 1} \delta$. The left-hand side of this inequality gives the expected payoff if i defects but the roll call would be won for sure if at least one member of the opposition defected as well. As defection by more than one member of the opposing party may be required for the roll call to be won in the end, this constitutes a weak upper bound on the true expected payoff from defecting. The right hand-side denotes the expected payoff from voting with the party line and having to rely on *all* remaining copartisans to carry the bill to passage (which is a weak lower bound on the true expected payoff from supporting the party). For $q = 0$, the left-hand side of the inequality equals zero and the right-hand side equals $\alpha + \delta > 0$. For $q = 1$, however, the left-hand side is equal to $\delta > 0$, whereas the right-hand side is $\alpha < 0$. Thus, by the Intermediate Value Theorem, there exists a $q_i^3 > 0$ at which equality holds. Furthermore since the left hand side is increasing in q and the right hand side is decreasing in q, the inequality is satisfied for all $q < q_i^3$.

Combining the previous three arguments, let $\bar{q}_i = \min\{q_i^1, q_i^2, q_i^3\}$ and further repeating the process at any node of the game tree where a member of D votes, let $\bar{q}^D = \min_{i \in D} \bar{q}_i$. if q is lower than the resulting $\bar{q}^D$ then senators of party D will defect if and only if $|Y|_{i'>i} + \left|\widetilde{Y}\right|_{i'>i} + 1 \neq \underline{y} - \overline{y}_i$, as in Proposition 1.

To determine the maximal amount of uncertainty under which senators from both parties would still play the same strategies as in Proposition 1, repeat the same steps for members of the Republican party and choose the smallest of $\bar{q}^D$ and $\bar{q}^R$.

Appendix E: Parties as Teams with Delay and Vote Changes

In this appendix, we consider two extensions to our model of "parties as teams." First, we allow for the possibility that a senator may skip her opportunity to vote when she is first called and instead votes in an exogenous order after all of her colleagues have been afforded an initial opportunity to cast their vote. Second, we allow for the possibility that after all senators have voted in order, an agent may change her vote, provided she obtains unanimous consent from all other senators. In what follows, we maintain the assumptions of Propositions 1–3 and modify the model as described.

Delayed Vote. By Senate rules, a senator who is not on the floor when called may still cast her vote after every other senator has been afforded an opportunity to vote. If we extend our model to account for the possibility of strategic delay, we can show that *(i)* our equilibrium of interest continues to hold, and *(ii)* if there is any positive cost to delay, then it continues to be the unique generic equilibrium.

The intuition for the robustness of our result is that skipping the opportunity to vote early is never beneficial and may be costly. If a senator is unconflicted, or if the outcome of the vote will be decided by more than one vote, then the possibility to delay does not affect behavior. Moreover, if a conflicted senator is presented with an opportunity to vote in which deviating would not cause her party to lose the roll call, then waiting may give a conflicted copartisan the opportunity to preempt the senator who chooses to delay. There are no gains from delaying a vote.

More formally, consider the following extension to our model of "parties as teams." When senator i is called to vote, she can either cast her vote on the spot or delay. If she chooses to delay, then she will be allowed to cast a "yea" or "nay" vote after the last senator (S) has been afforded his opportunity to vote.

Denote the number of senators who skip the vote by $K \geq 0$. We assume that the set of senators who choose to delay, vote in a new exogenous, random order, $\widetilde{i} = 1, ..., K$. There is no opportunity to skip the second vote. Let $(\bar{y}^S, \bar{n}^S)$ be the vote count after each senator has been afforded their first opportunity to vote. Note that the game that proceeds after senator S is afforded his opportunity to vote is equivalent to the game previously analyzed with a new yea threshold $\underline{y}^K = \underline{y} - \bar{y}^S$ and a new nay threshold $\underline{n}^K = \underline{n} - \bar{n}^S$. In particular, all the behavior described in Proposition 1 continues to hold. Let $\left|\widetilde{Y}\right|_{\hat{\imath}>i}$ be set of all conditional yea voters that have not yet voted.[2] Now, consider a senator i from the Democratic Party (the argument is the symmetric for Republicans) deciding whether or not to delay her vote. The senator will find herself in one of four cases:

(1) Senator i is unconflicted.

(2) Senator i is a conflicted type and $|Y|_{i'>i} + \left|\widetilde{Y}\right|_{\hat{\imath}>i} + 1 = \underline{y} - \overline{y}_i$.

(3) Senator i is a conflicted type and $|Y|_{i'>i} + \left|\widetilde{Y}\right|_{\hat{\imath}>i} + 1 < \underline{y} - \overline{y}_i$.

(4) Senator i is a conflicted type and $|Y|_{i'>i} + \left|\widetilde{Y}\right|_{\hat{\imath}>i} + 1 > \underline{y} - \overline{y}_i$.

In cases 1-3, the equlibrium behavior of senator i would be the same if she voted in the called order or if she delayed and voted in the "second" round.

In case 4, if $|Y|_{i'>i} < \underline{y} - \overline{y}_i$ so the vote of some conflicted senators is needed following i, then by delaying the initial vote and submitting her choice in the second round, one of three things are possible.

Possibility 1: Enough other conflicted voters have delayed their vote so that there is still an excess of conditional yea-voters choosing after i in the second round. Senator i can, therefore, still vote nay and see the bill pass.

Possibility 2: The bill fails, regardles of i's choice in the second round.

Possibility 3: The conflicted senator finds herself in a situation where her vote is required for passage.

Both possibility 2 and 3 are strictly worse than voting when called, while possibility 1 is equally good.

In case 4, if $|Y|_{i'>i} \geq \underline{y} - \overline{y}_i$ then no conflicted vote is required for the bill to pass. Conflicted voters have, therefore, a strict incentive to always vote nay.

In all cases, there is no incentive to delay the vote. Hence, if we assume that when indifferent senators do not delay their vote, or if there is $\epsilon > 0$ cost of delaying, then no senator will choose to do so. Therefore, $\left|\widetilde{Y}\right|_{\hat{\imath}>i} = \left|\widetilde{Y}\right|_{i'>i}$ and the equilibrium behavior in the extended game will be observationally equivalent to equilibrium of the original game, i.e., the game in the main text.

If there is no cost to delaying a vote, then there exists the possibility of multiple equilibria, in some of which unconflicted senators skip their vote. However, even in these equilibria voting early presents conflicted senators the opportunity to free-ride on those who are first allowed to vote after them.

Vote Changes. By Senate rules, after all votes have been cast, a senator may request to change her recorded vote. If and only if no other senator objects, the vote change is granted. We consider this possibility and show that since consequential vote changes would not be granted with unanimous consent, there is never a gain to casting a vote and changing it at a later time. Thus, if there is any cost to reversing one's choice—for example because senators are wary of being percieved as "flip-floppers"—then the equilibrium we consider in the main text will be the unique generic

[2]Note that if skipping is not permitted then $\left|\widetilde{Y}\right|_{\hat{\imath}>i} = \left|\widetilde{Y}\right|_{i'>i}$

equilibrium of the extended game.

Consider the following modification to our model of "parties as teams." After a complete call of the roll, each senator is offered the opportunity to change her vote (in some exogenously given order). If a senator does not ask for a vote change, nothing occurs. If a senator does request a vote change, then she pays a positive cost $\epsilon > 0$ and each of the remaining senators is simultaneously offered the opportunity to oppose the change at no cost. If at least one agent opposes the change, then the original vote remains in place. After every senator has been offered an opportunity to change their vote the game ends according to the final tally.

The first thing to note is that since, by assumption, there is always at least one senator that prefers the bill to pass and one senator that prefers the bill to fail, unanimous consent will be granted only if it does not alter the outcome of the bill. Additionally, even if a vote change would not reverse the outcome of the roll call, it is still an equilibrium for two or more of the remaining senators to oppose a vote change. Since vote changes do not affect senators' utility when they would not alter the ultimate outcome of the call, these "no change"-equilibria are robust to standard refinements of voting games, which limit players to weakly undominated strategies. Thus, in the vote-change subgame, there is always an equilibrium where no vote change is permitted, and, as consequence, our results in Propositions 1-3 remain unchanged even if $\epsilon = 0$.

However, if we assume that when indifferent senators permit a vote change, we can still show that our results continue to hold in such "permissive" equilibria. First, consider the behavior of an unconflicted Democratic senator who is considering whether to initially vote "nay" when all other players continue to play as in Proposition 1 (the argument is symmetric for unconflicted Republicans). We will show that voting "yea" continues to be a strictly dominant strategy if $\epsilon > 0$ and a weakly dominant strategy if $\epsilon = 0$.

After the initial call of the roll, an unconflicted Democratic senator will find herself in one of three cases:

(1) The bill passes by one or more votes.

(2) The bill fails by two or more votes.

(3) The bill fails by exactly one vote.

In cases 1 and 2, if the senator initially said "nay," she would find it optimal to ask for a vote change, which would be granted in a "permissive" equilibrium. However, doing so will cost her ϵ. In case 3, the senator would not be permitted to change a "nay"-vote.

If the senator had initiall voted "yea," the outcome would be unchanged in cases 1 and 2, but she would economize on ϵ. In case 3, the outcome of the vote may or may not change (depending on the choices of agents $i' > i$). If unchanged, the senator would still economize on ϵ, but if the outcome did change, then the senator would be strictly better off on the instrumental dimension *and* save ϵ. Thus, if $\epsilon = 0$ the senator is weakly better off voting her "true preference" when first called, and if $\epsilon > 0$ she is strictly better off. In either case, it remains an equilibrium for unconflicted Democratic senators to vote "yea."

Given that unconflicted senators continue to initially vote their preference, it follows that it is an equilibrium for conflicted senators to vote according to Proposition 2 when first called. Since the outcome of the roll call cannot change after the initial call of the roll, deviating from the behavior prescribed by Proposition 2 would lower the payoffs of conflicted senators.

In sum, we have shown that the equilibrium in the main text continues to hold if vote changes were permited, and that it is the unique generic equilibrium if there is any cost to changing one's vote.

Appendix F: Simulation Results and Additional Reduced-Form Evidence

F.1. *Simulated Comparative Statics*

In the main text we derive the comparative static of negative vote-order effects under the assumption that the ideological ideal points of members of both parties do not overlap. This restriction rules out situations in which Democrats *and* Republicans are conflicted. Although there is ample empirical evidence to suggest that our assumption is approximately satisfied for most of the period that we study, it is nonetheless restrictive. The example in Figure A.5, for instance, shows that when members of both parties are conflicted, then defection is not always monotonic in rank. It is, therefore, important to show that nonoverlapping preferences are not necessary for negative vote-order effects to obtain *on average*.

In Appendix Figures A.6 and A.7 we present simulation results that suggest that our conclusions about negative vote order effects and winning margins hold more generally. Each panel is based on 10 million roll calls in which 100 senators follow their equilibrium strategies in Proposition 1. For each roll call, the order in which senators vote is randomly determined. The panels on the left depict the average frequency with which agents in a given position deviate from the party line. The panels on the right show the total number of "yea" votes.

Figure A.6 depicts the situation predicted by simple agenda-setter models in political science (e.g., Romer and Rosenthal 1978): a bill that is unpopular with all members of the minority as well as some of the majority. The probability that any given majority party senator is conflicted equals 30%. In Figure A.7 we let members of *both* parties be conflicted with probability 30%. With nonoverlapping preferences, such a situation would never occur. Yet, provided that the margin of majority is neither too large nor too small, on average, deviations from the party line decrease with rank.[3] Moreover, the majority often wins with only a very small margin. Simulation results for other parameter combinations deliver qualitatively similar results and are available upon request.

Figure A.8 shows that when members of *both* parties would like to deviate *and* the margin of majority is small, then the average probability of defection need not be monotonic, even on average.[4]

[3]The rate of decline is lower in the latter figure because defection by minority-party senators lessens the need for members of the majority who come later in the order to support their own party's position.

[4]This observation may be surprising. It is due to the fact that defection by a conflicted member of the minority decreases y_i without lowering $\left|Y \cup \widetilde{Y}\right|_{i'>i}$. Thus, in any particular game, defection by majority-

Although typical seat margins in the U.S. Senate suggest that such scenarios are unlikely to be empirically important (cf. Appendix Figure A.9), it is worth noting that, provided that the majority party has more members than "yea" votes needed for passage, voting *very* early still confers an advantage.

F.2. *Additional Reduced-Form Evidence*

Comparative Statics First, Appendix Figure A. 10 presents some evidence on how often senators find themselves in a situation in which their vote is likely to be dynamically pivotal, i.e., in which $|Y|_{i'>i} + \left|\widetilde{Y}\right|_{i'>i} + 1 \neq \underline{y} - \overline{y}_i$, as in Proposition 1. Since we do not observe senators' preferences we rely on party membership to proxy for $|Y|_{i'>i} + \left|\widetilde{Y}\right|_{i'>i}$. That is, we assume that all senators who follow in the vote order would support the party line if need be. Consistent with the idea that senators early in the order are able to free ride on their colleagues, dynamically pivotal situations are somewhat more frequent towards the end.

According to our analytical results in the main text as well as the simulations in Figures A.6–A.8, rank and defection should be strongly correlated when the minority party is united and when the seat advantage of the majority is sizeable but not large enough for the roll call to be lopsided. By contrast, when the majority party has only a one- or two-seat advantage, then, in equilibrium, almost all of its conflicted members must stick with the party line or else the roll call will be lost. Under these circumstances one would not expect to see a large negative point estimate. In addition, the simulations suggest that for calls on which the minority is split, the correlation between rank and defection should only be modestly negative, if at all. This is because, defection by members of the minority lessens the need of conflicted majority party senators to stick with the party line.

By and large, these predictions are borne out in the results in Appendix Table A.7. Dividing roll calls by the median defection rate among members of the minority party and estimating λ on the sample of calls on which the minority was "split" shows that rank and defection are practically uncorrelated under these circumstances. The same is true when the majority party enjoys a very large or a very small seat advantage, but not in an intermediate range—as predicted by the theory.

Interestingly, $\widehat{\lambda}$ is estimated to be positive (but statistically insignificant) for cases in which the majority party owns only one or two seats more than the minority. Though the simulation results in Figure A.8 are consistent with a positive point estimate for certain parameter combinations, we note that the raw data indicate a small negative slope (cf. Appendix Figure A.11). In any case, we do not observe strongly negative vote-order effects when the majority party has barely more members than the minority, as predicted.

In Figure A.12, we show how $\widehat{\lambda}$ varies over time. Although political scientists disagree about how best to measure partisanship, they generally concur that parties were less important in the 1960s and 1970s, but much more so before and thereafter (see, e.g., Rhode 1991; Snyder and Groseclose

party members need not be monotonic in rank. When the seat advantage of the majority is small enough, this need not even be true on average.

2000). Viewed through the lens of our theory of "parties as teams," one would not expect to see much of a relationship between alphabetical rank and defection during the period in which partisanship was low. After all, when $\delta_p \approx 0$, senators have little incentive to backward induct. Point estimates before 1960 and after 1980, however, should be negative and large.

Although our estimates are imprecise and do not capture some of the more subtle patterns in Snyder and Groseclose's (2000) seminal work on party influence, we do observe markedly smaller vote-order effects during the middle of the twentieth century. Note, a related (but ultimately) distinct explanation for the lack of vote-order effects during this period is that the ideal points of Democratic and Repubilcan senators were more likely to overlap during the decades of low partisanship—as suggested by a large literature in political science (see, e.g., Poole and Rosenthal 1997). Nonetheless, the observed pattern is in line with our similation results, which predict lower vote-order effects when members of both parties are conflicted.

Appendix Table A.8 explores additional sources of heterogeneity in vote-order effects. Consistent with our theory, vote-order effects are larger among members of the majority than their minority-party counterparts. Although the estimated difference is perhaps somewhat smaller than one might have expected, it is statistically highly significant ($p < .01$). There are also small differences with respect to electoral incentives. Comparing the first four and the last two years of their terms, senators are slightly more likely to engage in strategic preemption when the next election is already on the horizon ($p = .107$). The remaining entries explore whether senators' response to changes in their alphabetical rank differs by age, gender, or formal education. Although the difference is barely statistically significant at conventional levels, if at all, the point estimates suggest that senators without a college education may actually be more likely to exploit the opportunity to preempt their colleagues than those without one.[5]

Lastly, in Appendix Table A.9 we differentiate between bills that were discussed in the *CQ Almanac* and those that were not. The *CQ Almanac* is a "nonpartisan reference work that chronicles and analyzes the major bills brought before Congress." Scholars of Congress generally agree that bills mentioned in the almanac are more important than those that are not (see, e.g., Clinton and Lapinski 2006; Howell et al. 2006; Volden and Wiseman 2014).[6] Interestingly, we find no meaningful difference in the order effects on both sets of votes. In particular, the estimated coefficients for both sets of bills are well contained within each other's 95%-confidence intervals.

Detecting Systematic Preemption. Our theory of "parties as teams" predicts that senators systematically preempt each other. If correct, then we would expect that deviations from the party line by senator i are *negatively* correlated with defection by whichever copartisan votes next. To

[5]One admittedly speculative explanation is that the few individuals making it to the Senate without being formally educated might possess higher-than-average innate intelligence, which could be especially conducive to recognizing the advantages conferred by being allowed to vote early.

[6]The data on bill mentions in the *CQ Almanac* come from Policy Agendas Project (2017) and Roberts et al. (2017).

test this prediction one may be tempted to estimate a model of the following kind

$$d_{i,p,r} = \mu_i + \psi d_{i+1,p,r} + \eta_{i,p,r}, \tag{F.1}$$

where $d_{i+1,p,r}$ denotes an indicator for whether the next senator of the same party defect. The problem with (F.1) is that $\widehat{\psi}$ will spuriously positive whenever unobserved roll-call characteristics cause senators from the same party to be skeptical about a bill. We, therefore, rely on the placebo approach outlined in the main text and ask whether $\widehat{\psi}$ is *smaller* than one would expect under the null hypothesis of no preemption.

The answer turns out to be "yes." Appendix Figure A.13 shows the null distribution. Consistent with our model of "parties as teams," our actual estimate of ψ is smaller than 96.1% of placebo coefficients. We, therefore, reject the null of no preemption.

Appendix G: Additional Structural Estimation Results and Robustness Checks

G.1. *Further Results*

For completeness, Appendix Figure A.16 plots the posterior mean of senators' ideal points in our *extended* model against the first dimension of Poole and Rosenthal's (1997) DW-Nominate scores, i.e., it replicates Figure 6 in the main text. As the comparison of both figures demonstrates, allowing for different "types" of legislators has almost no impact on the estimated ideal points.

Since the ternary plots in Figure 9 in the main text can be difficult to read, we show the corresponding marginal posterior densities in Appendix Figure A.17.

In addition to the extended structural model for which we present results in the text, we have also estimated a slightly more parsimonous extension of our baseline theory. This extension only allows for two types of legislators: *(i)* nonstrategic ones, and *(ii)* backward inductors. Results are shown in Appendix Figures A.18–A.20. Interestingly, the posterior distribution of π_2, i.e., the share of senators reasoning backwards, is very similar to that in the model with three types. It appears that allowing for an additional, "intermediate" type reduces primarily the estimated fraction of nonstrategic senators.

Similarly, we find that allowing for two or three types makes very little difference for the posterior of δ. This suggest that it is the nonstrategic rather than the "intermediate" type that causes the rightward shift in the posterior relative to our baseline model.

For each of the three structural models that we have estimated, Appendix Figures A.21 and A.21 depict the marginal posterior distributions of β and γ, respectively. Rather than showing the distribution for each bill-specific parameter, we pool over all roll calls. More detailed results are available from the authors upon request.

G.2. *Robustness Checks*

We have subjected our structural estimation results to an extensive set of sensitivity checks. In particular, we have probed the robustness of our findings with respect to alternative priors on the key parameters of the model as well as a different error structure.

Logit Errors. Appendix Figure A.23 shows the posterior distribution of δ when we replace the normally distributed error terms in our baseline model, i.e., equation (2), with a logit error—leaving all other assumptions in place. Probit and logit are by far the most common parametric restrictions on the error term in structural discrete choice models. To ensure comparability with the results in the main text, we scale the estimates resulting from this robustness check by the standard deviation of the logit distribution. Reassuringly, we obtain qualitatively and quantitalively very similar results.

The same holds true for our extended model with three "types." Appendix Figure A.24 shows the posterior of δ, given a logit-error structure, while Appendix Figure A.25 plots the joint posterior distribution of the type shares.

Alternative Priors. To demonstrate that our finding of strategic behavior and backward reasoning of senators is not due to a particular choice of prior, we show in Appendix Figures A.26 and A.27 how the posterior in our baseline model would change if were to work with an informative prior on δ, i.e., one that is reasonably tightly centered around zero. Such a prior may for instance be motivated with the conventional wisdom of sincere voting in the political science literature on Congress.

The posterior distribution in Figure A.26 results from the prior $N(0,1)$, while that in Figure A.27 is based on $N\left(0,\frac{1}{4}\right)$ instead. Although the mean and mode of the posterior do, of course, move towards zero as the prior becomes more informative, the bulk of the probability mass continues to be concentrated on positive values of δ. It thus appears that the results in the main text are not particularly sensitive to our choice of an uninformative prior on δ. In other words, the information in the data swamp the prior.

Appendix Figures A.28–A.37 probe the sensitivity of the results from our extended model with respect to the prior on (π_1, π_2, π_3). We continue to assume that the prior on these parameters is given by a Dirichlet distribution, but we vary its (hyper)parameters so that it corresponds to very different substantive beliefs.

In particular, Figure A.28 is based on a Dirichlet prior with parameters $(.5, .5, .5)$. This prior resembles the belief that almost all senators are of a single type, without prioritizing one type over the over. Nonetheless, the resulting posterior of (π_1, π_2, π_3) is qualitatively very similar to that in Figure 8, i.e., to the distribution shown in main text. For completeness, Appendix Figure A.29 presents the posterior of δ. Again, changing the Dirichlet prior in this fashion has very little impact on this key parameter.

Appendix Figures A.30 and A.31 are based on a Dirichlet prior with parameters to $(2, 2, 2)$. This configuration corresponds to the belief that there are senator of all types, with roughly equal proportions. Figures A.32–A.33, A.34–A.35, and A.36–A.37 use Dirichlet priors with parameter vectors $(2,1,1)$, $(1,1,2)$, and $(1,2,1)$, respectively. Each corresponds to a situation in which the researcher believes a particular type to be substantially more prevalent than the other two. Reassurinly, the posterior distribution of (π_1, π_2, π_3) as well as that of δ change very little in response to using these alternative priors instead of those in the main text. We, therefore, conlclude that our substantive conclusions are not driven a particular choice of prior.

Appendix H: Detecting Herding

H.1. *Some Monte Carlo Evidence*

This subsection speaks to the ability of our placebo approach to detect "abnormal serial correlation," as predicted by theories of herding and learning about common values. In Appendix F.2 above, we have already shown that our approach detects lower-than-expected serial correlation in defection—consistent with systematic preemption. In what follows, we present evidence from a Monte Carlo study, which we have conducted to directly evaluate the ability of the method to detect herding in small- to medium-sized data sets.

We first describe the data-generating process for our Monte Carlo study and the simulation procedure. We then present results on the frequency of type I and type II errors.

Data-Generating Process. Each Monte Carlo simulation features 100 senators, of which 60 belong to the majority party. Senator i votes "yea" on roll call c if and only if her intrinsic utility from doing so exceeds zero, i.e.,

$$y_{i,c} = \begin{cases} 1 & \text{if } y^*_{i,c} \geq 0 \\ 0 & \text{otherwise} \end{cases},$$

with $y^*_{i,c}$ given by

$$y^*_{i,c} = \alpha_p + \theta \overline{y}_{<i,c} + \varepsilon_{i,c}.$$

As in the main text, $\overline{y}_{<i,c}$ stands for the share of i's copartisans who chose "yea" before it was i's turn to vote. By construction, there are herding effects in senators' choices whenever $\theta \neq 0$. The absolute value of α governs how similar copartisans' preferences are absent any cue-taking behavior. For simplicity, α varies only by party, with $\alpha_D = -\alpha_R$. The noise term $\varepsilon_{i,c} \sim N(0, \sigma)$ denotes senator i's idiosyncratic preference regarding bill c. It is independently distributed across senators and roll calls. Thus, by construction, we have that

$$\Pr\left(y_{i,c} = 1 | \overline{y}_{<i,c}\right) = \Phi\left(\frac{\alpha_p + \theta \overline{y}_{<i,c}}{\sigma}\right),$$

where Φ denotes the standard normal CDF.

Estimation & Simulation. For each simulation run, i.e., each parameter combination (a, θ), we use the data-generating process above to create a "reference data set." We then estimate

$$y_{i,c} = \varkappa + \varphi \overline{y}_{<i,c} + \eta_{i,c}$$

on the reference data in order to obtain the "true" point estimate, i.e., $\widehat{\varphi}$. To determine whether $\widehat{\varphi}$ is larger than one would expect under the null of no bandwagon effects, we apply our placebo approach to the reference data set, relying on 1,000 "randomly reshuffled" vote orders to construct the null distribution, as explained in the main text. Lastly, for each simulation run, we record whether the null hypothesis would have been rejected at the 5%-significance level, i.e, whether $\widehat{\varphi}$ is greater than 95% of placebo coefficients.

Since the purpose of this exercise is to evaluate the ability of surrogate data testing to detect bandwagon effects across a range of alternative settings, we vary the parameters $\alpha = \{0, .1\sigma, .2\sigma, .3\sigma\}$ and $\theta = \{0, .05\sigma, .1\sigma, .15\sigma\}$, as well as the number of available observations per senator, $|c| = \{500, 1{,}000, 10{,}000\}$. For each parameter combination, we conduct 1,000 independent Monte Carlo runs. The resulting rejection rates are reported in Appendix Table A.10.

Results. The evidence in Table A.10 indicates that our placebo approach is more likely to detect herding effects (i.e., less likely to commit a type II error) when the econometrician observes more roll calls per senator and when preferences are less correlated among copartisans (i.e., when α is lower). The intuition for the latter pattern is simple: a higher correlation increases the likelihood that copartisans would "naturally" vote in the same way, which makes it more difficult to pick up any potential effect of one legislator on the other.

Importantly, the Monte Carlo results suggest that our approach exhibits adequate to excellent statistical power, even in relatively small data sets. In particular, as long as the econometrician observes 500 to 1,000 roll calls ($\sim$1 Congress), moderate bandwagon effects (i.e., $\theta = .15\sigma$) are reliably detected. Only for small and very small herding effects (i.e., $\theta = .05\sigma$ or $\theta = .1\sigma$) to be picked up does it take up to 10,000 roll calls.

For comparison, in our data for the 35th–112th Congress, we observe nearly 40,000 roll calls. We, therefore, suspect that our approach would have detected herding among senators with high probability, if it existed.

It is also reassuring to note that theoretical and actual rates of committing a type I error are fairly closely aligned (cf. the column for $\theta = 0$).

H.2. *Do Senators Learn from Observing the Votes of the Leadership?*

Another possibility that might result in herding is that rank-and-file senators look specifically to the votes of the respective party leader and / or whip for cues. In order to test for this form of learning, we ask whether the vote of a senator who ranks *after* the leader (whip) in the voter order is more correlated with the leader's (whip's) choice than that of a colleague who ranks ahead the leader (whip). Since the latter was not able to observe the leader's (whip's) vote before casting her own, we would expect their choices to be less correlated.

The results in Appendix Table A.11, however, show that this is not the case. The correlation between the votes of rank-and-file members and that of the party leadership is essentially independent of whether the former were able to observe the choice of the latter. Given the precision of our estimates, we can even rule out moderately-large differences. We, therefore, find no support for the idea that senators learn from observing the votes of others.

H.3. *Do Senators React to Opponents Who Come Later in the Vote Order?*

Perhaps the best way to rule out explanations that build on backward- rather than forward-looking agents is to show that senators anticipate the *future* choices of at least some of their opponents, and that they change their own behavior in response. To this end, consider the following econometric model:

$$d_{i,p,r,c} = \mu_i + \sum_t \delta_t \, \mathrm{E}\left[d_{i+t,p,r,c}\right] \times 1\left[p \neq p_{i+t}\right] + \sum_t \gamma_t 1\left[p \neq p_{i+t}\right] + \phi \overline{d}_{p,r,c} + \varepsilon_{i,p,r,c}, \tag{H.1}$$

where $\mathrm{E}\left[d_{i+t,p,r,c}\right]$ denotes i's expectation about whether senator $i+t$ (i.e. the one who ranks t positions after i) will defect, p and p_{i+t} respectively denote i's and $i+t$'s party affiliation, and $\overline{d}_{p,r,c}$ is the mean defection rate among other senators of party p (excluding i). All other symbols are as defined as in the main text.

The coefficients of interest are δ_t. They indicate whether senators react to the expected choices of their opponents. In particular, if senator i reasons that the roll call is less likely to be lost when his own defection and that of $i+t$ offset each other, then one should observe that $\widehat{\delta}_t > 0$.

By including $\overline{d}_{p,r,c}$, equation (H.1) controls for unobserved heterogeneity across roll calls, i.e., some calls being intrinsically more controversial among the members of party p than others. Identification is due to situations in which senators i and $i+s$ face some of the same opponents, but each of them is s positions further removed from the former than from the latter. Moreover, senator i may vote in close proximity to multiple opponents, some of whom are more likely to defect from their own party's position than others.

The issue with implementing equation (H.1) empirically is that senators' expectations, i.e. $\mathrm{E}\left[d_{i+t,p,r,c}\right]$, are not observed. It is, therefore, necessary to find an appropriate proxy. One possibility would be to use the actual choices of other senators, i.e. $d_{i+t,p,r,c}$. The obvious concern with such an approach, however, is reverse causality. That is, senator $i+t$ might decide to deviate from the party line *because* i lowered the expected cost of doing so by having defected before him. In order to avoid this problem, the results in the middle panel of Appendix Table A.12 proxy for agents' expectations with the defection probabilities implied by senators' DW-Nominate scores (Poole 2005; Poole and Rosenthal 1997).

DW-Nominate is a scaling technique that predicts Congressmen's choices based on their history of roll call voting. It is widely used by scholars in the field of American Politics and typically thought to recover politicians "ideological ideal points." Although the results in the main text imply that such an interpretation is not literally correct—after all, strategic considerations appear to exert a nontrivial influence—it is noteworthy that DW-Nominate scores are an excellent predictor of actual votes.[7] In particular, they correctly predict about 85% of individual choices. Since senators' past votes are easily observable by their colleagues, it seems plausible that defection probabilities based on their "empirical ideal points" might be a good proxy for others' expectations thereof. Importantly, given that the average senator participates in more than two thousand roll calls, the effect of any particular choice on the estimated forecast, and thus the degree of endogeneity, is likely negligible.

Regardless of whether we proxy for $\mathrm{E}\left[d_{i+t,p,r,c}\right]$ with opponents actual choices or the defection probabilities implied by their DW-Nominate scores, the evidence in Appendix Table A.12 shows that senators do react to the choice of the next two opponents in the vote order. Interestingly, the results in the lower panel of the same table suggest that, conditional "DW-Nominate defection probabilities," opponents' actual choices exert *no* effect. This suggests that senators do, in fact, react to the predictable component of their opponents' behavior, consistent with the idea behind our theory.

In summary, based on the finding of systematic preemption in Appendix F.2 and the result that senators react to exptected choices of opponents who follow them in the order, we conclude that theories based on herding and learning cannot rationalize the entirety of the evidence. The same patterns in the data allow us to also rule out models of vote buying in which the vote buyer is not forward-looking or not reasonably well-informed about the likely choices of senators that have not yet voted.

Appendix I: A Model of Sequential Vote Buying

I.1. *Extension of Dekel et al. (2009)*

This appendix extends the model of Dekel et al. (2009) (henceforth DJW) to the sequential voting structure of the Senate to show that their result that winning coalitions are made up of the lowest cost legislators continues to hold. Thus, in order for vote buying to account for the patterns we observe, senators would need to become intrinsically more supportive of a measure as their alphabetical rank increases.

In what follows we closely adhere to the model of DJW while introducing an explicit sequential legislator-vote-round structure with an exogenous and fixed vote order.

In particular, the model is as follows. Let there be a legislature of size S odd deciding between a new policy x supported by party X and a status quo z supported by party Z. Policy x is

[7] By contrast, the results in this paper provide support for the view of Snyder and Groseclose (2000), according to which ideal point estimates ought to be based on "lopsided" roll calls, as Senators are only free to vote according to their preferences when a single vote is unlikely to be decisive.

implemented if it receives a majority of the votes. Each legislator i is characterized by her utility for voting for each policy, $u_i(x)$ and $u_i(z)$, and the order in which she votes j.[8] We denote the net benefit they get from voting for the policy x by $v_i^{\Delta x} = u_i(x) - u_i(z)$ and, analogously, from voting for z, $v_i^{\Delta y} = u_i(z) - u_i(x)$. We assume that $v_i^{\Delta x}$ is strictly decreasing in i, so that $i = 1$ is the greatest natural supporter of x and $i = S$ is the greatest supporter of z (or least supporter of x.)[9] Let $J(i)$ denote the position in the order in which legislator i votes and let $I(j)$ be the rank in terms of support for the legislator who votes j^{th}. So, legislator i votes in the $J(i)^{th}$ position and the legislator voting in the j^{th} position is the $I(j)^{th}$ most intense supporter of policy x. Given the ordering of legislators by i, let m be the *preference median* legislator and assume that $v_m^{\Delta x} > 0$ so that a majority of legislators prefers voting for x. As mentioned, there are two vote buyers X and Z, where X can generally be thought of the majority party and Z the opposition or an outside group. X would like to pass the bill and values policy x over z at W_X. Z would like the status quo to remain and values it W_Z more than policy x.[10] Let $n_X = |\{i : v_i^{\Delta x} > 0\}|$ be the number of legislators that would support passage without vote buying.

We assume that there is a smallest increment of resources that can be transfered ϵ and that $v_i^{\Delta x}$, W_X, and W_Z are not integer multiples of ϵ in order to avoid indifference. For a positive number k, let $\lceil k \rceil_\epsilon$ (resp. $\lfloor k \rfloor_\epsilon$) be the smallest multiple of ϵ greater than k (resp. largest multiple of ϵ smaller than k). If k is negative, then $\lceil k \rceil_\epsilon = -\lceil |k| \rceil_\epsilon$ and $\lfloor k \rfloor_\epsilon = -\lfloor |k| \rfloor_\epsilon$

Following DJW, we assume that both parties can make binding offers that depend only on a senator's choice. Let $BX(j)$ and $BZ(j)$ denote the final offers made to the legislator voting in position j by parties X and Z respectively. We also assume that voter $v_i^{\Delta x}$ votes for x if and only if $BX(J(i)) + v_i^{\Delta x} > BZ(J(i))$. If party Z were to prevail, the cheapest possible way for them to do so would be the purchase a minimum majority buying the votes of the weakest supporters of x which given the incremental unit ϵ is equal to $\bar{v}^{\Delta X} = \sum_{i=m}^{S} \lceil v_i^{\Delta x} \rceil_\epsilon$. Further, we let $\lfloor W_Z \rfloor_\epsilon > \bar{v}^{\Delta X}$ to avoid the trivial situation where Z can never prevail.

Vote buying proceeds exactly as in DJW but instead of bids being made simultaneously to all legislators, vote buying proceeds round by round. Thus, bidding occurs for the vote of legislator j after all votes from legislators in position 1 to $j - 1$ have been recorded. After bidding, legislator j announces her vote and the game proceeds to the next round with legislator $j + 1$ until all votes have been cast.

The particulars are as follows: In each round, parties alternate in making offers and observe all past offers. So for example, party X begins by making an offer and Z can counter X's offer at which point X can counter and so on. Once an offer has been made by a party, the subsequent offer by that

[8]We assume that legislators only care about their vote for a few reasons. First, we want to adhere to the structure of DJW who make the same assumption. Second, Dal Bo (2007) shows that a vote buyer can eliminate any legislator's pivotality concerns by buying one more vote than needed. Finally, we want to contrast the vote buying setting with our parties as teams model.

[9]This assumption is slightly stronger than DJW who assume that $v_i^{\Delta x}$ is non-increasing. Our stronger assumption rules out indifference between winning coalitions.

[10]Alternatively, we could think of these quantities as their effective budgets.

same party cannot be lower. Again following DJW, we assume there is a small positive cost, $\gamma > 0$, of making an offer in any round. The bidding round ends when two consecutive offer opportunities pass where the vote of the legislator does not change. We say that a party drops out if it does not raise its bids between rounds and the buying round ends. Thus, a party's strategy specifies how much more is offered over the previous offer given the vote count and previous offers. Following DJW, we assume that parties employ strategies that are analogous to their Least Expensive Majority (LEM) strategies in which at each time the party bids (that is the party does not drop out), it purchases the vote of the legislator in the least expensive manner provided the bid does not exceed the net value of winning. Given these strategies, each round amounts to a complete information English auction with bidding costs and jump bidding permitted for the legislator's vote.

If the game does not end, then the payoff for both parties is $-\infty$. Otherwise, party k's payoff if she wins is W_k minus the total cost of votes purchased and the cost of bidding. If party k loses, then her final payoff is a loss equal to the cost of votes purchased and the cost of bidding.

Our solution concept is subgame perfect equilibrium. In order to account for position of play in the game tree, we introduce additional notation. Let $cy(j) = 1$ $(= 0)$ denote a yea (nay) vote in round j. Legislator i votes for the final offer that gives him the highest total utility and so votes yea if $BX(J(i)) + v_i^{\Delta X} > BZ(J(i))$ and votes nay if $BX(J(i)) + v_i^{\Delta X} < BZ(J(i))$. Let $R(j) = \frac{S+1}{2} - \sum_{t=1}^{j-1} cy(t)$ be the number of yea votes needed for passage and $M(j) = S + 1 - j$ be the remaining votes at the beginning of round j.

Denote by $V_X(R, M)$ and $V_Z(R, M)$ the continuation value of the buyers when there are R yes votes needed and M rounds remaining. Finally, let $\widetilde{W}_X(R(j), M(j))$ be the gross benefit to party X of prevailing (winning the round) absent of the cost of buying a vote in round j when there are R yea votes remaining and note that $\widetilde{W}_X(R(j), M(j)) = V_X(R-1, M-1) - V_X(R, M-1)$. Similarly, $\widetilde{W}_Z(R(j), M(j))$ is party Z's gross value of prevailing in round j and is equal to $V_Z(R, M-1) - V_Z(R-1, M-1)$.

Proposition 4 is nearly identical to proposition 1 in DJW (2009) and follows from the same logic of backward induction, limits to rational bids and perfect information. The result extends trivially to our setting as backward induction implies that each voting round amounts to a simple version of their game in which the parties compete for a single vote (rather than seeking a majority), that the values of the parties in the rounds are $\widetilde{W}_Z(R(j), M(j))$ and $\widetilde{W}_X(R(j), M(j))$, and that we can we can work backwards through the rounds. We present the proof, but note that it follows their logic closely and should not be construed as a novel contribution.

Proposition 4: *Given positive costs to making an offer $\gamma > 0$, the vote buying game has a pure strategies equilibrium. In every equilibrium, the same party wins and the losing party never makes any offers for votes.*

Proof: First, note that any legislator-bidding round j with gross winning benefits of winning $\widetilde{W}_X(R(j), M(j))$ and $\widetilde{W}_Z(R(j), M(j))$ is equivalent to a simple one legislator game with terminal payoffs $(\widetilde{W}_X(R(j),M(j)),0)$ if X prevails and $(0,\widetilde{W}_Z(R(j),M(j)))$ if Z prevails. We first show that for any game equivalent to the round j legislator bidding round, there is an equivalent truncated game with bounded bidding. In each bidding period within a round, the offer of at least one party must strictly increase or else the round ends. So after h rounds the minimum offer to the legislator is $h\epsilon$ and for h large enough, the offer to the legislator will be greater than $\max\{\widetilde{W}_X(R(j),M(j)),\widetilde{W}_Z(R(j),M(j))\}$. In order for k to make an offer greater than $\widetilde{W}_k(R(j),M(j))$ it must be that k is certain that the other party will outbid them. However, it cannot be the case that both parties are certain to lose and be outbid, so after some finite number of rounds both parties will quit bidding. This implies there is a finite bound to each party's bids and so we can truncate bids to some finite maximum bid for each party and preserve all equilibria of the game.

Now, the truncated equivalent to each round is a finite, sequential-move game of perfect information and so the existence of equilibrium follows as one can be found by backward induction. Furthermore, there is no indifference anywhere in the game due to our assumption that values are not ϵ-multiples. In particular, this implies that each terminal node in the round has a unique winner. As each party prefers to win in the truncated game, in any subgame working backwards by induction from those nodes that precede terminal nodes, there is a unique winner. Since there is a unique winner, it cannot be the case that the losing party makes an offer as they could deviate to offering nothing an save the cost of bidding and making payments.

Having shown that every voting round with gross benefits of winning $\widetilde{W}_X(R(j),M(j))$ and $\widetilde{W}_Z(R(j),M(j))$ has a unique winner, we can apply the same set of inductive arguments on the game as whole and show that there is a unique winner and the loser never bids as required. *Q.E.D.*

The following remark highlights the first steps of our proof as a useful result to reference later.

Remark 1: *Given positive costs to making an offer $\gamma > 0$, each round j with gross winning benefits of winning $\widetilde{W}_X(R(j),M(j))$ and $\widetilde{W}_Z(R(j),M(j))$ has a unique winner and the losing party never makes any offers.*

Having established that in our environment the same party wins in all equilibria, we now characterize the least costly equilibrium for the winning party and show that it is characterized by a threshold type where all legislators below the threshold (if X wins) or above the threshold (if Z wins) vote with the winning party. These are the same winning coalitions that emerge in the DJW model, where offers are made to all legislators simultaneously. Our key contribution is to show that these coalitions extend to the sequential setting.

Let $WB_k(i)$ denote the bid (possibly 0) paid by k to i when party k wins and i is part of her winning coalition. The winner's objective is to collect the cheapest winning coalition or to

$$\text{(P)} \qquad \min_{A\subseteq I}\sum_{i\in A} WB_k(i) \text{ subject to } |A| \geq \frac{S+1}{2}$$

To show that the solution to this program will be characterized by a threshold, we characterize $WB_k(i)$ for every legislator.

Lemma 1: *For $\gamma > 0$ and sufficiently small, in any legislator round j, party k wins the round if and only if $\lfloor \widetilde{W}_k((R(j), M(j)) \rfloor_\epsilon + v^{\Delta k}_{I(j)} > \lfloor \widetilde{W}_\ell((R(j), M(j)) \rfloor_\epsilon$*

Proof: First note that since $v^{\Delta k}_{I(j)}$ is not an integer multiple of ϵ, so it is never the case that $\lfloor \widetilde{W}_k((R(j), M(j)) \rfloor_\epsilon + v^{\Delta k}_{I(j)} = \lfloor \widetilde{W}_\ell((R(j), M(j)) \rfloor_\epsilon$. To see that k wins if the condition holds, let k bid her maximum willingness to pay $\lfloor \widetilde{W}_k((R(j), M(j)) \rfloor_\epsilon$, which she will be willing to do if $\gamma > 0$ is sufficiently small. The most ℓ would be willing to bid is $\lfloor \widetilde{W}_\ell((R(j), M(j)) \rfloor_\epsilon$ so legislator j votes for k's position if $\lfloor \widetilde{W}_k((R(j), M(j)) \rfloor_\epsilon + v^{k}_{I(j)} > \lfloor \widetilde{W}_\ell((R(j), M(j)) \rfloor_\epsilon + v^{\ell}_{I(j)}$ which is exactly our condition. Since k has a winning strategy, by Remark 1, k wins the round. To see that the condition is necessary, assume that it does not hold. Then by the same reasoning as above it must be the case that ℓ wins the round. *Q.E.D.*

Corollary 1: *If k wins the round and $\lfloor \widetilde{W}_k((R(j), M(j)) \rfloor_\epsilon \leq \lfloor \widetilde{W}_\ell((R(j), M(j)) \rfloor_\epsilon$ then $v^{\Delta k}_{I(j)} > 0$*

Lemma 2: *If k wins the round and ℓ must make a strictly positive offer to win the vote of legislator j, then ℓ will quit the round immediately without making an offer.*

Proof: Our argument is by induction. First note that ℓ quits at any node when she must bid greater than $\lfloor \widetilde{W}_\ell((R(j), M(j)) \rfloor_\epsilon$. Assume that ℓ will quit at any node where the minimum bid needed to win is $q\epsilon$. Now consider a node α where ℓ must spend $(q-1)\epsilon$. If ℓ makes the offer and the game proceeds to node α'. Because we assume that k wins the round, k can respond by adding ϵ to her offer and be in a position to win the vote. The game then proceeds to node α'' where ℓ must now bid $q\epsilon$ to prevail and so by our induction hypothesis, ℓ will quit. Thus, at node α' the continuation equilibrium must have k winning. Thus, ℓ's offer at α leads to a loss of γ, the cost of making an offer and so ℓ would have been better off quiting at node α. *Q.E.D.*

Lemma 3: *If k wins round j then $WB_k(I(j)) = \max\{0, -v^{\Delta k}_{I(j)}\}$*

Proof: We proceed by cases. Case 1: k wins and $\lfloor \widetilde{W}_k((R(j), M(j)) \rfloor_\epsilon \geq \lfloor \widetilde{W}_\ell((R(j), M(j)) \rfloor_\epsilon$ and $\bar{v}^{\Delta k}_{I(j)} > 0$. By lemma 2, ℓ never bids and so k can win for a bid of 0. Case 2: k wins and $\lfloor \widetilde{W}_k((R(j), M(j)) \rfloor_\epsilon \geq \lfloor \widetilde{W}_\ell((R(j), M(j)) \rfloor_\epsilon$ and $v^{\Delta k}_{I(j)} < 0$. If k never bids a positive amount then round ends with ℓ winning, a contradiction of our premise. When k first bids a positive amount it must bid at least $\lceil \bar{v}^{\Delta k}_{I(j)} \rceil_\epsilon$ or else again ℓ wins. Now in order for ℓ to win, it must make an offer equal to $\lceil \bar{v}^{\Delta k}_{I(j)} - \lceil \bar{v}^{\Delta k}_{I(j)} \rceil_\epsilon \rceil_\epsilon = \epsilon > 0$ and so by Lemma 2 drops out. Case 3: k wins and $\lfloor \widetilde{W}_k((R(j), M(j)) \rfloor_\epsilon < \lfloor \widetilde{W}_\ell((R(j), M(j)) \rfloor_\epsilon$ and $v^{\Delta k}_{I(j)} < 0$ is ruled out by Corollary 1. Case 4: k wins and $\lfloor \widetilde{W}_k((R(j), M(j)) \rfloor_\epsilon < \lfloor \widetilde{W}_\ell((R(j), M(j)) \rfloor_\epsilon$ and $v^{\Delta k}_{I(j)} > 0$. By assumption k wins and so by Lemma 2 ℓ will never make an offer, so k wins with a bid of 0. Collecting the cases it is clear that k's winning bid is equal to $\max\{0, -v^{\Delta k}_{I(j)}\}$. *Q.E.D.*

Proposition 5: *If X prevails then the lowest cost winning coalition is defined by an $\bar{n}_X \geq n_X$ where i votes for x if and only if $i \leq \bar{n}_X$. If Z prevails then the lowest cost coalition is defined by $\underline{n}_Z \leq m$ such that i votes for z if and only if $i \geq \underline{n}_Z$*

Proof: The proposition follows immediately from the vote buyer's program (P) and the characterization of $WB_k(I(j))$ in Lemma 3 *Q.E.D.*

Proposition 5 is the chief result of this appendix. It establishes that the coalitions that emerge from vote buying should include those most likely to support the winner's position absent vote buying. Thus, unless changes in senators' position-taking preferences are correlated with changes in the vote order, vote buying cannot generate our main empirical results on order effects.

I.2. *Legislators with Instrumental and Expressive Preferences*

In order to adhere as closely as possible to the DJW model, we adopted their preferences. We now describe how the analysis would change if we instead we assumed the preferences of legislatures had both an instrumental and an expressive component, as in our main model of parties as teams. In particular, we now assume that each legislator i has preference for voting for the bill α_i and an individual specific benefit to the bill passing δ_i. We further assume that for each legislator i neither α_i nor $\alpha_i + \delta_i$ is an integer multiple of ϵ. We also assume that there exists a strict ordering of α_i and $\alpha_i + \delta_i$, and that, under these orderings there is at least a gap of ϵ between each legislator. This ensures that the vote buyers will always have a strict preference for buying an uncontested vote of one legislator over another. Finally, assume that there exists a bound on how much each party can offer each legislator in any round, but that this bound is greater than the total value that the party attaches to winning. That is for each legislator i, there exists a maximum amount q_i^k that party k can offer and that $q_i^k > W_k$. Substantively, this implies that if parties are not limited by feasibility of bidding on legislators—they would never want to bid more than the value they attach to the policy—but there are limits to how much it is possible to bid. The existence of a finite bound on bidding, simplifies our analysis because it ensures that there is a unique equilibrium to the game. The structure of the game otherwise remains the same. We now prove a new proposition that is equivalent to Proposition 4 in our new environment with instrumental and expressive preference and finite bounds on bids.

Proposition 6: *Given positive costs to making an offer $\gamma > 0$, the vote buying game when voters have both expressive and instrumental preferences and there is a finite bound on bids has a unique equilibrium. The unique equilibrium is in pure strategies and the losing party never makes any offers for votes.*

Proof: Because legislators' preferences with or without pivotality concerns are never integer multiples of the bidding unit ϵ they can never be made indifferent between voting actions. Similarly, as the value that parties attach to winning (W_x and W_y) are not integer multiples of ϵ and, for a given vote total, they are never indifferent between whom they bid for, they will never be indifferent between bidding and not bidding. We, therefore, have a sequential move game with perfect information, finite actions and no indifference. Hence, there will be a unique equilibrium, and it will be in pure strategies. Finally, since there is a unique winner, it cannot be the case that the losing party makes an offer as they could deviate to offering nothing and save the cost of bidding, $\gamma > 0$, and potentially the cost of any payments to legislator. *Q.E.D.*

While in the previous section, our proof that there was a unique winner relied on induction, given our additional assumption on bounded strategies, we can instead directly appeal to the structure of the game to generate a unique winner. However, it is useful for later results to prove that each stage game also has a unique winner and that the loser does not make any offers.

LEMMA 4: *Given positive costs to making an offer $\gamma > 0$, each round j with gross winning benefits of winning $\widetilde{W}_X(R(j), M(j))$ and $\widetilde{W}_Z(R(j), M(j))$ has a unique winner and the losing party never makes any offers.*

PROOF: Follows the same argument as the proof to Proposition 6 above. *Q.E.D.*

Having established that there is a unique equilibrium, we now show that the winning party's coalition will consist of the lowest cost collection of legislators, which can be characterized by a threshold type as in DJW. The complication that now arises is that what constitutes the lowest cost coalition depends on whether or not the yea voters are pivotal, the nay voters are pivotal or no one is pivotal. In particular, given that there is a unique equilibrium it will either be the case that at the final vote count (i) the bill passes with exactly $\frac{S+1}{2}$ yea votes, (ii) the bill fails with exactly $\frac{S+1}{2} - 1$ yea votes, or (iii) the bill passes with greater than $\frac{S+1}{2} + 1$ yea votes or fails with less than $\frac{S+1}{2} - 2$ yea votes. We proceed by examining each case in turn.

Case (i): The final number of yea votes is $\frac{S+1}{2}$. In this case, party X wins with a coalition of exactly size $C = \frac{S+1}{2}$ where each member of C votes according to the party's preference. As before, $WB_k(i)$ denotes the bid (possibly 0) paid by k to i when party k wins. Equilibrium will coincide with party X assembling the lowest cost coalition. If not, X would have a profitable deviation and we could not have been at an equilibrium. Thus, X's objective is to assemble the least cost coalition generates a coalition of size $\frac{S+1}{2}$ or:

$$(P') \qquad \min_{A \subseteq I} \sum_{i \in A} WB_X(i) \text{ subject to } |A| = \frac{S+1}{2}$$

Because there are exactly as many yea votes as needed to pass, every yea voter considers themselves pivotal and votes yea if $BX(J(i)) + \alpha_i + \delta_i > 0$ (since party Z loses $BZ(J(i)) = 0$ by Proposition 6) and votes no if $BX(J(i)) + \alpha_i + \delta_i < 0$. As in the main model, let Y be the set of unconflicted voters for whom $\alpha_i > 0$ and $\delta_i > 0$ and let $\widetilde{Y}$ be the set of conflicted voters for whom $\alpha_i < 0$ and $\alpha_i + \delta_i > 0$. If $|Y| + |\widetilde{Y}| \geq \frac{S+1}{2}$ then X is able to win for sure without bidding. Thus, the bidding strategy of X will be uncorrelated with the vote order and cannot generate our empirical regularity. If $|Y| + |\widetilde{Y}| < \frac{S+1}{2}$ then every member of Y and $\widetilde{Y}$ will be part of the winning coalition and X must additionally induce $\frac{S+1}{2} - |Y| + \widetilde{Y}$ legislators from N and $\tilde{N}$ to vote yea. Let $\widetilde{S} = S \backslash (Y \cup \widetilde{Y})$ be set of aligned and conditional no voters and note that a winning coalition must include legislators from this set. Let $\tilde{v}_{\tilde{i}}^{\Delta x} = \alpha_i + \delta_i$ be net utility of voting for policy x as every $i \in \widetilde{S}$ is pivotal and note that for all $i \in \widetilde{S}$, $\tilde{v}_{i}^{\Delta x} < 0$. Let $\widetilde{I}$ be a strict ordering such that $\tilde{v}_{\tilde{i}}^{\Delta x} > \tilde{v}_{\tilde{j}}^{\Delta x}$ if and only if $\tilde{\imath} < \tilde{\jmath}$; so a lower index is associated with greater support. We now characterize the prices for legislators in the set $\widetilde{S}$.

Lemma 5: *For $\gamma > 0$ and sufficiently small, in any legislator round j, party X wins the round if and only if $\lfloor \widetilde{W}_X((R(j), M(j)) \rfloor_\epsilon + \tilde{v}_{\tilde{i}}^{\Delta x} > \lfloor \widetilde{W}_Z((R(j), M(j)) \rfloor_\epsilon$*

Proof: First note that since $\tilde{v}_{\tilde{i}}^{\Delta x}$ is not an integer multiple of ϵ, so it is never the case that $\lfloor \widetilde{W}_X((R(j), M(j)) \rfloor_\epsilon + \tilde{v}_{\tilde{i}}^{\Delta x} = \lfloor \widetilde{W}_Z((R(j), M(j)) \rfloor_\epsilon$. To see that X wins if the condition holds, let X bid her maximum willingness to pay $\lfloor \widetilde{W}_X((R(j), M(j)) \rfloor_\epsilon$, which she will be willing to do if $\gamma > 0$ is sufficiently small. The most Z would be willing to bid is $\lfloor \widetilde{W}_Z((R(j), M(j)) \rfloor_\epsilon$ so legislator j votes for X's position if $\lfloor \widetilde{W}_X((R(j), M(j)) \rfloor_\epsilon + \tilde{v}_{\tilde{i}}^{\Delta x} > \lfloor \widetilde{W}_Z((R(j), M(j)) \rfloor_\epsilon$ which is exactly our condition. Since X has a winning strategy, by Lemma 4, X wins the round. To see that the condition is necessary, assume that it does not hold. Then by the same reasoning as above it must be the case that Z wins the round. *Q.E.D.*

Corollary 2: *If X wins the round and $\lfloor \widetilde{W}_X((R(j), M(j)) \rfloor_\epsilon \leq \lfloor \widetilde{W}_Z((R(j), M(j)) \rfloor_\epsilon$ then $\tilde{v}_{\tilde{i}}^{\Delta x} > 0$*

Lemma 6: *If k wins the round and ℓ must make a strictly positive offer to win the vote of legislator j, then ℓ will quit the round immediately without making an offer.*

Proof: Our argument is by induction. First note that ℓ quits at any node when she must bid greater than $\lfloor \widetilde{W}_\ell((R(j), M(j)) \rfloor_\epsilon$. Assume that ℓ will quit at any node where the minimum bid needed to win is $q\epsilon$. Now consider a node α where ℓ must spend $(q-1)\epsilon$. If ℓ makes the offer and the game proceeds to node α'. Because we assume that k wins the round, k can respond by adding ϵ to her offer and be in a position to win the vote. The game then proceeds to node α'' where ℓ must now bid $q\epsilon$ to prevail and so by our induction hypothesis, ℓ will quit. Thus, at node α' the continuation equilibrium must have k winning. Thus, ℓ's offer at α leads to a loss of γ, the cost of making an offer and so ℓ would have been better off quiting at node α. *Q.E.D.*

Lemma 7: *If X wins round j and $I(j) \in \widetilde{S}$ then $WB_X(I(j)) = -\tilde{v}_{\tilde{i}}^{\Delta x}$, else $WB_X(I(j)) = 0$*

Proof: We proceed by cases. Case 1: $I(j) \in Y \cup \widetilde{Y}$ then since X wins and the loser never bids , so X can win with a bid of 0. Case 2: X wins and $\lfloor \widetilde{W}_X((R(j), M(j)) \rfloor_\epsilon \geq \lfloor \widetilde{W}_Z((R(j), M(j)) \rfloor_\epsilon$ and $\tilde{v}_{\tilde{i}}^{\Delta x} < 0$. If X never bids a positive amount then the round ends with Z winning, a contradiction of our premise. When X bids a positive amount it must bid at least $\lceil \tilde{v}_{\tilde{i}}^{\Delta x} \rceil_\epsilon$ or else again Z wins. Now in order for Z to win, it must make an offer equal to $\lceil \tilde{v}_{\tilde{i}}^{\Delta x} - \lceil \bar{\tilde{v}}_{\tilde{i}}^{\Delta x} \rceil_\epsilon \rceil_\epsilon = \epsilon > 0$ and so by Lemma 6 drops out. Case 3: X wins and $\lfloor \widetilde{W}_X((R(j), M(j))\rfloor_\epsilon < \lfloor \widetilde{W}_Z((R(j), M(j))\rfloor_\epsilon$ and $\tilde{v}_{\tilde{i}}^{\Delta x} < 0$ is ruled out by Corollary 2. *Q.E.D.*

Proposition 7: *If $|Y| + |\widetilde{Y}| \geq \frac{S+1}{2}$ then X prevails without bidding. If $|Y| + |\widetilde{Y}| < \frac{S+1}{2}$ then the lowest-cost winning coalition consists of $Y \cup \widetilde{Y}$ and a subset of $\widetilde{S}$ defined by a threshold $\tilde{n}_X$ in the order $\widetilde{I}$, where $\tilde{i} \in \widetilde{S}$ votes for x if and only if $\tilde{i} \leq \tilde{n}_X$.*

Proof: The proposition follows immediately from the vote buyer's program (P') and the characterization of $WB_k(I(j))$ in Lemma 7 *Q.E.D.*

Case (ii): The final number of yea votes is $\frac{S+1}{2} - 1$. Case (ii) proceeds exactly as case (i), but from the perspective of party Z and the bill fails by exactly 1 vote.

Case (iii): The final number of yea votes is strictly less than $\frac{S+1}{2} - 1$ or strictly more than $\frac{S+1}{2}$. In case this case, the winner will assemble a coalition of size C where $C > \frac{N+1}{2}$ and each member of C votes according to the party's preference. Conditional on the size of the coalition, the winner's objective is to assemble the coalition in the lest costly way or to:

$$(P'') \qquad \min_{A \subseteq I} \sum_{i \in A} WB_k(i) \text{ subject to } |A| = C$$

As every legislator anticipates that their vote will not be pivotal and so votes yea if $BX(J(i)) + \alpha_i > BZ(J(i))$ and votes nay if $BX(J(i)) + \alpha_i < BZ(J(i))$. This is exactly the voter behavior examined with only expressive preferences and so the analysis proceed exactly in the subsection above. In particular, Proposition 5 continues to hold and the winning coalition that emerges includes those whose expressive preferences are most in line with the winner's position. Again, unless these position taking preferences are correlated with order, it cannot be the case that they generate our main empirical result on order effects.

To sum up, we have established that the coalition that emerges from vote buying should include those most likely to support the winner's position absent vote buying *conditional on the size of the winning coalition*. Note that this does not imply that there are never order effects in the presence of vote buying, but instead that the vote buying, when it occurs, will target the cheapest legislators and will not condition on order per se. In fact, there may be order effects in the presence of vote buyers that are induced by the "party as teams" preferences. For example, consider a situation where $|Y| + |\widetilde{Y}|$ constitutes a strict majority and both buyers have no budget or the preference advantaged buyer has a sufficiently large budget so that the opponent would never bid. In such a case, no votes are purchased, but early conditional-yea voters free ride on later conditional-yea voters as in our main model.

Appendix J: Data Appendix

This appendix provides a description of all data used in the paper, as well as precise definitions together with the sources of all variables.

J.1. *Universe of Roll Calls, 1857–2013*

Data on all roll call votes in the United States Senate were kindly provided by Keith Poole.[11] They are based on careful codings of the Congressional Record. The data contain senators' names, home states, party affiliation, and final votes. They neither indicate the actual order in which votes were submitted, nor do they contain any information on whether a given senator changed or withdrew his initial vote. Unfortunately, this information is not part of the Congressional Record. The analysis in this paper restricts attention to the votes of Democratic and Republican senators since the emergence of the two-party system, i.e. from the 35th to the 112th Congress (1857–2013). The following variables are being used:

Party Line is defined for each roll call vote that a senator submits. It equals the vote choice of the simple majority of *other* senators from the same party (not including the senator for whose vote it is calculated).

Deviate is an indicator variable equal to one if a senator's vote differs from the party line, as defined above. It is zero otherwise and undefined for senators who did not participate in a given roll call.

Alphabetical Rank is defined as $\frac{s_i-1}{S-1}$, where S denotes the number of senators who particiapte in a given roll call, and s_i is senator i's raw alphabetical rank among participants. s_i is constructed based on senators' last names, as contained in the raw data. Roughly speaking, the variable Rank corresponds to senators' alphebatical percentile ranking among their colleagues (divided by 100).

"Close" vs "Lopsided" Roll Calls are categorized as in Snyder and Groseclose (2000). That is, a roll call is said to be "lopsided" whenever more than 65% or less than 35% of Senators voted "yea." For votes that require a supermajority, e.g., treaties and cloture votes, the corresponding cutoffs are 51.7% and 81.7% (i.e. 66.7% $\pm$ 15%). Data on supermajority requirements come from Snyder and Groseclose (2000) and have been manually extended through the 112th Congress.

Divisive is an indicator variable equal to one if the majority of one party votes in the opposite direction of the majority of the other party. It is zero otherwise.

"Split" Minority is an indicator variable equal to one if fewer than the median percentage of minority-party senators deviate from the party line on a particular roll call. It is zero otherwise.

Seat Advantage is defined as the difference in the number of senators between the majority and minority parties who participate in a given roll call.

[11] They are publicly accessible at http://www.voteview.com.

DW-Nominate Scores were kindly provided by Keith Poole. For a description of the DW-Nominate estimation procedure, see Poole (2005).

Experience is defined as the total number of roll call votes that a senator had ever submitted before a particualr roll call was conducted.

J.2. *Roll Calls in the 112th Congress*

As described in the main text, a team of research assistants collected information above and beyond what is contained in the Congressional Record for the 112th Congress. To this end, the C-SPAN network was asked to furnish video recordings of all 486 roll calls held during this period, i.e., from January 3, 2011 to January 3, 2013. Due to technically difficulties C-SPAN was only able to produce videos for 484 out of 486 calls. We retrieved the remaining two recordings from the National Archives in Washington, D.C.

A research assistant watched each recording, pausing frequently to transcribe every senator's choice, whether it was submitted during the alphabetical call of the roll, as well as any subsequent changes. Recording senators' votes during the initial period is facilitated by the fact that, directly after the call of the roll, the clerk summarizes the vote by respectively identifying those who voted "yea" and "nay." The research assistant also recorded the exact order in which votes were submitted (or changed) after the clerk had stopped calling the roll, but before the roll call had been officially closed. Doing so is relatively straightforward, as the clerk generally announces the name of the senator who just submitted his vote, followed by his choice. A second research assistant then rewatched each video in order to check the transcription for errors and to ensure that senators' final choices in the data match the official Congressional Record.

J.3. *Senator Characteristics*

Raw data on senators' characteristics come from the Database of Congressional Historical Statistics and were obtained through the Inter-University Consortium for Political and Social Research (ICPSR 3371). The data were manually checked for errors and extended to cover all senators who served before the end of the 112th Congress. Whenever the information in the Biographical Directory of the U.S. Congress differed from the raw data, the latter was changed to conform to the former.[12] Throughout the analysis, the following variables are used:

Age is defined as a senator's age (in years) at the beginning of a particular Congress.

Gender is defined as the senator's biological sex.

College Educated is an indicator variable equal to one if the Biographical Directory of the U.S. Congress indicates that the senator graduated from college. It is zero otherwise.

End of Term refers to the last two years of a senator's regular 6-year term.

[12] The Biographical Directory of the U.S. Congress is available at http://bioguide.congress.gov/.

References

Clinton, J., S. Jackman, and D. Rivers (2004). "The Statistical Analysis of Roll Call Data." *American Political Science Review*, 98(2): 355–370.

———, and J. S. Lapinski (2006). "Measuring Legislative Accomplishment, 1877–1994." *American Journal of Political Science,* 50(1): 232–249.

Dal Bo, E. (2007). "Bribing Voters." *American Journal of Political Science*, 51(4): 789–803.

Dekel, E., M. O. Jackson, and A. Wolinsky (2008). "Vote Buying: General Elections." *Journal of Political Economy*, 116(2): 351–380.

———, ———, and ——— (2009). "Vote Buying: Legislatures and Lobbying." *Quarterly Journal of Political Science*, 4(2): 103–128.

Frisch, R., and F. V. Waugh (1933). "Partial Time Regressions as Compared with Individual Trends." *Econometrica*, 1(4): 387–401.

Howell, W., S. Adler, C. Cameron, and C. Riemann (2000). "Divided Government and the Legislative Productivity of Congress, 1945–94." *Legislative Studies Quarterly*, 25(2): 285–312.

Policy Agendas Project at the University of Texas at Austin (2017). Congressional Quarterly Almanac Data. retrieved from <www.comparativeagendas.net>

Poole, K. T. (2000). "Nonparametric Unfolding of Binary Choice Data." *Political Analysis*, 8(3): 211–237.

——— (2005). *Spatial Models of Parliamentary Voting*. Cambridge, UK: Cambridge University Press.

———, and H. Rosenthal (1997). *Congress: A Political-Economic History of Roll Call Voting.* New York: Oxford University Press.

Roberts, J., D. Rohde, and M. Crespin (2017). Political Institutions and Public Choice Senate Roll-Call Database. retried from <http://cacexplore.org/pipcvotes/>

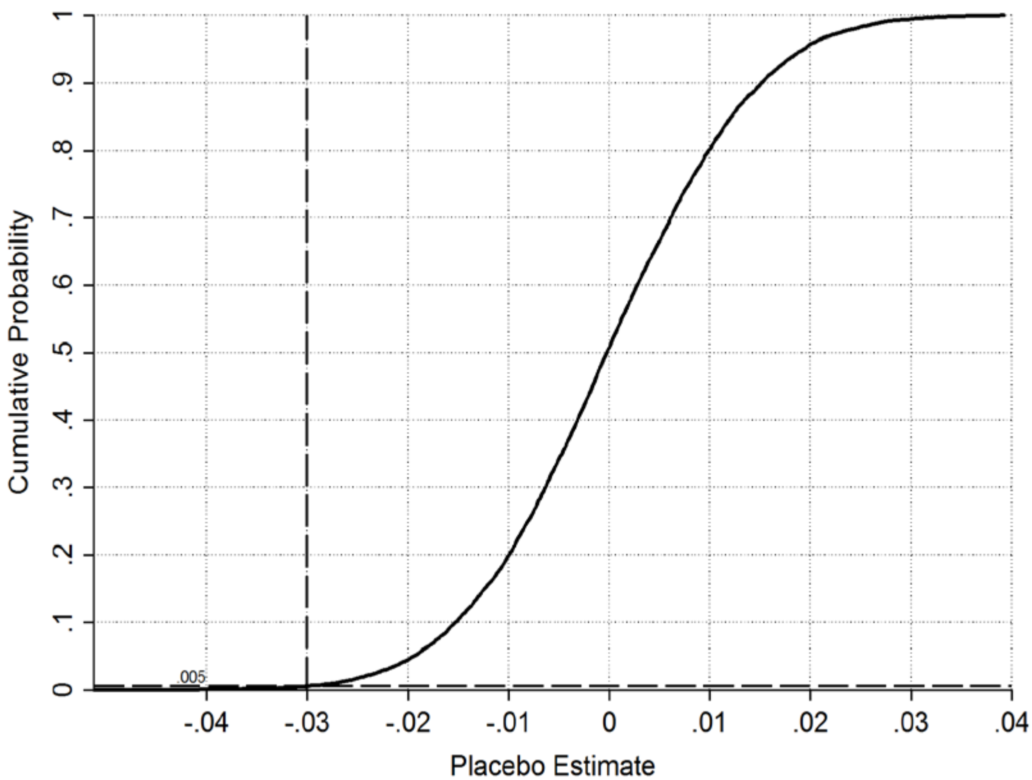
Rhode, D. W. (1991). *Parties and Leaders in the Postreform House.* Chicago: University of Chicago Press.

Romer, T., and H. Rosenthal (1978). "Political Resource Allocation, Controlled Agendas, and the Status Quo." *Public Choice*, 33(4): 27–43.

Snyder, J. M., and T. Groseclose (2000). "Estimating Party Influence in Congressional Roll-Call Voting." *American Journal of Political Science*, 44(2): 193–211.

Volden, C., and A. E. Wiseman (2014). *Legislative Effectiveness in the United States Congress: The Lawmakers*. Cambridge, UK: Cambridge University Press.

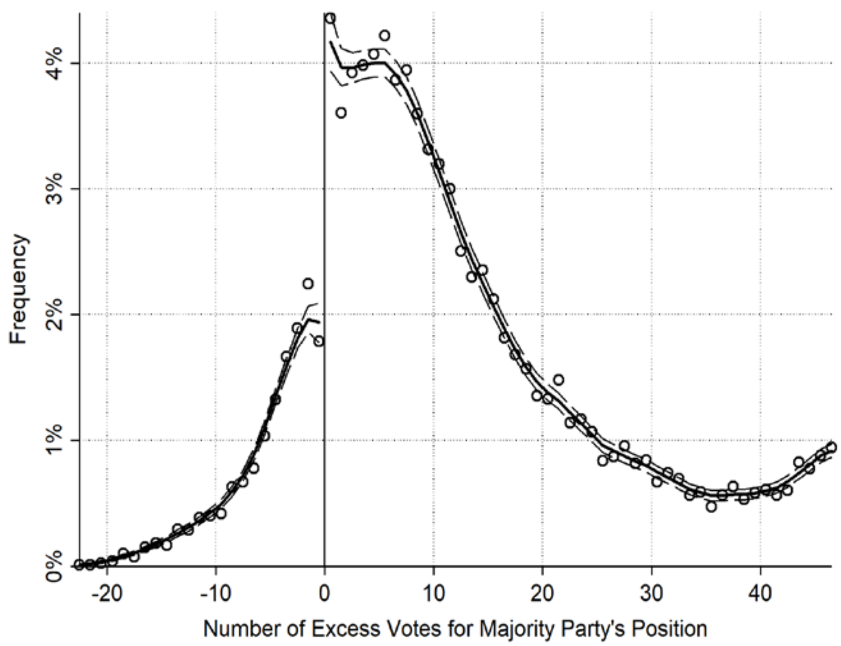
Figure A.1: Empirical CDF of Placebo Estimates for λ, Controlling for Copartisans' Choices



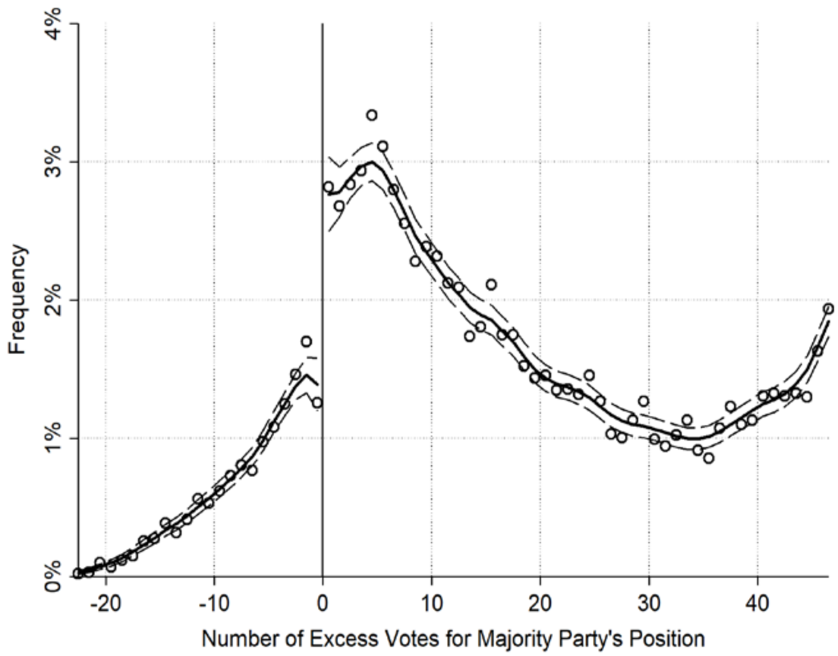
Notes: Figure shows the empirical cumulative distribution function for placebo estimates of λ in equation (1), controlling for the mean deviation rate among copartisans. Estimates are based on 10,000 randomly generated placebo orderings. The vertical line indicates the point estimate in the original data.

Figure A.2: Winning Margings, by Party Affiliation of Vice President

A. Vice President Belongs to Majority Party

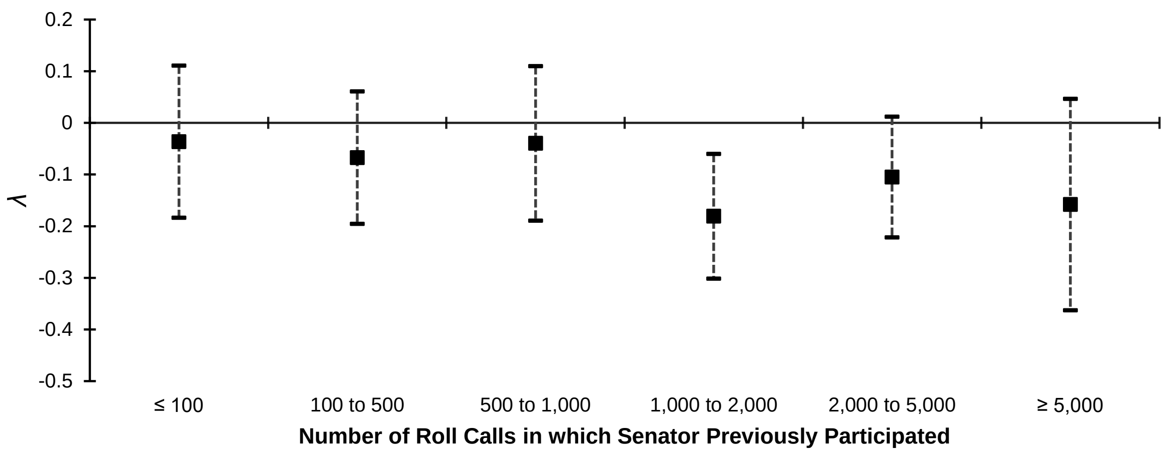


B. Vice President Doesn't Belong to Majority Party



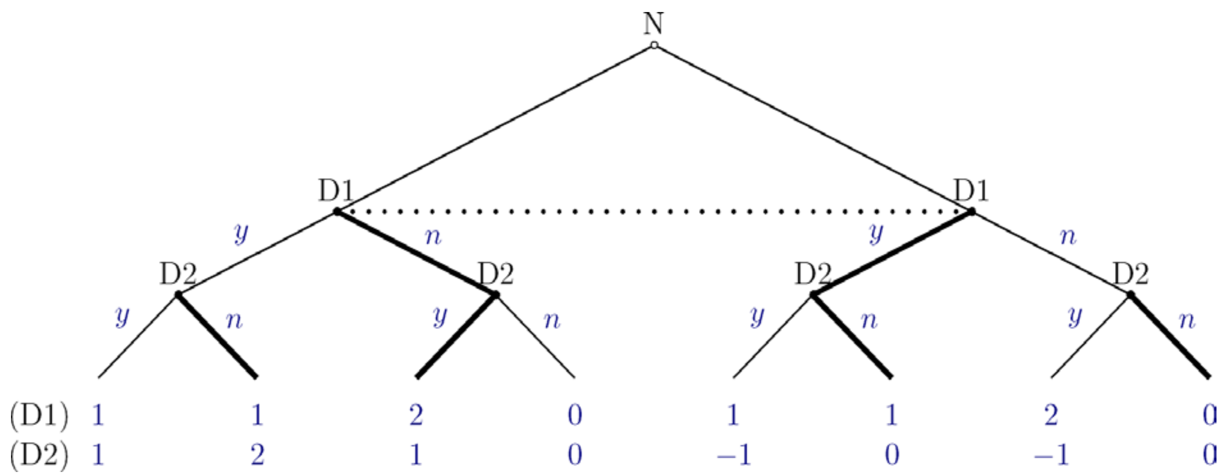
Notes: Figure shows the frequency of excess votes (relative to the threshold required for passage) in favor of the position held by the Senate's majority party, as well as the estimated density function and the associated 95%-confidence intervals. The upper panel restricts attention to situations in which the Vice President belongs to the majority party, whereas the lower one presents results for the opposite case. The underlying data come from roll calls in the U.S. Senate that required a simple majority and were held during the 35th–112th Congresses. Density estimates are based on local linear regressions with a bandwidth of 4, applied separately on each side of the cutoff. See McCrary (2008) for details on the estimation procedure.

Figure A.3: Estimated Order Effects by Senators' Prior Experience, Controlling for Time Served in the Senate



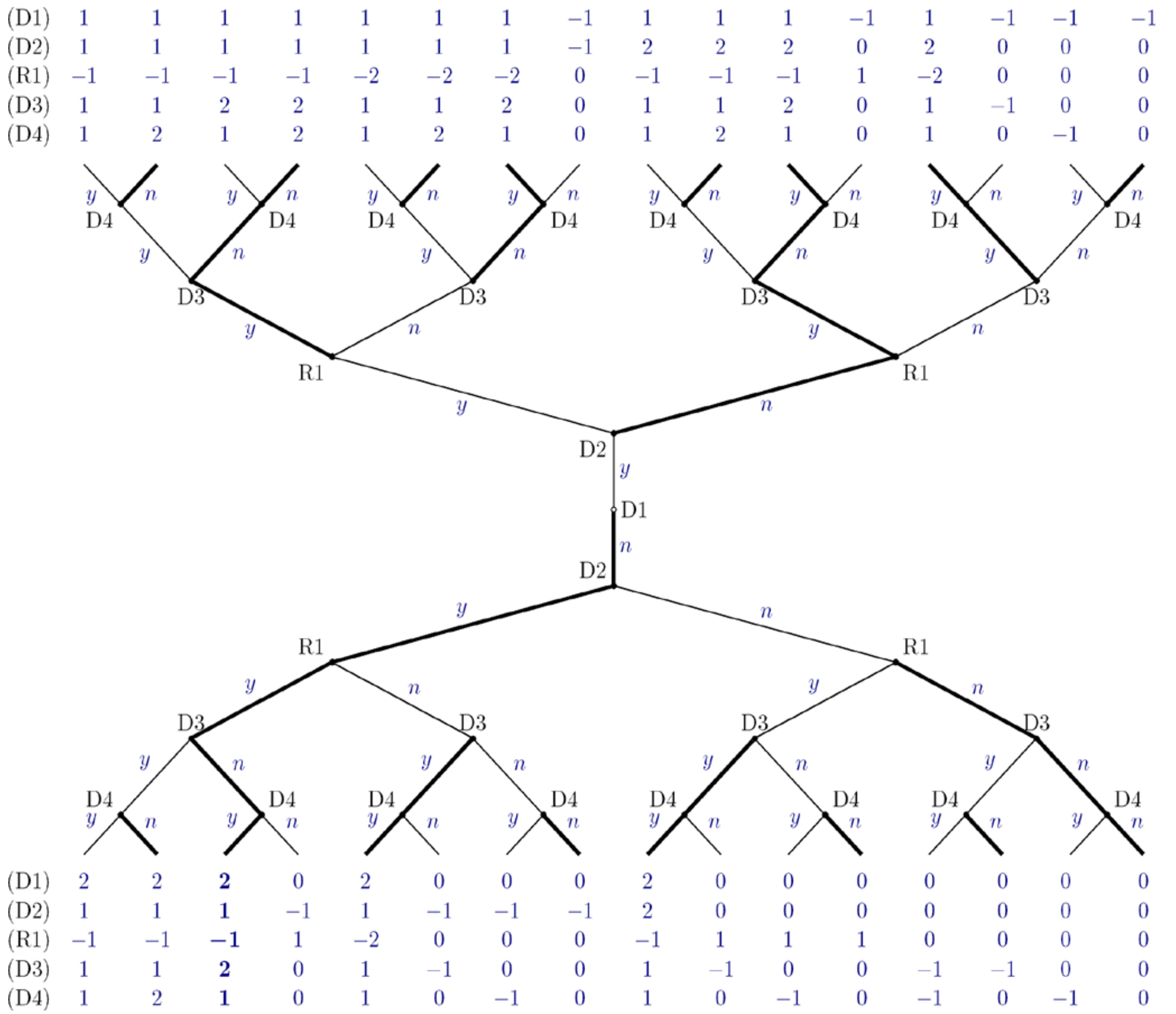
Notes: Figure shows point estimates and the associated 95%-confidence intervals for λ in equation (1), controlling for the number of Congresses in which a senator had previously served. Estimates are based on senators' roll call–specific rank. Confidence intervals account for heteroskedasticity and clustering of the residuals at the Congress level.

Figure A.4: Example of Sequential Voting Game with Uncertainty



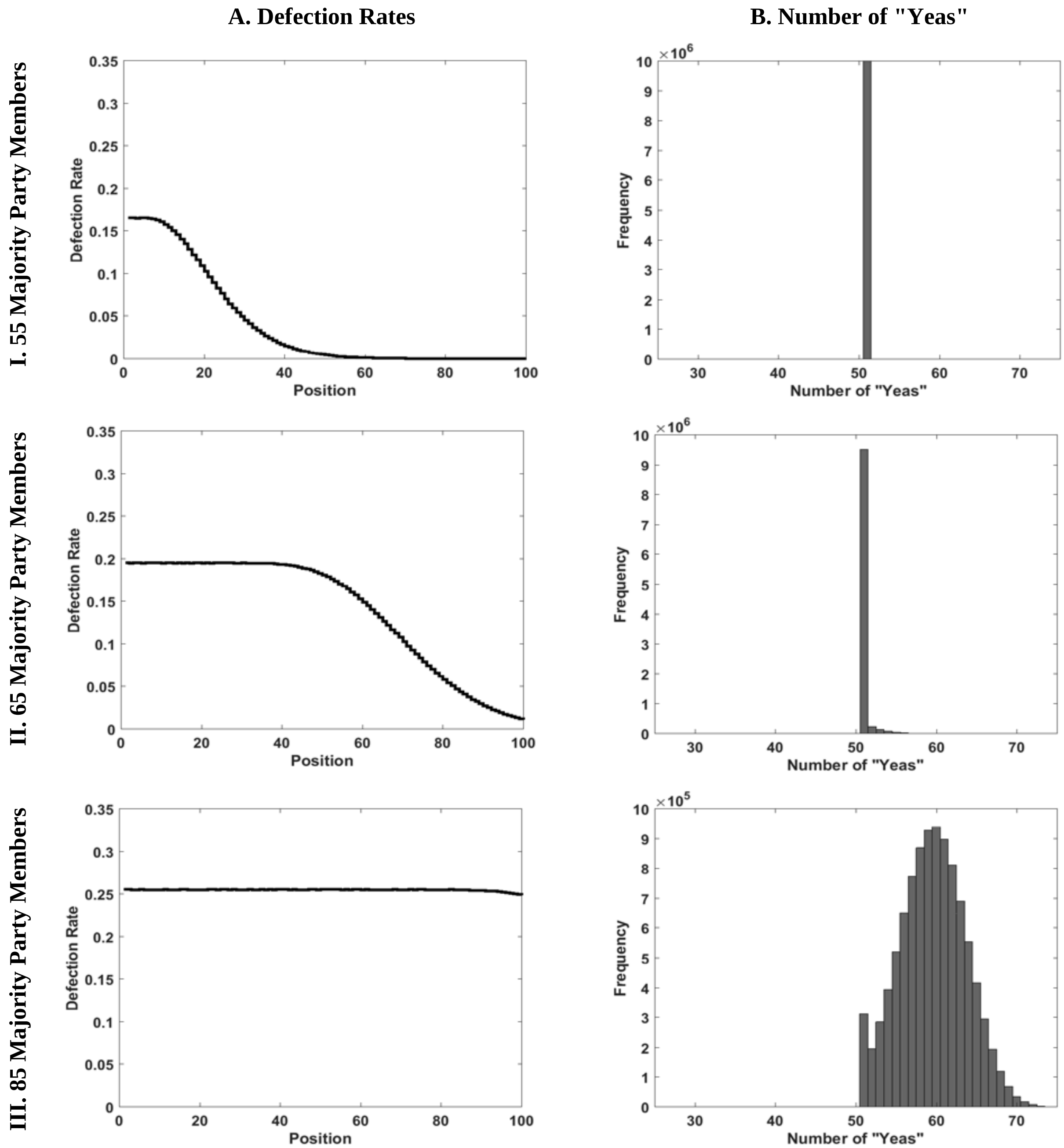
Notes: Figure shows an example of our sequential voting game with uncertainty about preferences. For simplicity, there is one party and two players. One "yea" vote is required for passage. Holding the outcome of the roll call fixed, both D1 and D2 would like to go on the record opposing the bill. D1, however, would always rather vote "yea" than let the measure fail for sure. By contrast, D2's payoffs are determined by nature and unknown to D1. In one state of nature, D2 will always vote "nay" regardless of whether this causes the bill to fail. In the other state, D2 is willing to support the party line if need be. The thick lines indicate each player's optimal action at a particular node in the game tree.

Figure A.5: Example with Nonmonotonic Defection



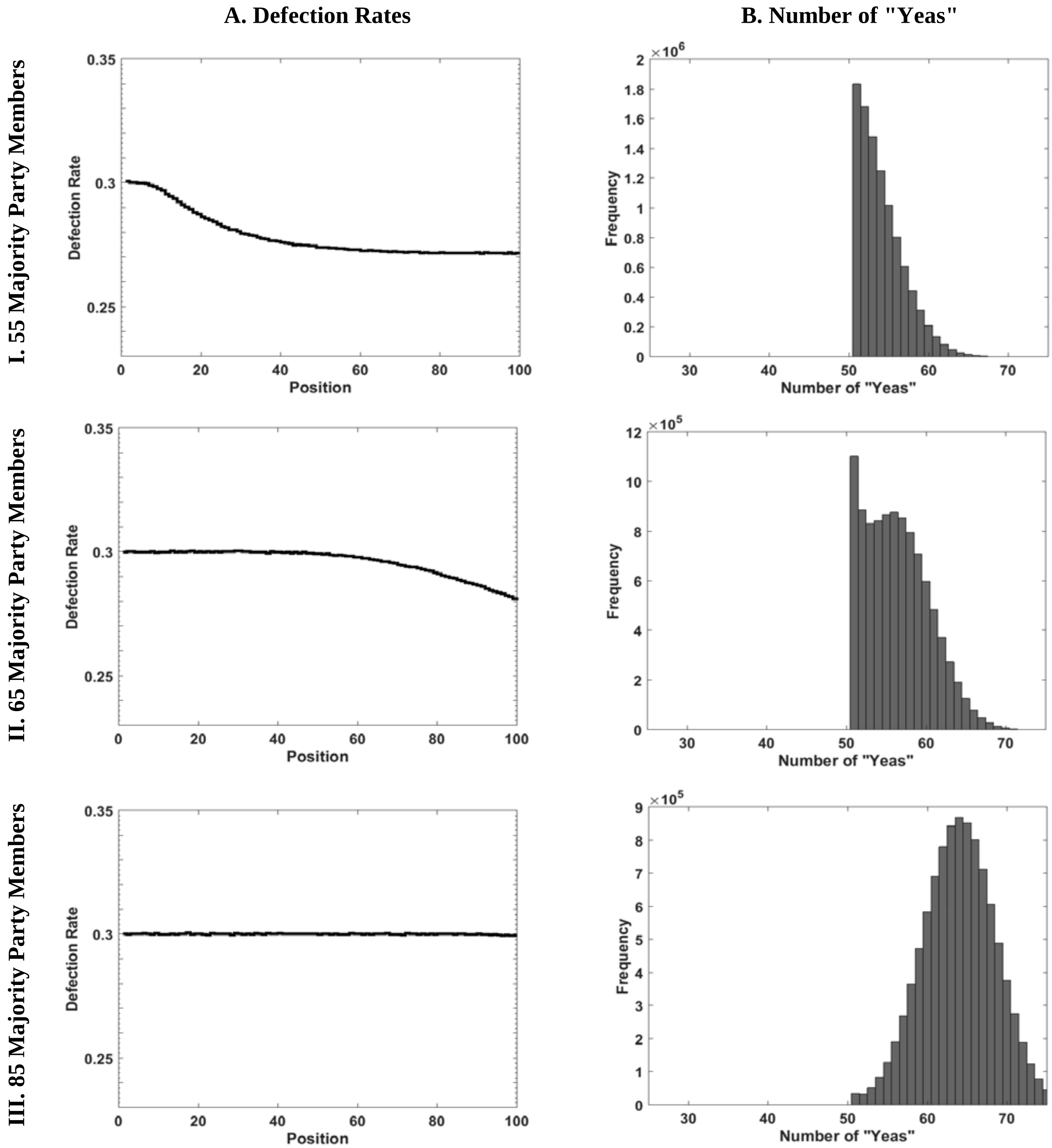
Notes: Figure shows an example of our sequential voting game with two parties and five Senators. All senators are conflicted. Specifically, Democrats receive payoff α = -1 if they vote "yea" and δ = 2 if the bill ends up being approved, whereas Republicans receive α = 1 and δ = -2. Three "yea" votes are needed for passage. The thick lines indicate each player's optimal action at a particular node in the game tree.

Figure A.6: Simulation Results for Situations when Only Majority Party Senators are Conflicted



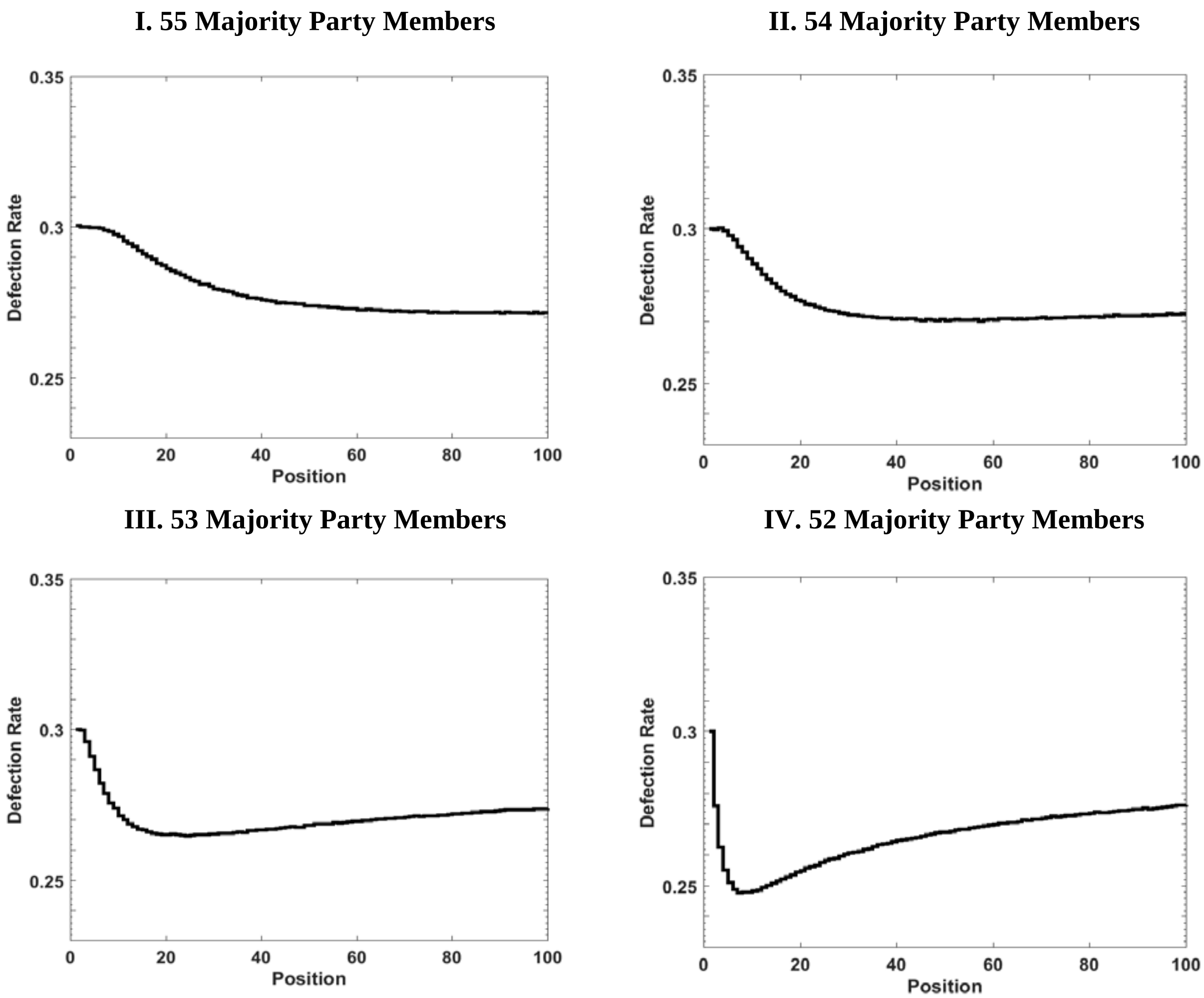
Notes: Panels on the left depict the average rate of defection as a function of when a senator gets to cast her vote. Panels on the right show a histogram of the total number of "yea" votes on a simulated roll call. All results are based on 10 million simulations in which 100 senators play the equilibrium strategies prescribed by Proposition 1, assuming that the bill passes when a simple majority assents. For each simulation, the order in which agents vote is randomly determined. The probability that a particular member of the majorty party is "conflicted" equals 30%. The preferences of all other senators are aligned with their parties' opposing stances. In the upper two panels, 55 senators belong to the majority. In the middle and lower two panels, the number of majority party members is 65 and 85, respectively.

Figure A.7: Simulation Results for Situations when Senators from Both Parties are Conflicted



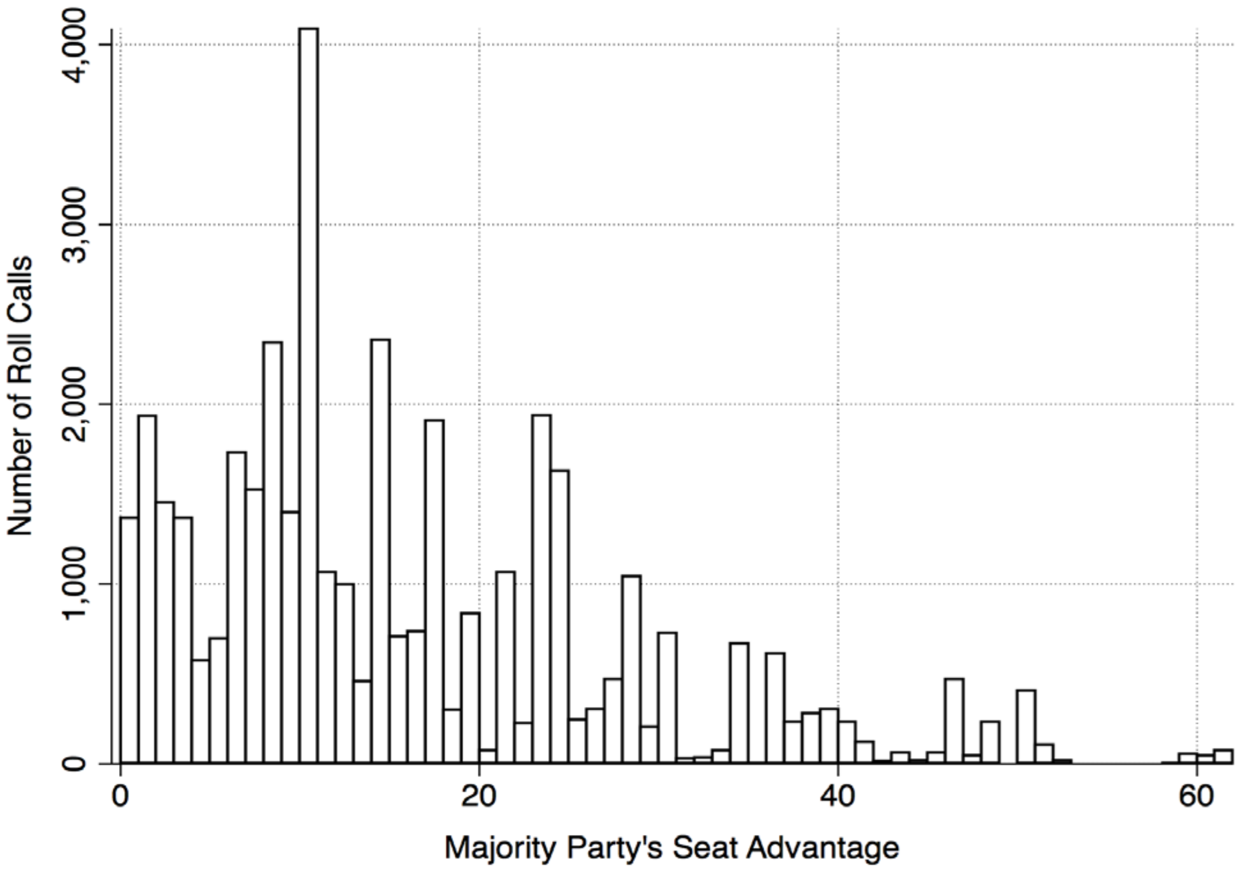
Notes: Panels on the left depict the average rate of defection as a function of when a senator gets to cast her vote. Panels on the right show a histogram of the total number of "yea" votes on a simulated roll call. All results are based on 10 million simulations in which 100 senators play the equilibrium strategies prescribed by Proposition 1, assuming that the bill passes when a simple majority assents. For each simulation, the order in which agents vote is randomly determined. The probability that a member of *either* party is "conflicted" equals 30%. The preferences of all other senators are aligned with their parties' opposing stances. In the upper two panels, 55 senators belong to the majority. In the middle and lower two panels, the number of majority party members is 65 and 85, respectively.

Figure A.8: Simulation Results for Situations in which Both Parties are Split and Almost Equal in Size



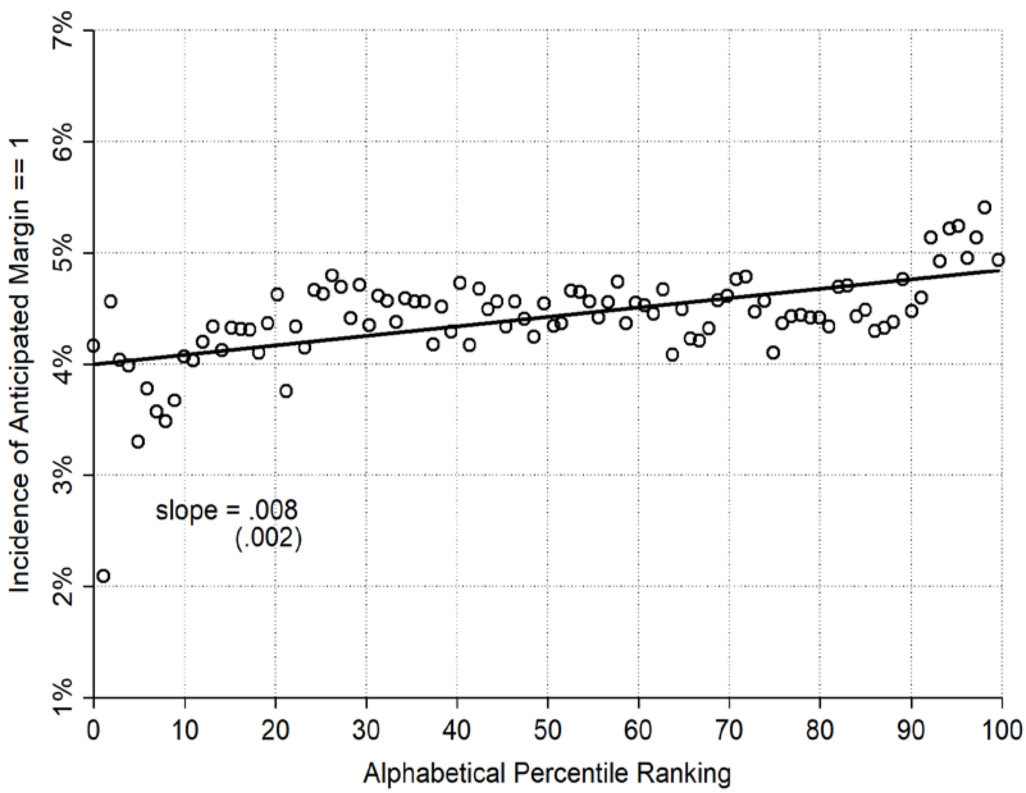
Notes: Figure depicts the expected average rate of defection as a function of when a senator gets to cast her vote. The results in each panel are based on 10 million simulated roll calls in which 100 senators follow the equilibrium strategies in Proposition 1. For each roll call, the order in which agents vote is randomly determined. The majority party's preferred outcome obtains whenever a simple majority votes "yea." In all simulations, there is a 30% probability that any given senator is "conflicted." The preferences of all other agents are aligned with their parties' opposing stances. Panels I–IV lower the number of senators who belong to the majority party from 55 to 52.

Figure A.9: Distribution of the Majority Party's Roll Call-Specific Seat Advantage, U.S. Senate 1857–2013



Notes: Figure shows the distribution of the majority party's seat advantage during the 35th–112th Congresses, restricting attention to senators who participated in a given roll call. The majority party is defined by roll call, i.e., the party with the most senators participating.

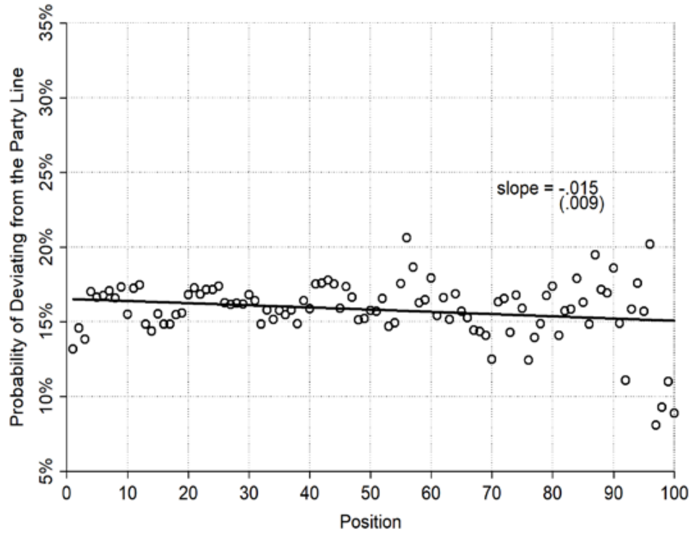
Figure A.10: Frequency of Dynamically Pivotal Situations



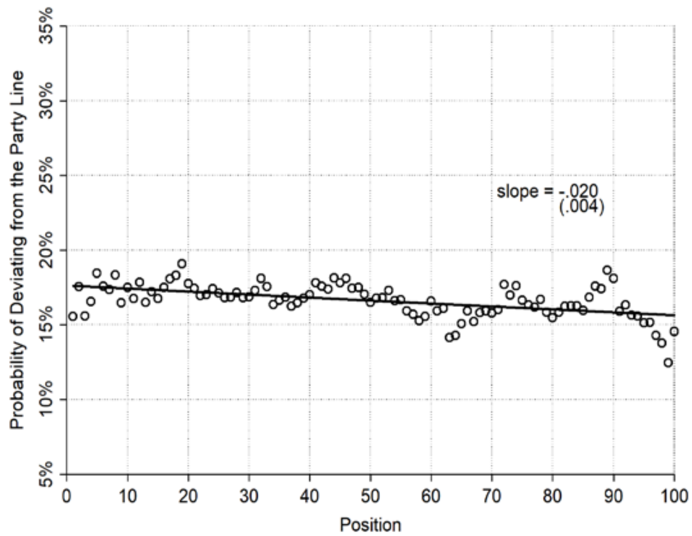
Notes: Figure shows the frequency with $|Y|_{i'>i} + |\tilde{Y}|_{i'>I} + 1 = \underline{y} - \overline{y}_i$ (cf. Proposition 1) as a function of senators' ranking in the vote order. Since we do not observe senators' preferences we rely on party membership to proxy for $|Y|_{i'>i} + |\tilde{Y}|_{i'>i}$. That is, we assume that all senators who follow in the vote order would support the party line if need be.

Figure A.11: Deviations from the Party Line in the Raw Data, by Majority Party's Seat Advanatge

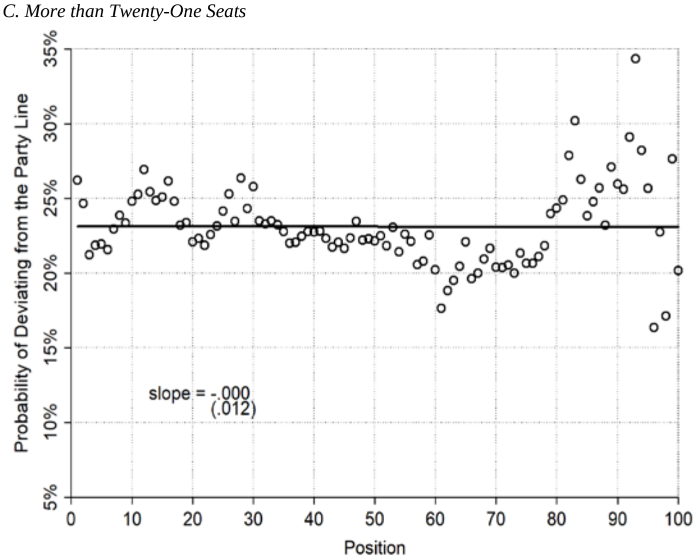
A. Less than Two Seats



B. Three to Twenty Seats

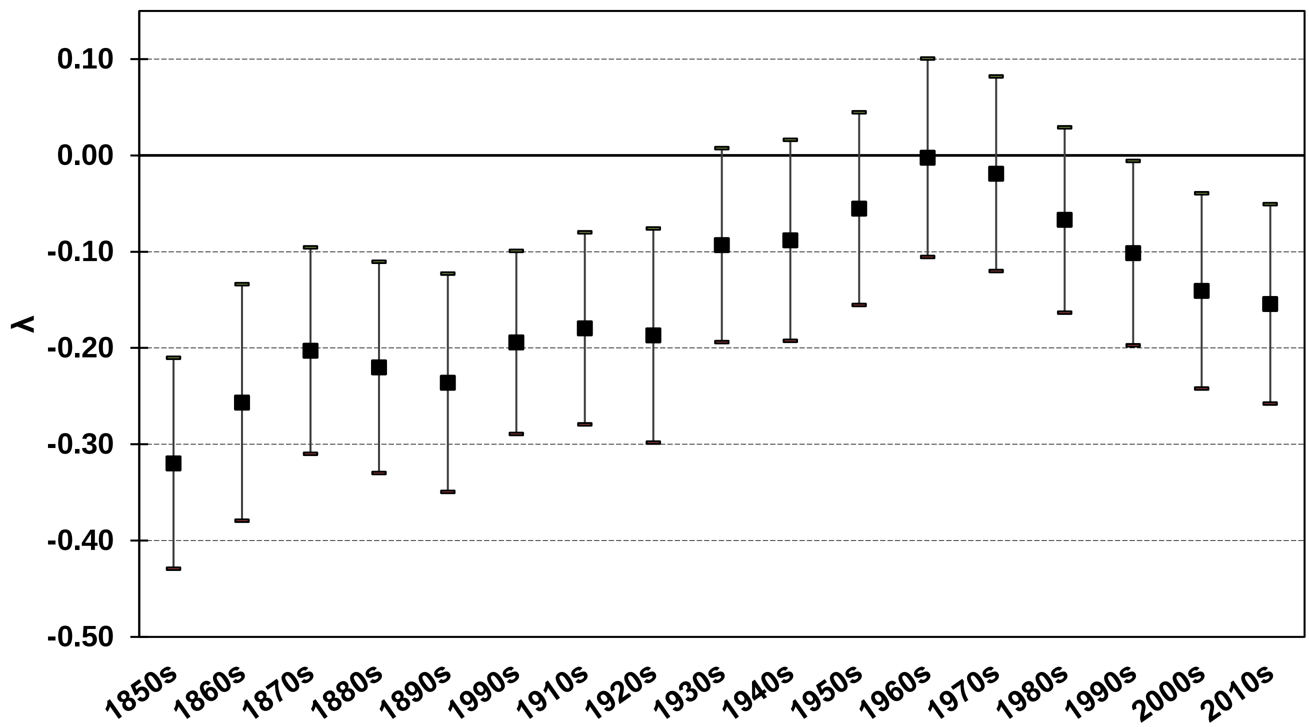


C. More than Twenty-One Seats



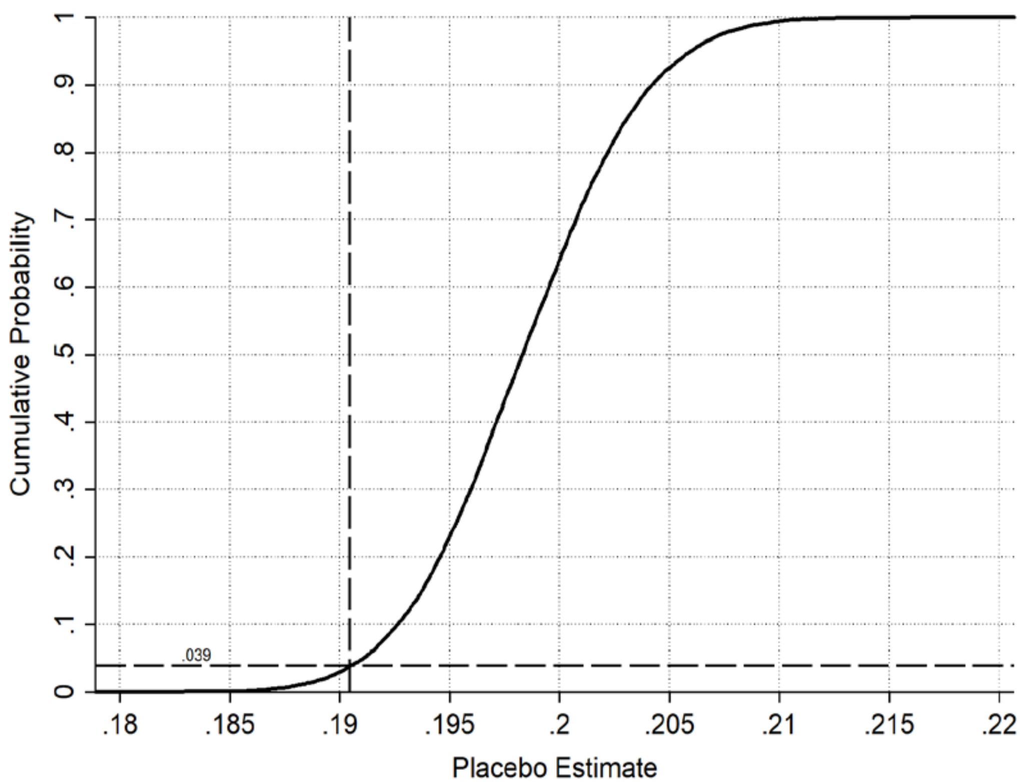
Notes: Figure shows the average frequency with which senators deviate from the majority of their copartisans, depending on their position in the vote order. The upper panel restricts attention to roll calls in which the majority party had a seat advantage of at most 2 while the remaining two panels focus on roll calls with seat advantages of 3 to 20 and at least 21, respectively.

Figure A.12: Estimated Order Effects, by Decade



Notes: Figure shows point estimates and the associated 95%-confidence intervals for λ, estimated decade by decade. Estimates are based on equation (1) and senators' roll call–specific rank. Confidence intervals account for heteroskedasticity and clustering of the residuals at the Congress level.

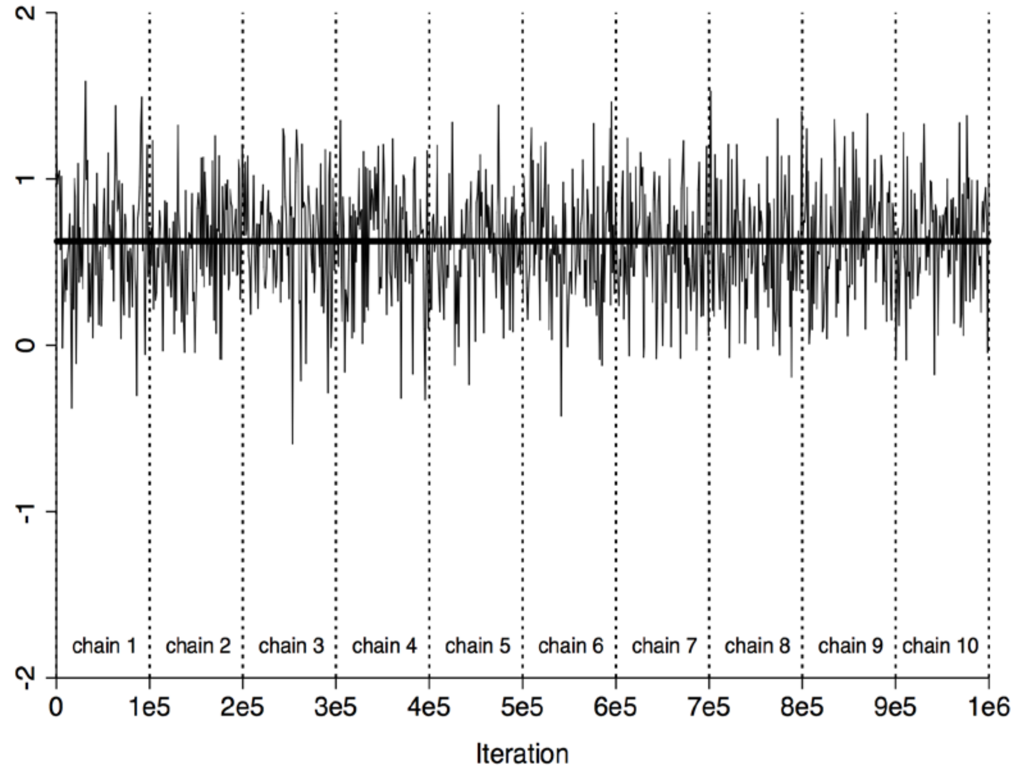
Figure A.13: Null Distribution for ψ in Equation (F.1)



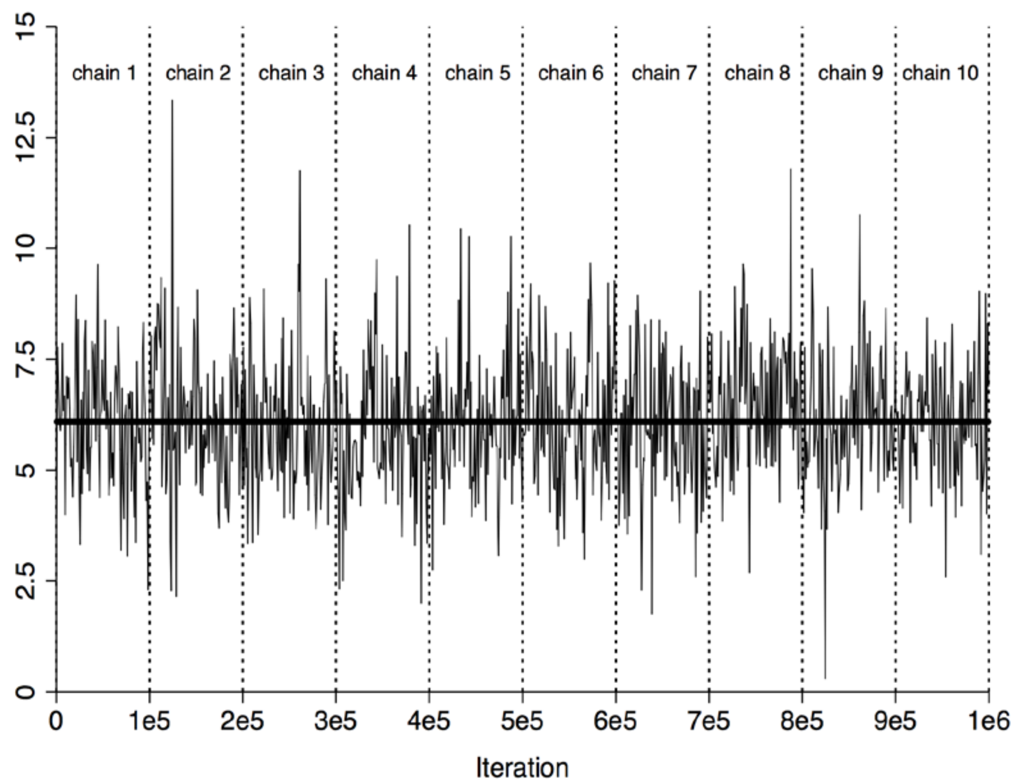
Notes: Figure shows the empirical cumulative distribution function for placebo estimates of ψ in equation (F.1) under the null hypothesis of no systematic preemption (solid line) vis-à-vis the actual point estimate (dashed line). Estimates are based on 10,000 randomly generated placebo orderings.

Figure A.14: Iterative History of δ in the MCMC Algorithm

A. Model with Homogenous Senators (Baseline Model)



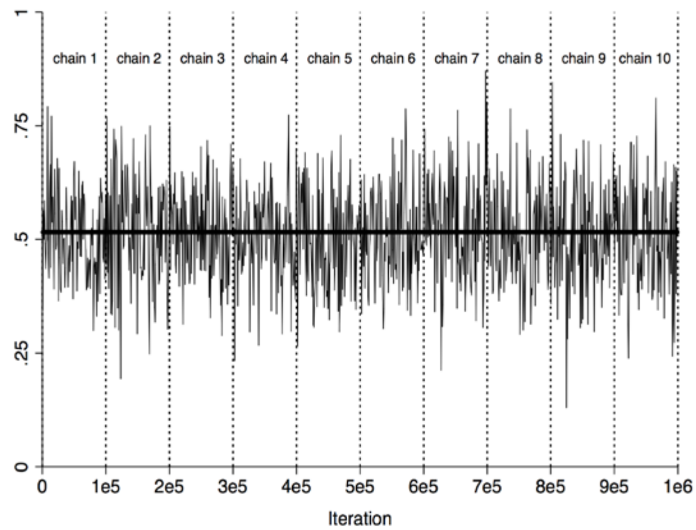
B. Model with 3 Types of Senators (Extended Model)



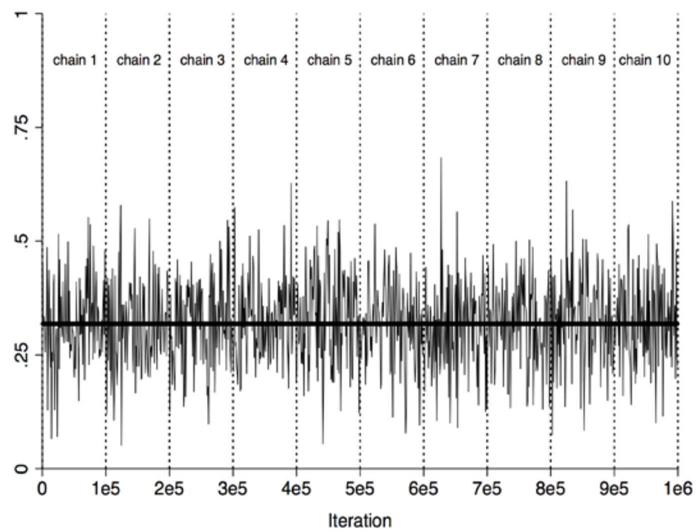
Notes: Figure shows the iterative history of δ in our MCMC algorithm (thin line) and its posterior mean (thick line). The upper panel pertains to our baseline model, while the lower one refers to our extended structural model. All parallel chains consist of 120,000 iterations, of which the first 20,000 are discarded as burn-in period. For this figure we thin each chain by a factor of 1,000.

Figure A.15: Iterative History of Type Shares in the MCMC Algorithm

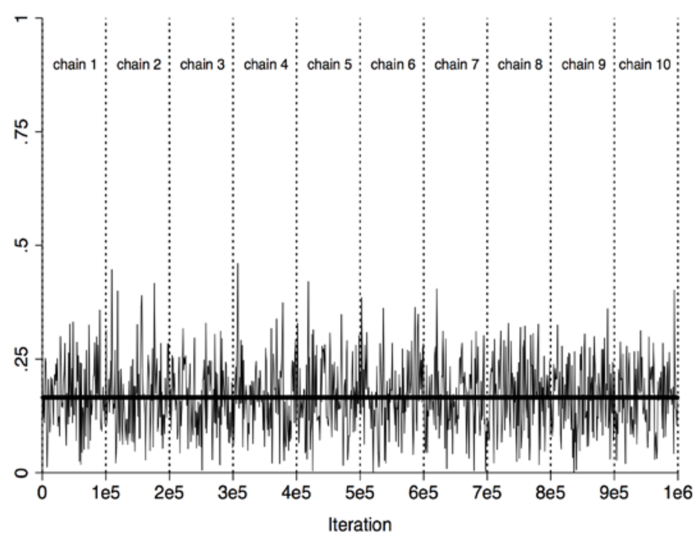
A. Share of Myopic Senators (π_1)



B. Share of Backward Reasoning Senators (π_2)

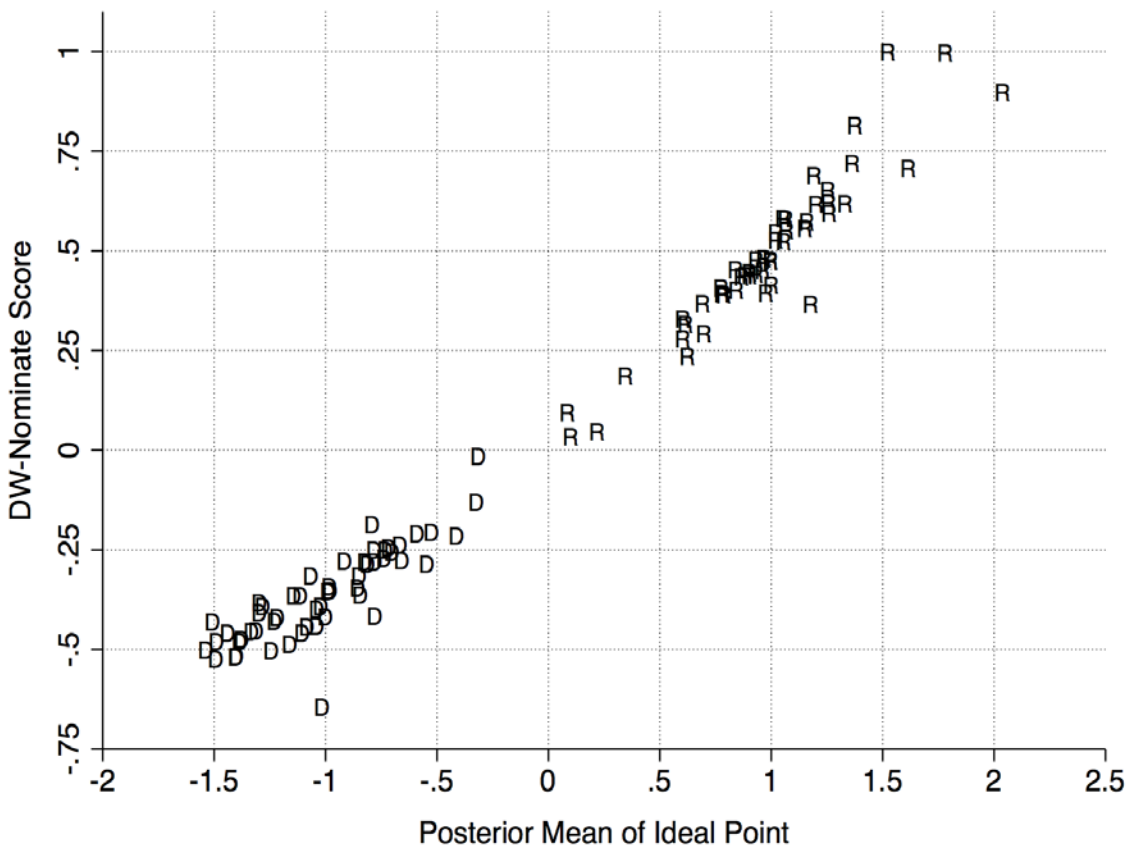


C. Share of "Intermediate Type" (π_3)



Notes: Figure shows the iterative history of π_1, π_2, and π_3 in the MCMC algorithm for our extended model (thin line) and their posterior mean (thick line). All parallel chains consist of 120,000 iterations, of which the first 20,000 are discarded as burn-in period. For this figure we thin each chain by a factor of 1,000.

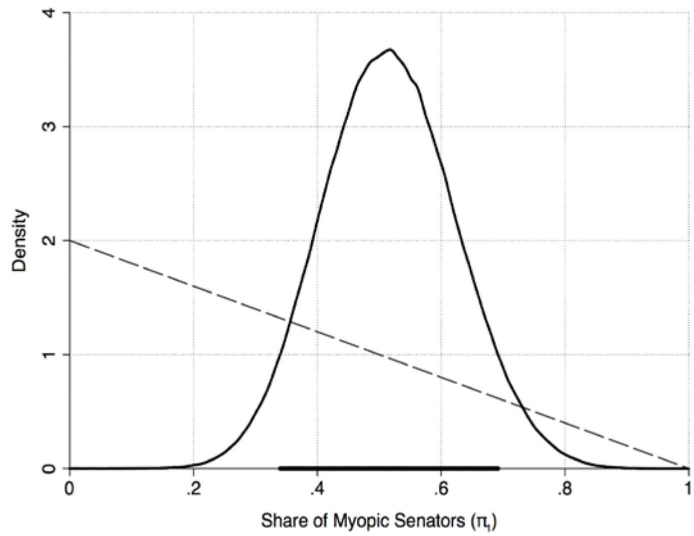
Figure A.16: Posterior Mean of Senators' Ideal Points, Extended Model



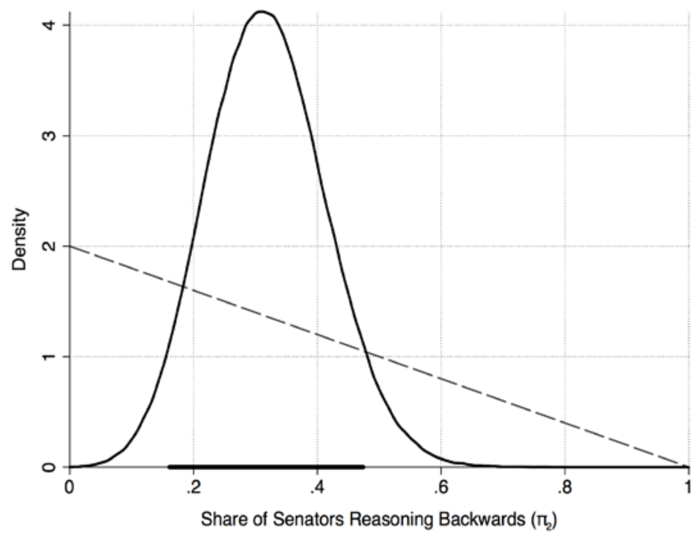
Notes: Figure plots the posterior mean of senators' ideal points in our extended structural model against the left–right dimension of DW-Nominate scores. "D" and "R" respectively denote Democratic and Republican senators. Estimates are based on 1,000,000 MCMC draws.

Figure A.17: Marginal Posterior of Type Shares, Extended Model

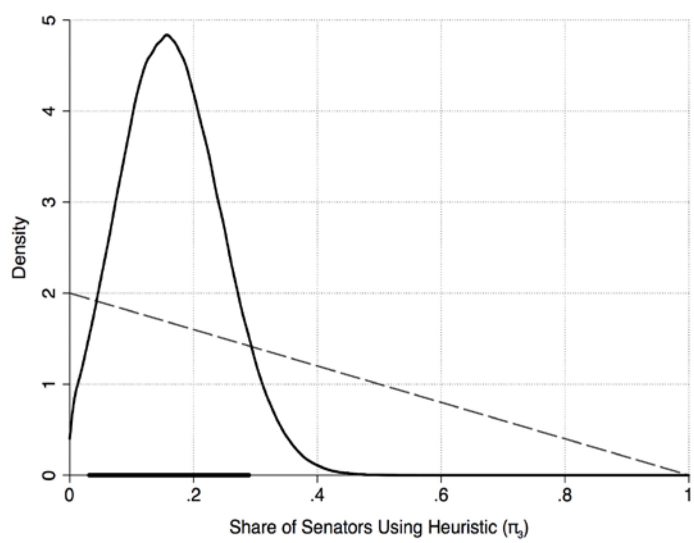
A. Posterior Share of Myopic Senators



B. Posterior Share of Backward-Reasoning Senators

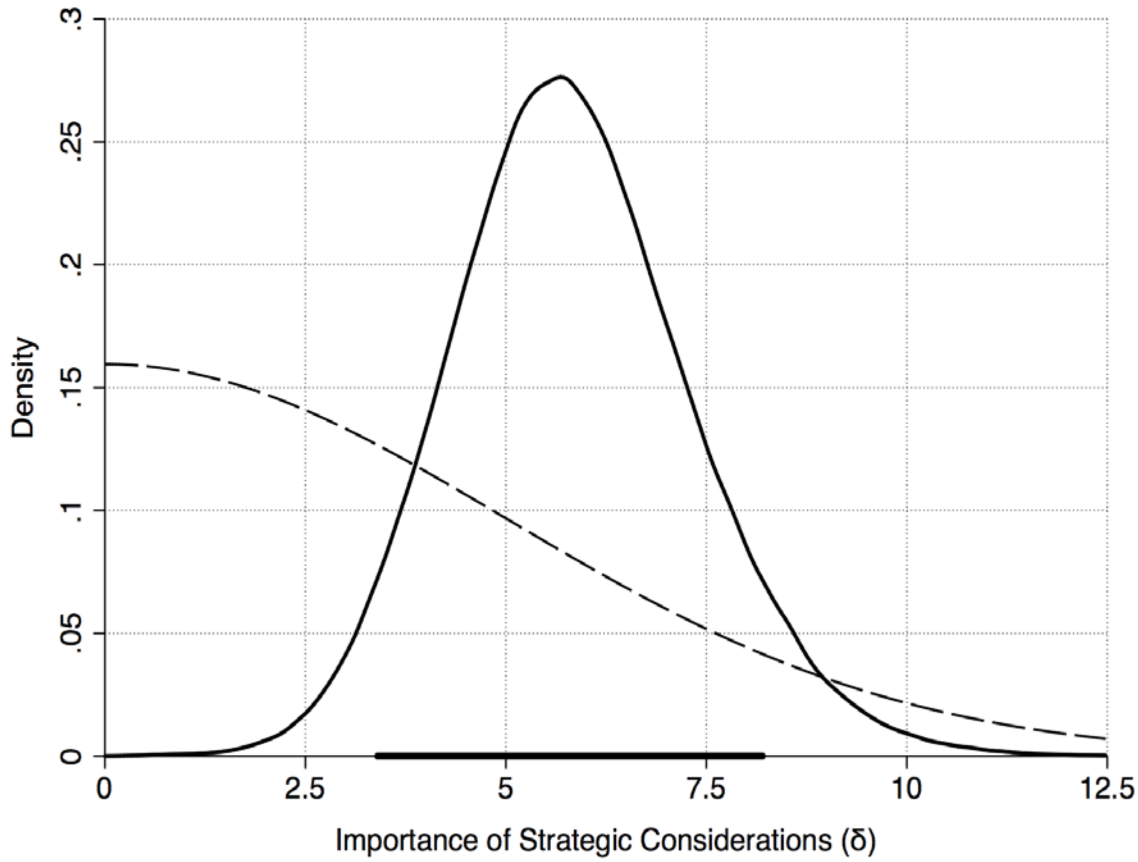


C. Posterior Share of "Intermediate Type"



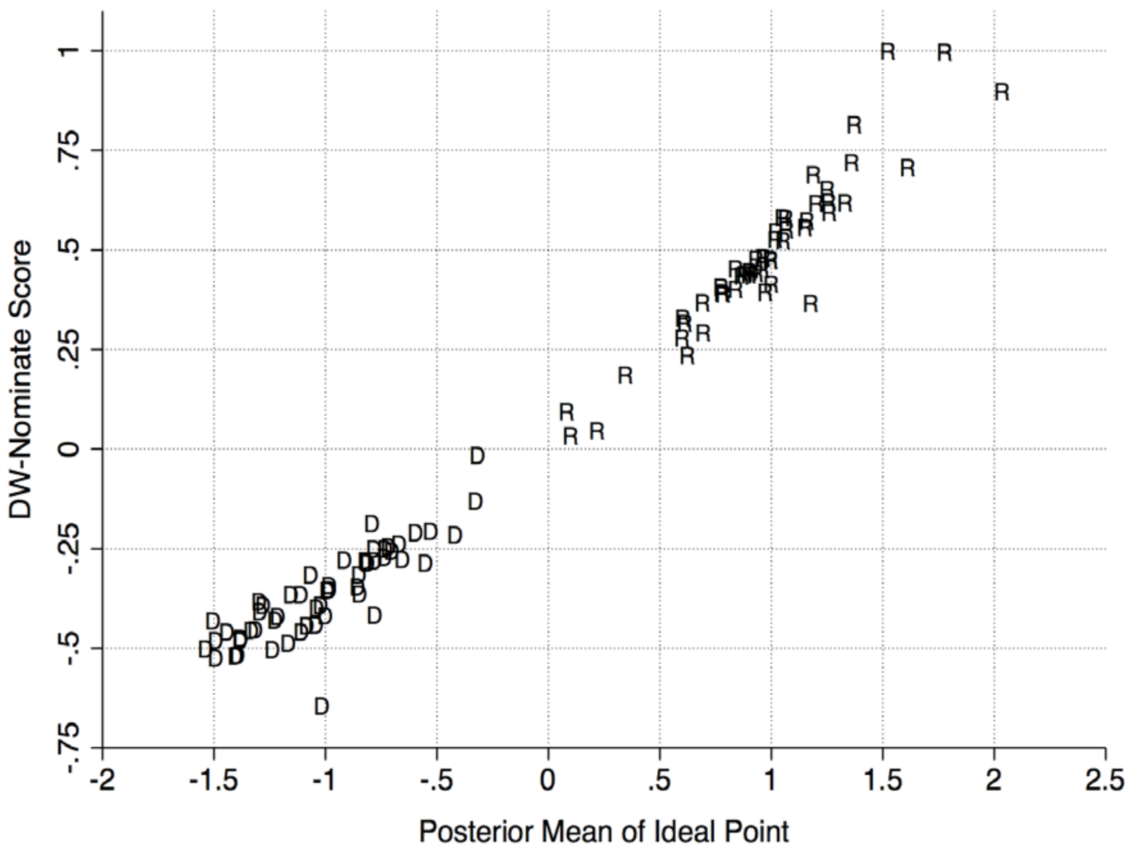
Notes: Figure shows the marginal posterior distributions for the type shares in our extended model (solid lines) as well as the associated priors (dashed lines). The upper panel refers to the share of myopic agents (π_1). The middle panel shows the fraction of backward inductors (π_2), while the lower one pertains to the share of senators using the heuristic outlined in the main text (π_3). All estimates are based on 1,000,000 MCMC draws and smoothed using a Gaussian kernel with a bandwidth of .005.

Figure A.18: Posterior Distribution of δ in a Model without "Intermediate Type"



Notes: Figure shows the marginal posterior density of δ in a variant of our structural model with myopic types and backward inductors only (solid line) as well as the associated prior (dashed line). The thick line indicates the 90% highest posterior density region. Estimates are based on 1,000,000 MCMC draws and smoothed using a Gaussian kernel with a bandwidth of .1.

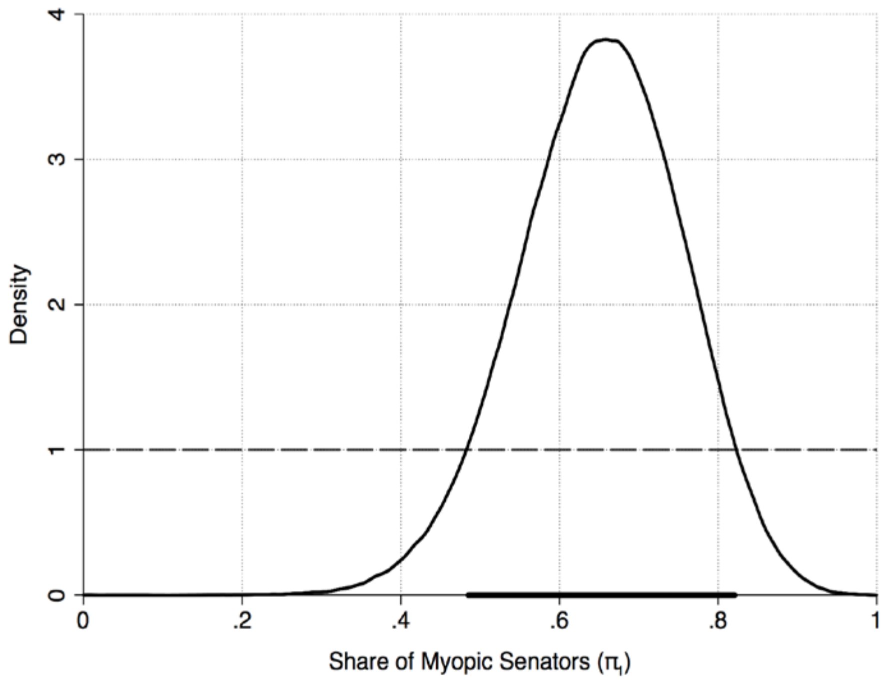
Figure A.19: Posterior Mean of Senators' Ideal Points, Model without "Intermediate Type"



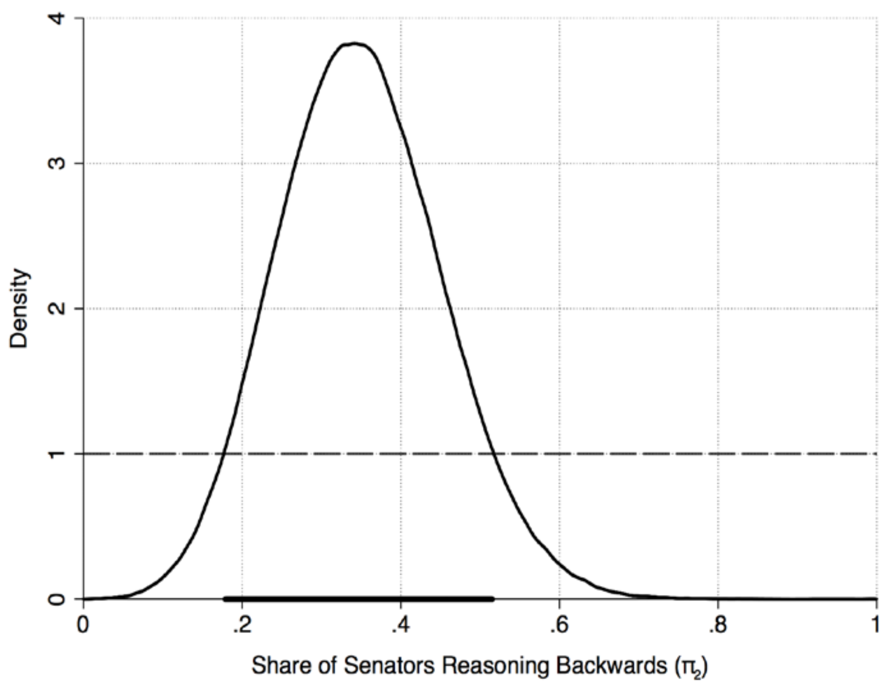
Notes: Figure plots the posterior mean of senators' ideal points in a variant of our structural model with myopic types and backward inductors only against the left-right dimension of DW-Nominate scores. "D" and "R" respectively denote Democratic and Republican senators. Estimates are based on 1,000,000 MCMC draws.

Figure A.20: Marginal Posterior of Type Shares in a Model without "Intermediate Types"

A. Marginal Posterior Share of Myopic Senators



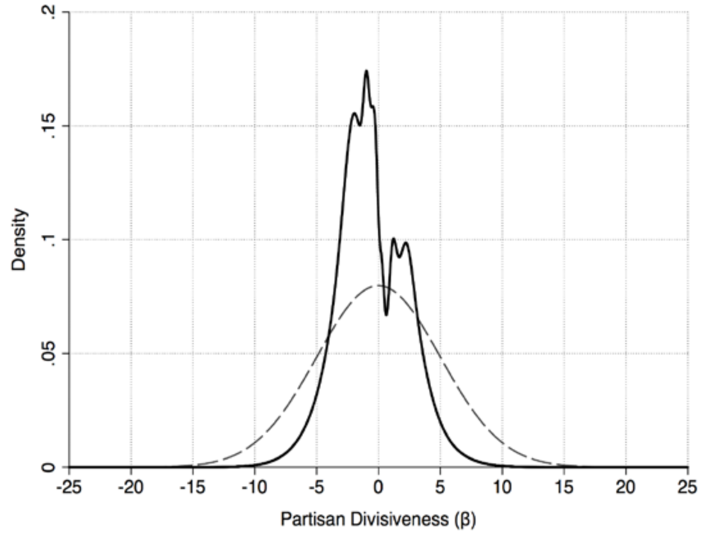
B. Marginal Posterior Share of Backward Reasoning Senators



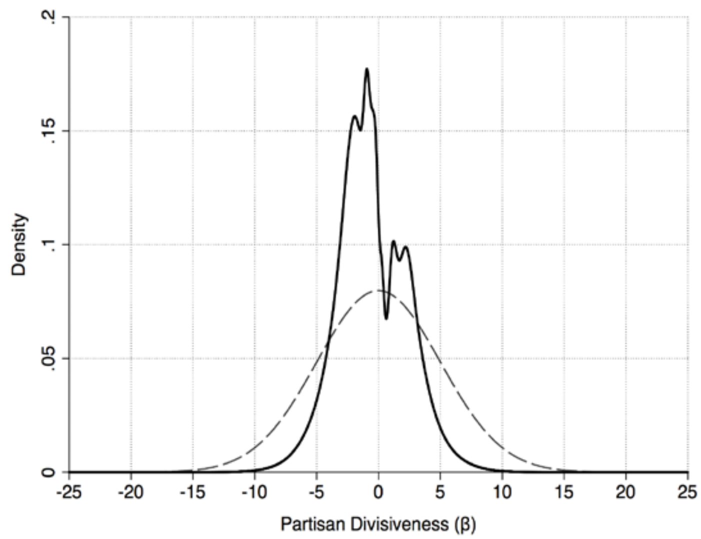
Notes: Figure shows the marginal posterior distributions for the type shares in a variant of our structural model with myopic types and backward inductors only (solid lines) as well as the associated priors (dashed lines). The upper panel refers to the share of myopic agents (π_1), while the lower one pertains to the share of senators reasoning backwards (π_2). All estimates are based on 1,000,000 MCMC draws and smoothed using a Gaussian kernel with a bandwidth of .005.

Figure A.21: Posterior of β, Pooling over Roll Calls

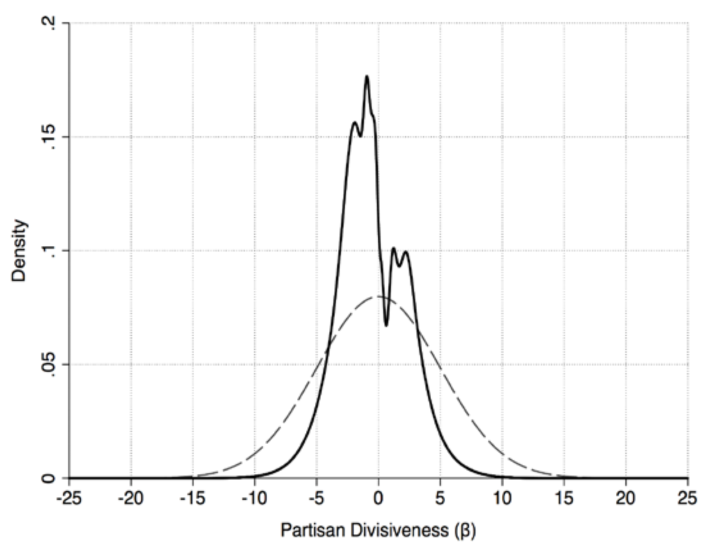
A. Baseline Model



B. Extended Structural Model



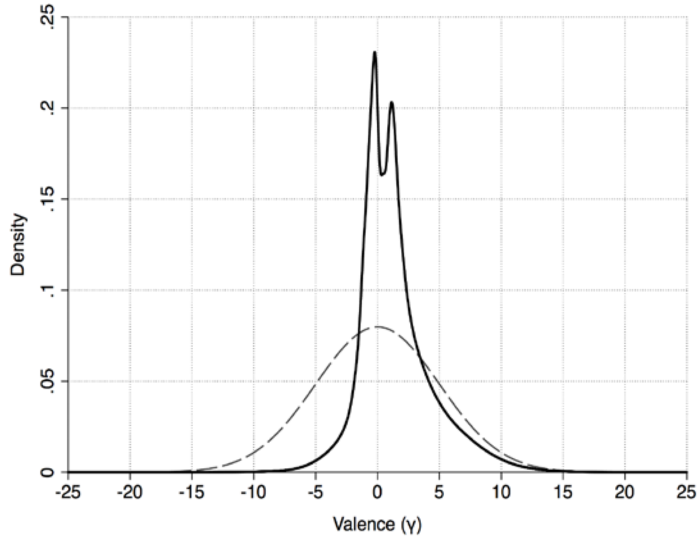
C. Model without "Intermediate" Type



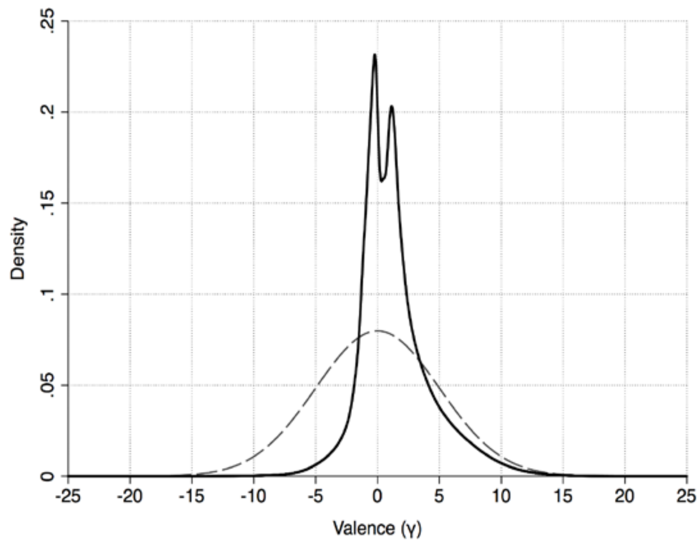
Notes: Figure shows the marginal posterior distribution of β pooling over all roll calls during the 112th Congress (solid lines) as well as the associated priors (dashed lines). The upper panel refers to our baseline model, while the middle and lower panels pertain to our extended structural model and a variant thereof with myopic types and backward inductors only. All estimates are based on 1,000,000 MCMC draws and smoothed using a Gaussian kernel with a bandwidth of .1.

Figure A.22: Posterior of γ, Pooling over Roll Calls

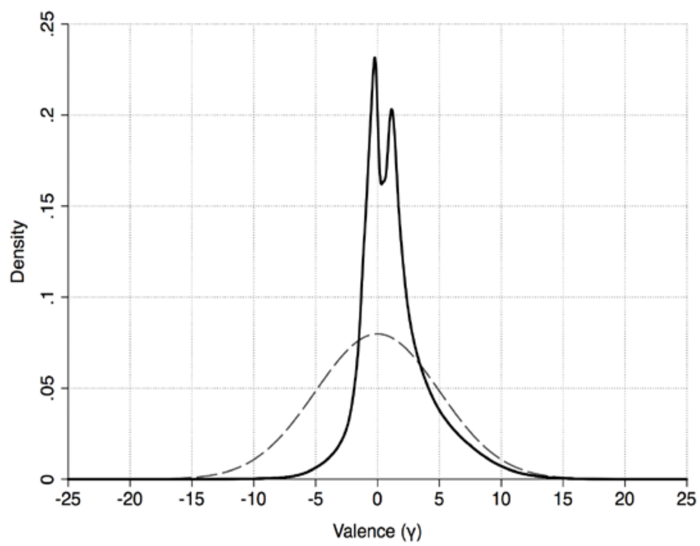
A. Baseline Model



B. Extended Structural Model

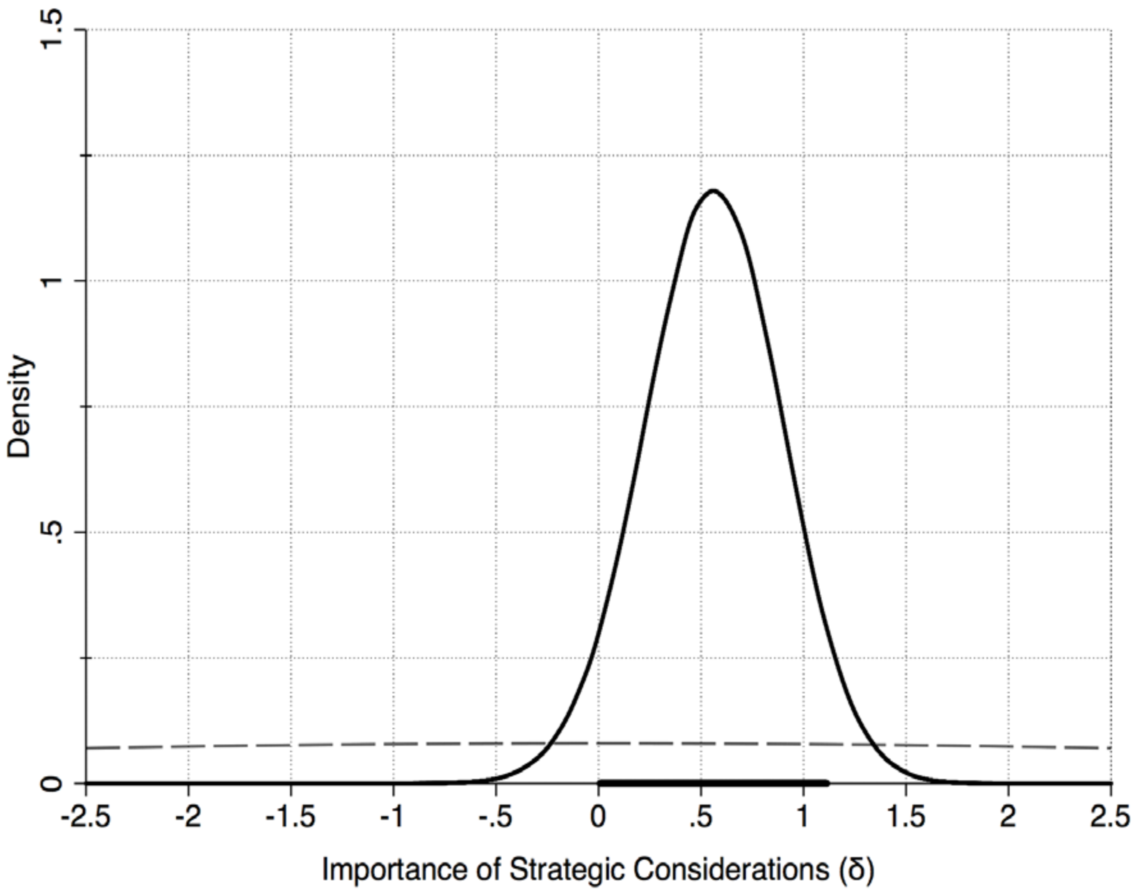


C. Model without "Intermediate" Type



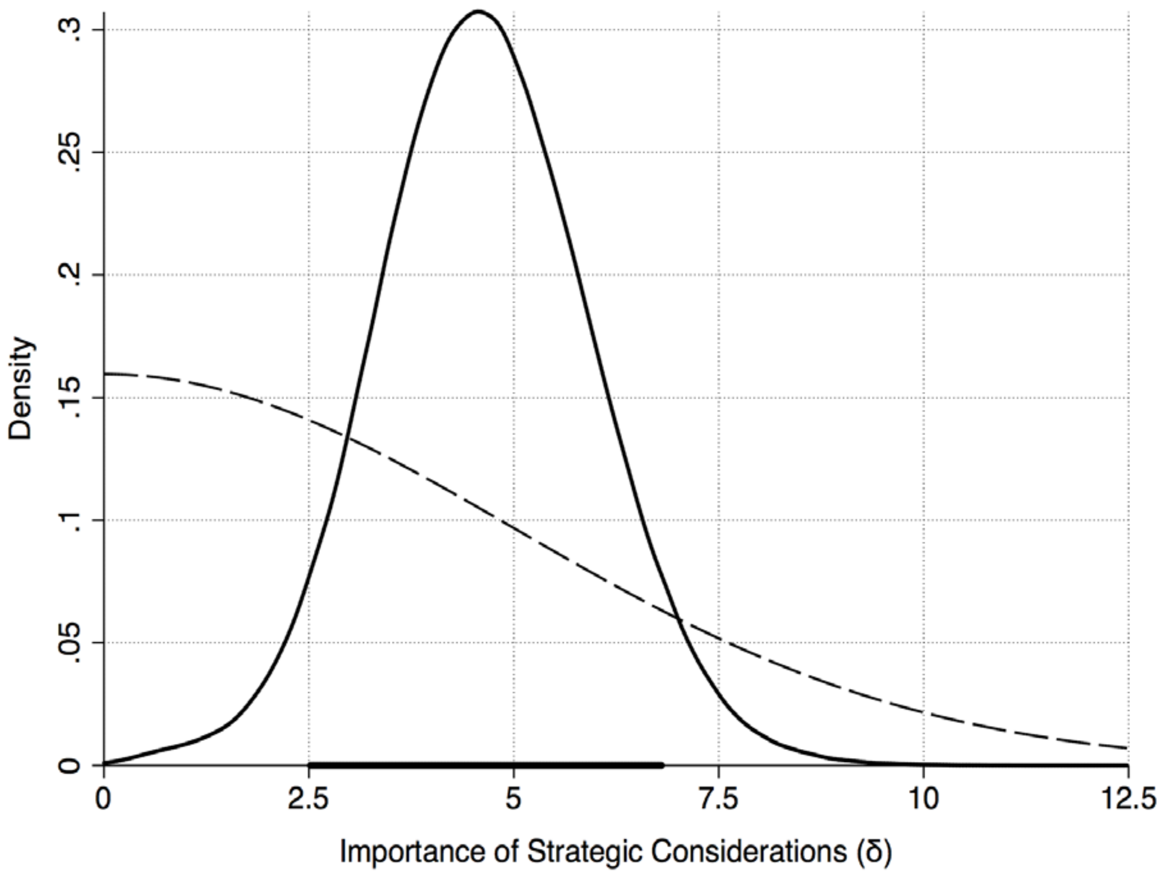
Notes: Figure shows the marginal posterior distribution of γ pooling over all roll calls during the 112th Congress (solid lines) as well as the associated priors (dashed lines). The upper panel refers to our baseline model, while the middle and lower panels pertain to our extended structural model and a variant thereof with myopic types and backward inductors only. All estimates are based on 1,000,000 MCMC draws and smoothed using a Gaussian kernel with a bandwidth of .1.

Figure A.23: Posterior Distribution of δ, Baseline Model with Logit Errors



Notes: Figure shows the marginal posterior density of δ (solid line) as well as the associated prior (dashed line) in a variant of our baseline model with logit errors. To ensure comparability with the results in the main text, estimates are scaled by the standard deviation of the logit distribution. The thick line indicates the 90% highest posterior density region. Estimates are based on 1,000,000 MCMC draws and smoothed using a Gaussian kernel with a bandwidth of .1.

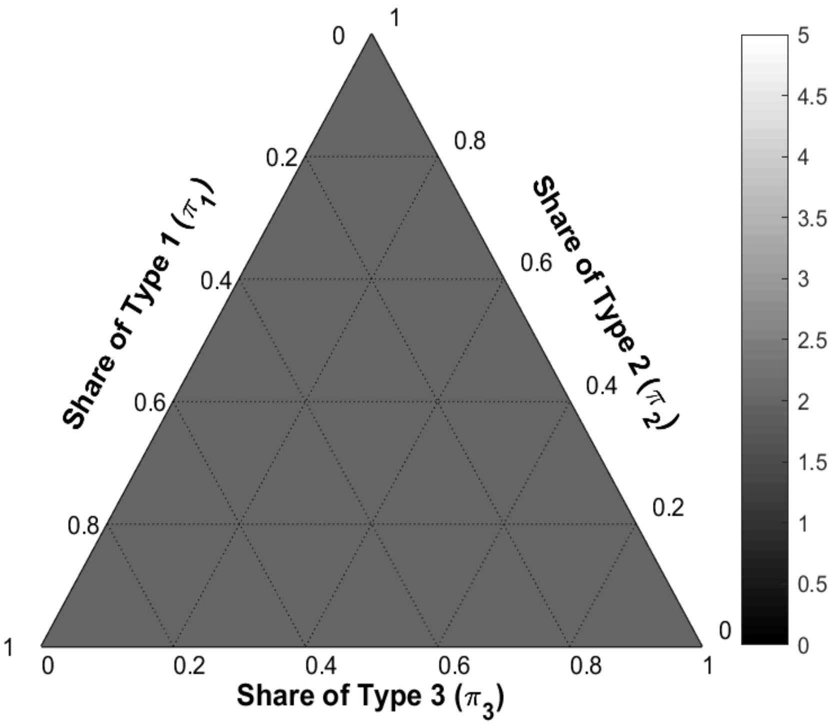
Figure A.24: Posterior Distribution of δ, Extended Model with Logit Errors



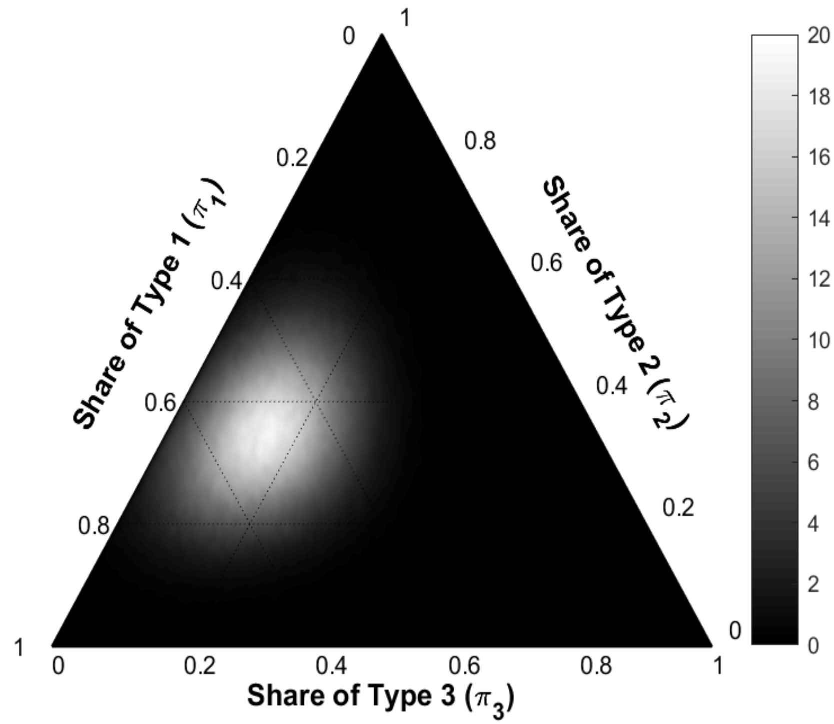
Notes: Figure shows the marginal posterior density of δ (solid line) as well as the associated prior (dashed line) in a variant of our extended model with logit errors. To ensure comparability with the results in the main text, estimates are scaled by the standard deviation of the logit distribution. The thick line indicates the 90% highest posterior density region. Estimates are based on 1,000,000 MCMC draws and smoothed using a Gaussian kernel with a bandwidth of .1.

Figure A.25: Joint Distribution of Type Shares, Extended Model with Logit Errors

A. Prior

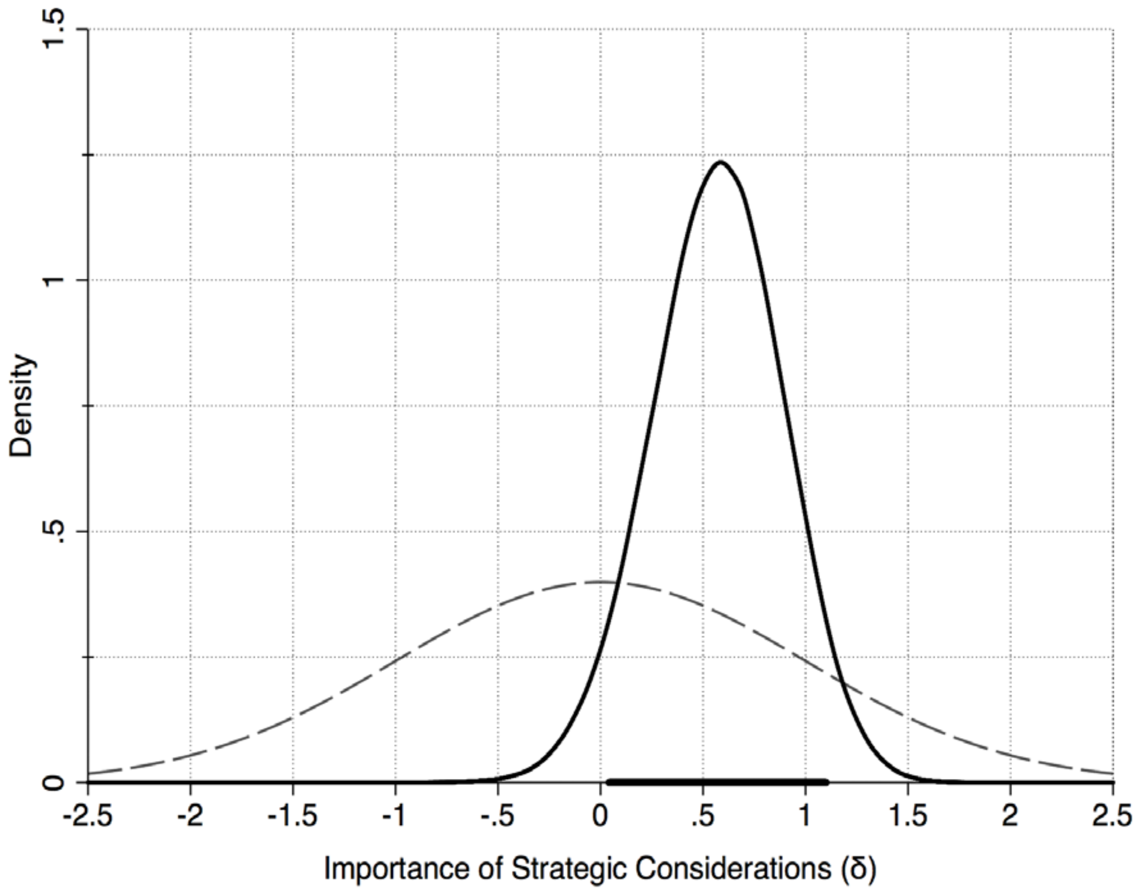


B. Posterior



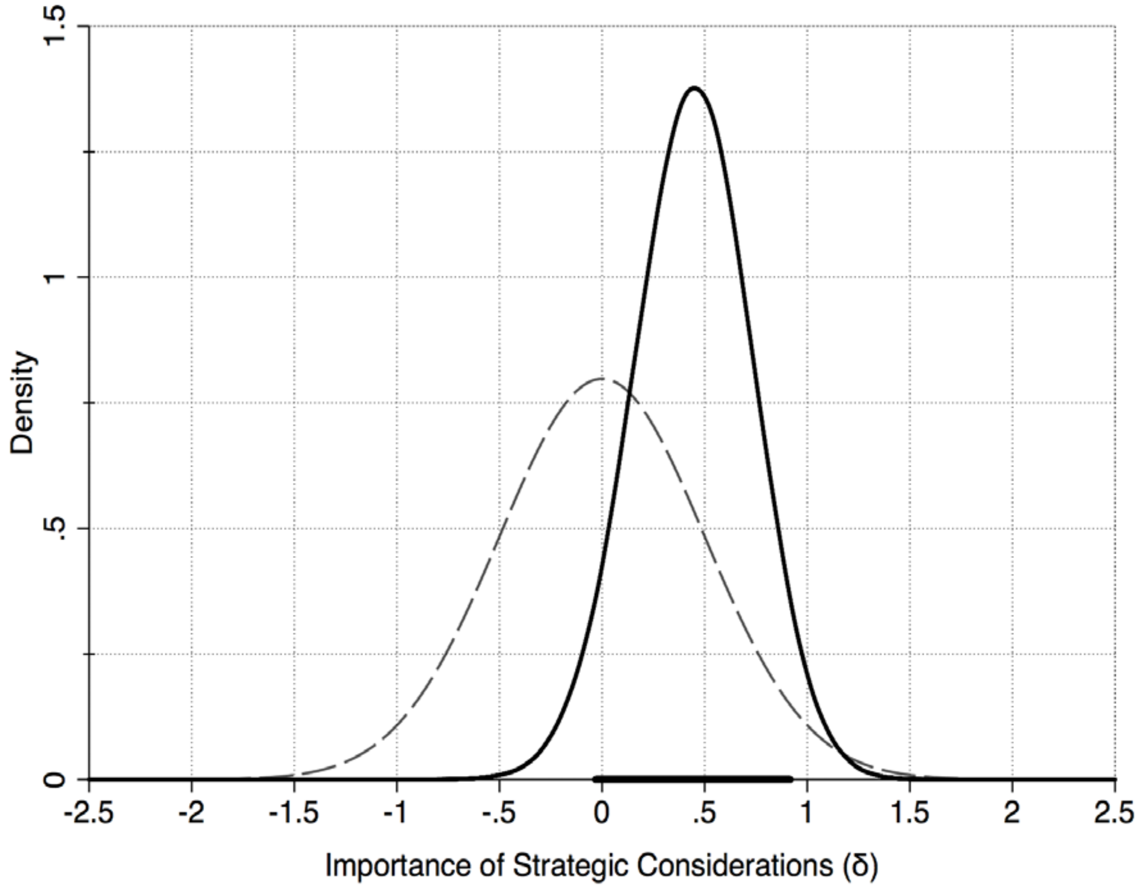
Notes: Figure shows the prior (left) and posterior (right) joint distribution for the type shares in a variant of our extended model with logit errors. Estimates of the posterior are based on 1,000,000 MCMC draws and smoothed using a Gaussian kernel with a bandwidth of .005.

Figure A.26: Posterior Distribution of δ, Baseline Model with Alternative Prior (I/II)



Notes: Figure shows the marginal posterior density of δ (solid line) as well as the associated prior (dashed line) in a variant of our baseline model that uses a normal normal distribution with mean 0 and variance 1 as prior on δ. The thick line indicates the 90% highest posterior density region. Estimates are based on 1,000,000 MCMC draws and smoothed using a Gaussian kernel with a bandwidth of .1.

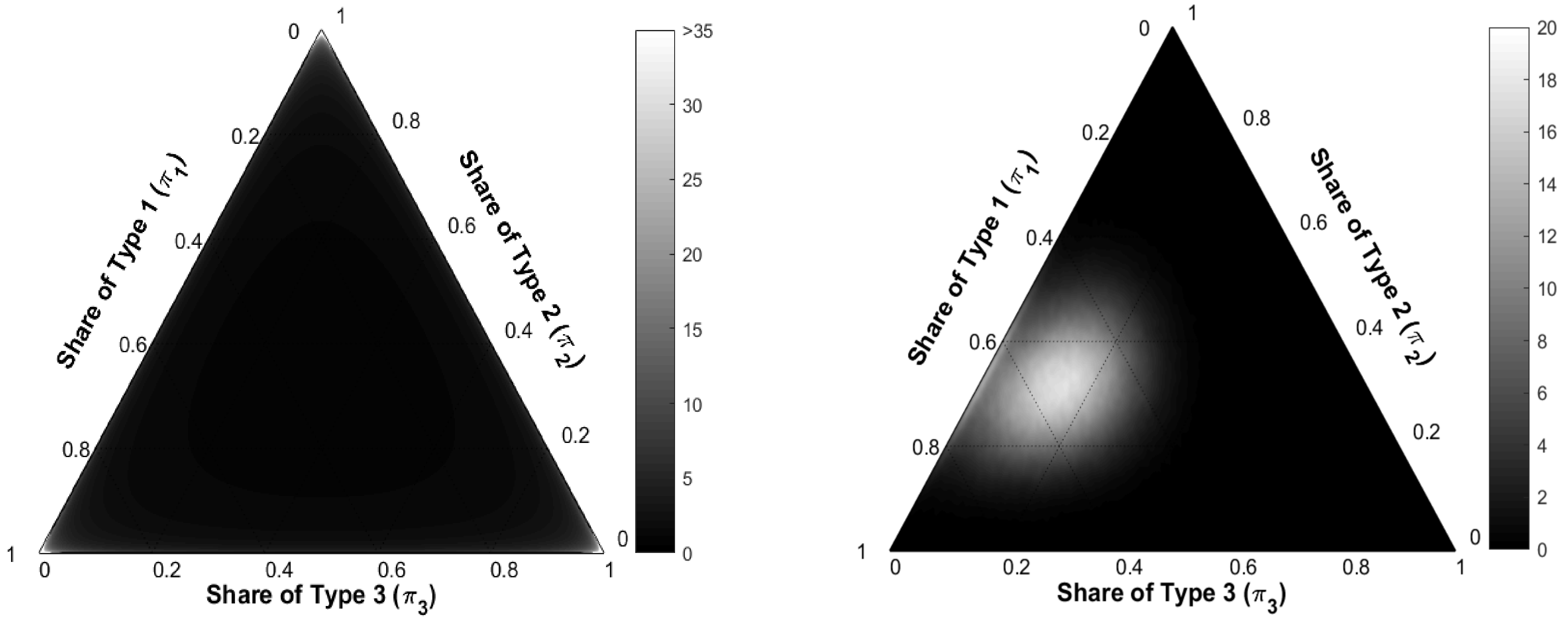
Figure A.27: Posterior Distribution of δ, Baseline Model with Alternative Prior



Notes: Figure shows the marginal posterior density of δ (solid line) as well as the associated prior (dashed line) in a variant of our baseline model that uses a normal normal distribution with mean 0 and variance .25 as prior on δ. The thick line indicates the 90% highest posterior density region. Estimates are based on 1,000,000 MCMC draws and smoothed using a Gaussian kernel with a bandwidth of .1.

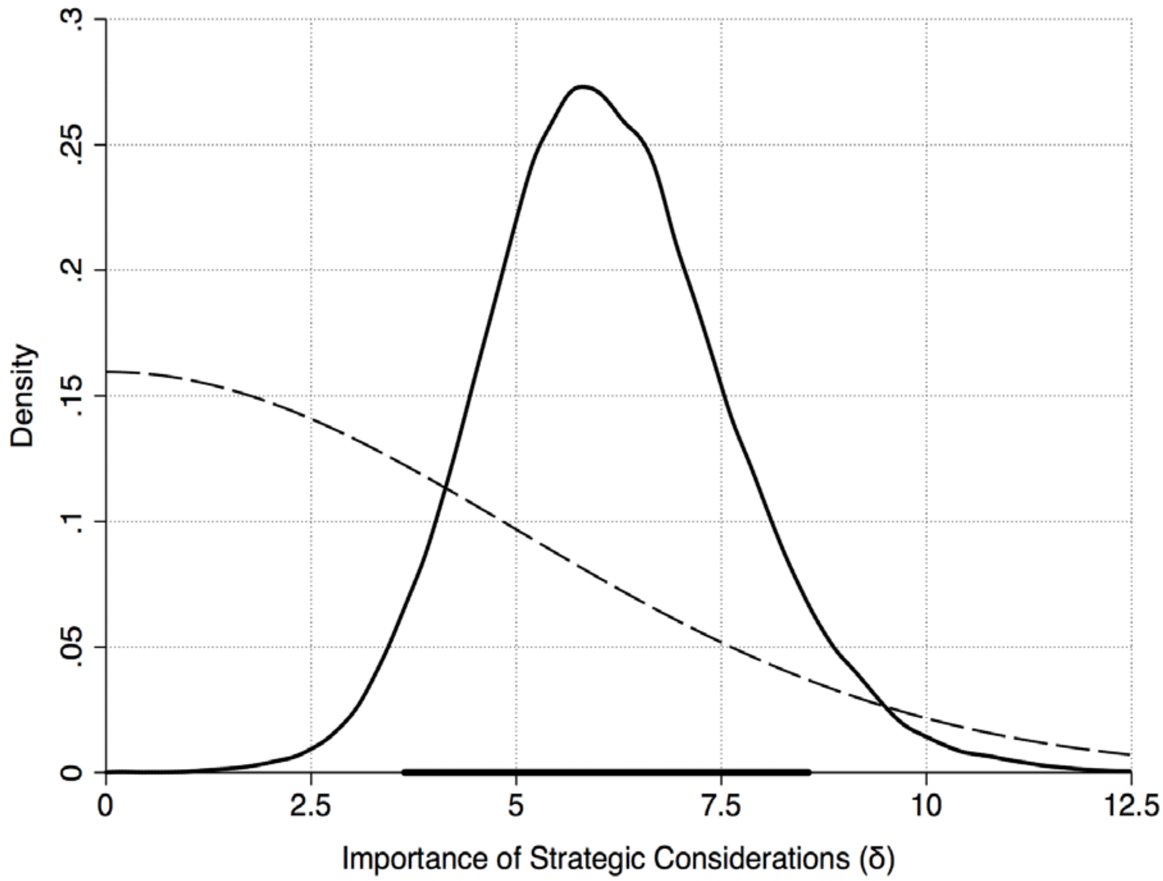
Figure A.28: Joint Distribution of Type Shares for Alternative Prior (I/V)

A. Prior *B. Posterior*



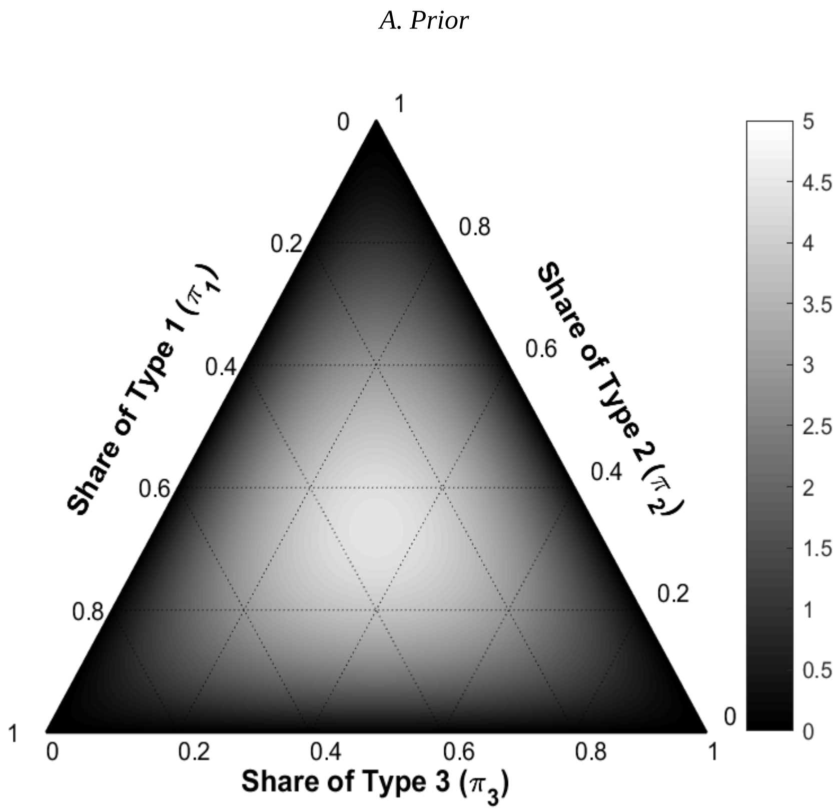
Notes: Figure shows the prior (left) and posterior (right) joint distribution for the type shares in our extended model, using a Dirichlet distribution with parameters (.5,.5,.5) as an alternative prior for (π_1,π_2,π_3). Estimates of the posterior are based on 1,000,000 MCMC draws and smoothed using a Gaussian kernel with a bandwidth of .005.

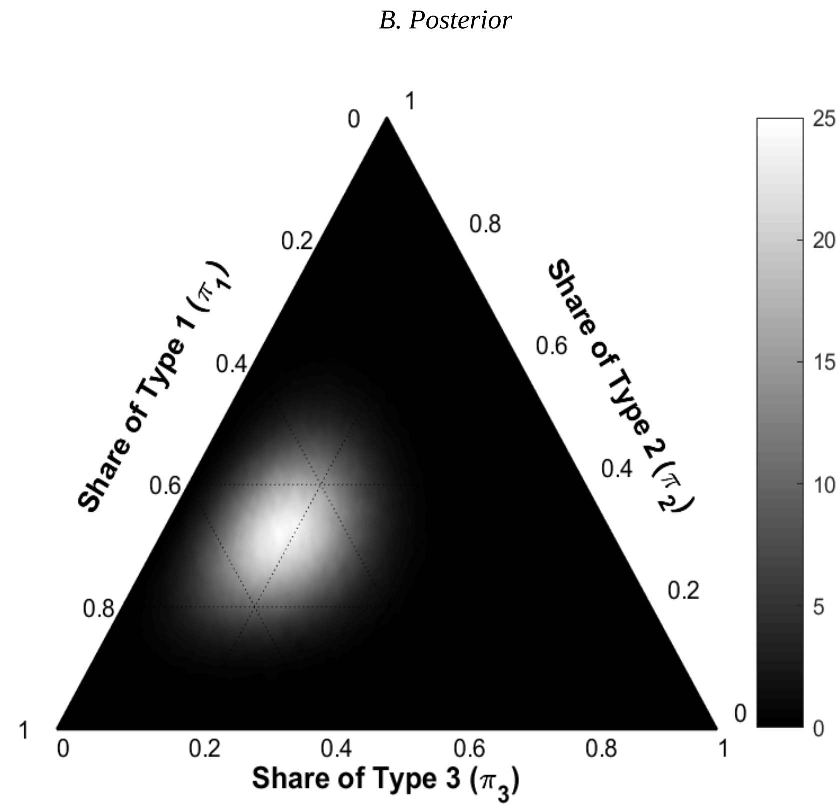
Figure A.29: Posterior Distribution of δ, Extended Model with Alternative Prior on (π_1,π_2,π_3) (I/V)



Notes: Figure shows the marginal posterior density of δ (solid line) as well as the associated prior (dashed line) in a variant of our extended model with a Dirichlet distribution with parameters (.5,.5,5) as alternative prior on (π_1,π_2,π_3). The thick line indicates the 90% highest posterior density region. Estimates are based on 1,000,000 MCMC draws and smoothed using a Gaussian kernel with a bandwidth of .1.

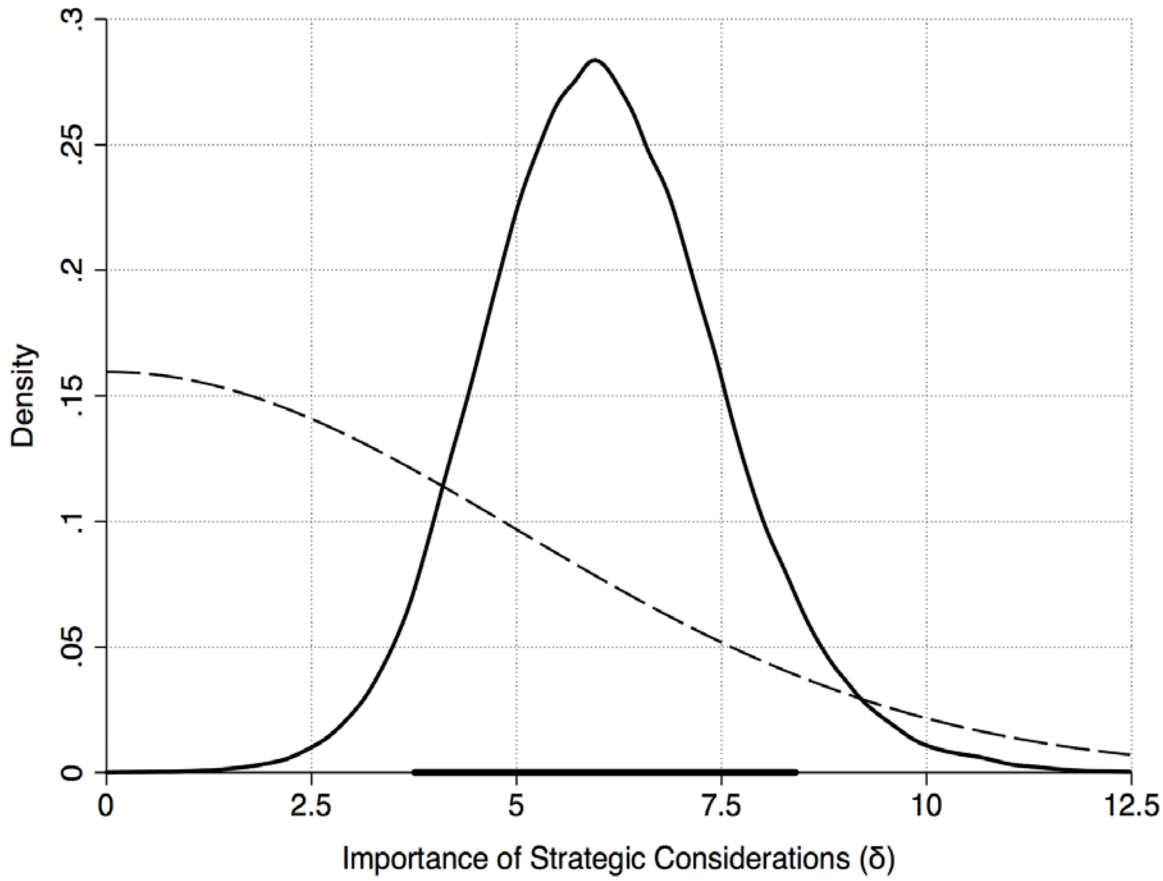
Figure A.30: Joint Distribution of Type Shares for Alternative Prior (II/V)





Notes: Figure shows the prior (left) and posterior (right) joint distribution for the type shares in our extended model, using a Dirichlet distribution with parameters (2,2,2) as an alternative prior for (π_1,π_2,π_3). Estimates of the posterior are based on 1,000,000 MCMC draws and smoothed using a Gaussian kernel with a bandwidth of .005.

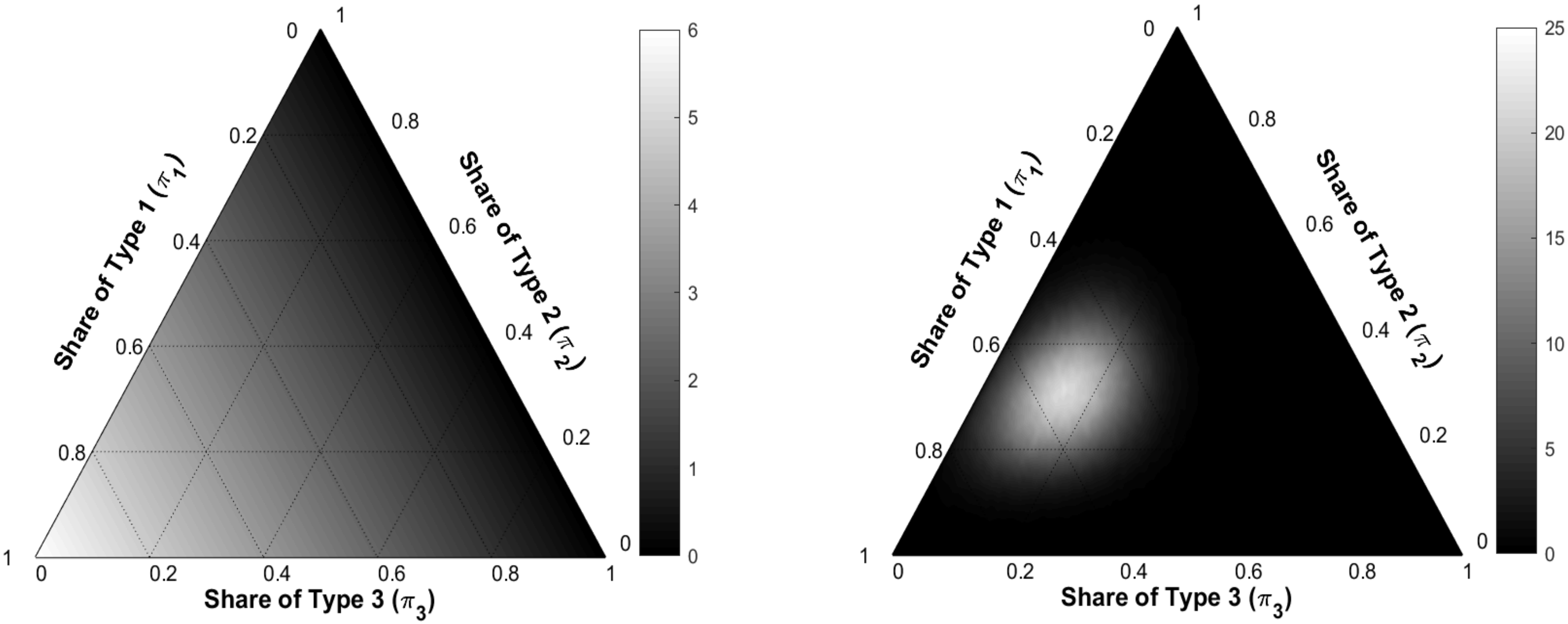
Figure A.31: Posterior Distribution of δ, Extended Model with Alternative Prior on (π_1,π_2,π_3) (II/V)



Notes: Figure shows the marginal posterior density of δ (solid line) as well as the associated prior (dashed line) in a variant of our extended model with a Dirichlet distribution with parameters (2,2,2) as alternative prior on (π_1,π_2,π_3). The thick line indicates the 90% highest posterior density region. Estimates are based on 1,000,000 MCMC draws and smoothed using a Gaussian kernel with a bandwidth of .1.

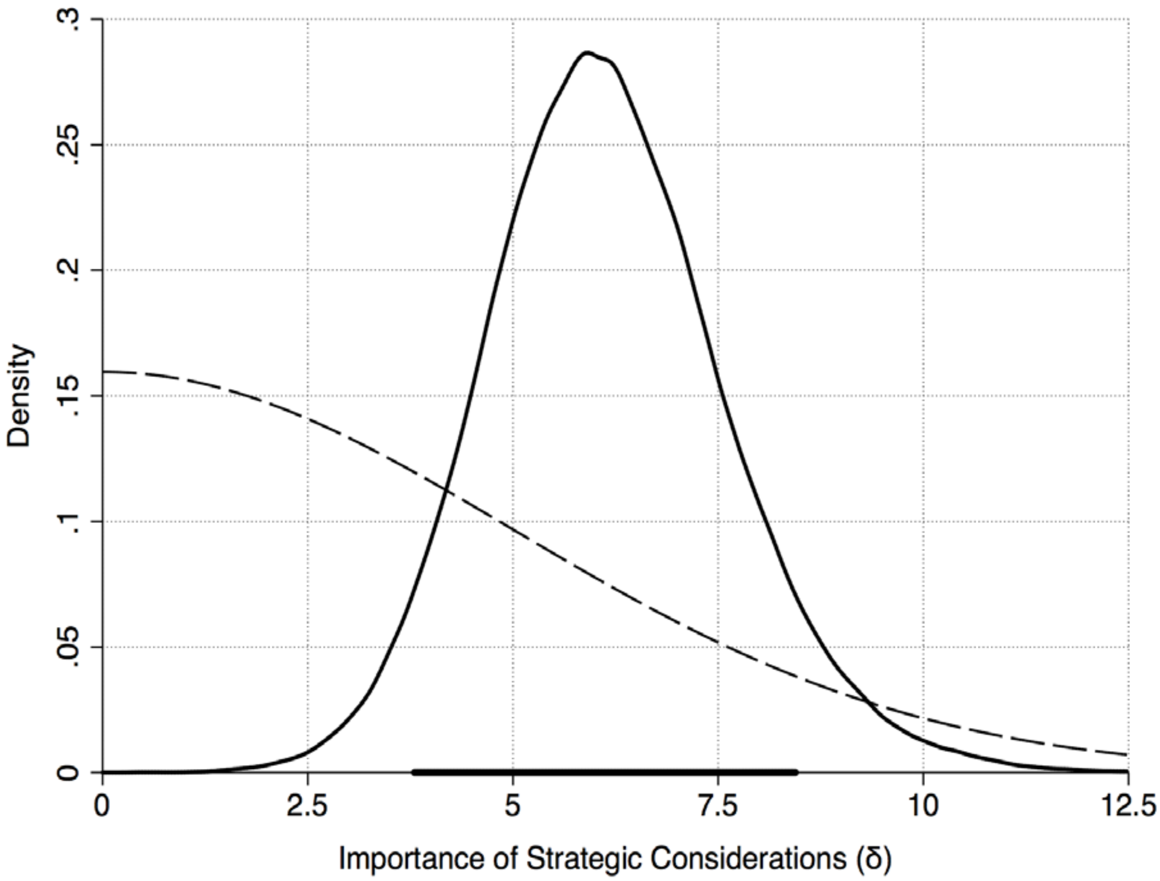
Figure A.32: Joint Distribution of Type Shares for Alternative Prior (III/V)

A. Prior *B. Posterior*



Notes: Figure shows the prior (left) and posterior (right) joint distribution for the type shares in our extended model, using a Dirichlet distribution with parameters (2,1,1) as an alternative prior for (π_1,π_2,π_3). Estimates of the posterior are based on 1,000,000 MCMC draws and smoothed using a Gaussian kernel with a bandwidth of .005.

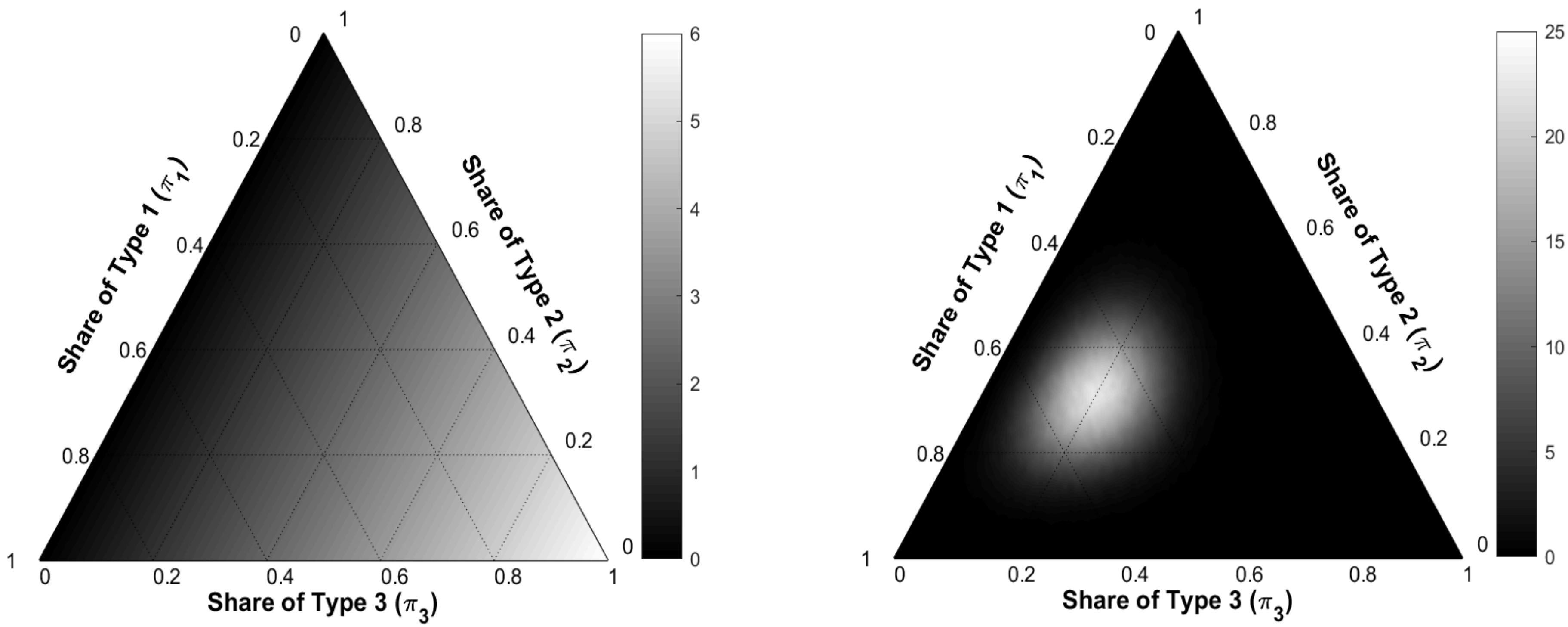
Figure A.33: Posterior Distribution of δ, Extended Model with Alternative Prior on (π_1,π_2,π_3) (III/V)



Notes: Figure shows the marginal posterior density of δ (solid line) as well as the associated prior (dashed line) in a variant of our extended model with a Dirichlet distribution with parameters (2,1,1) as alternative prior on (π_1,π_2,π_3). The thick line indicates the 90% highest posterior density region. Estimates are based on 1,000,000 MCMC draws and smoothed using a Gaussian kernel with a bandwidth of .1.

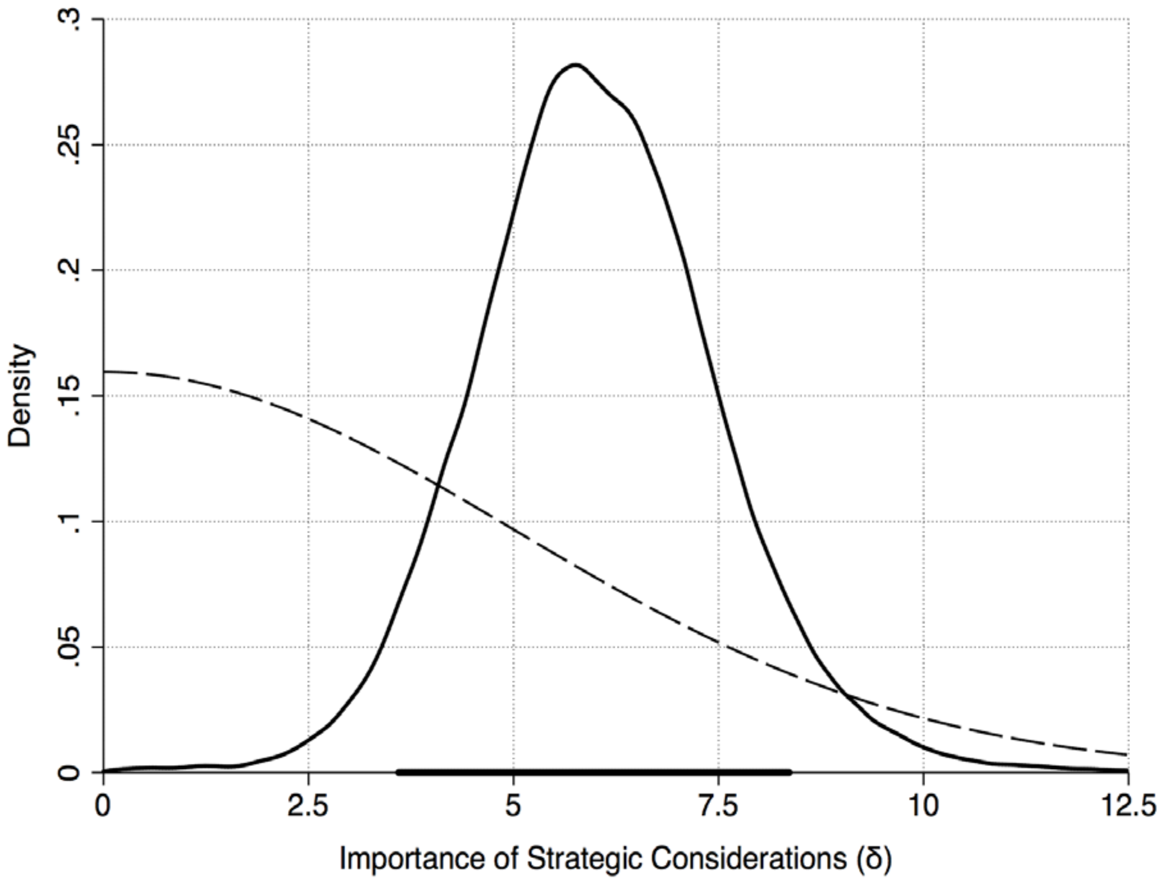
Figure A.34: Joint Distribution of Type Shares for Alternative Prior (IV/V)

A. Prior *B. Posterior*



Notes: Figure shows the prior (left) and posterior (right) joint distribution for the type shares in our extended model, using a Dirichlet distribution with parameters (1,1,2) as an alternative prior for (π_1,π_2,π_3). Estimates of the posterior are based on 1,000,000 MCMC draws and smoothed using a Gaussian kernel with a bandwidth of .005.

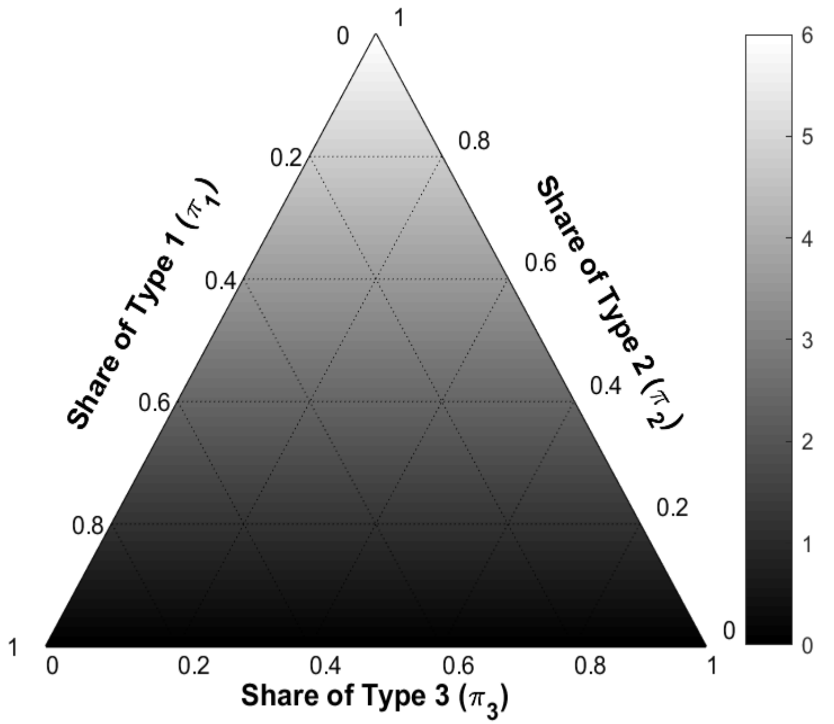
Figure A.35: Posterior Distribution of δ, Extended Model with Alternative Prior on (π_1,π_2,π_3) (IV/V)



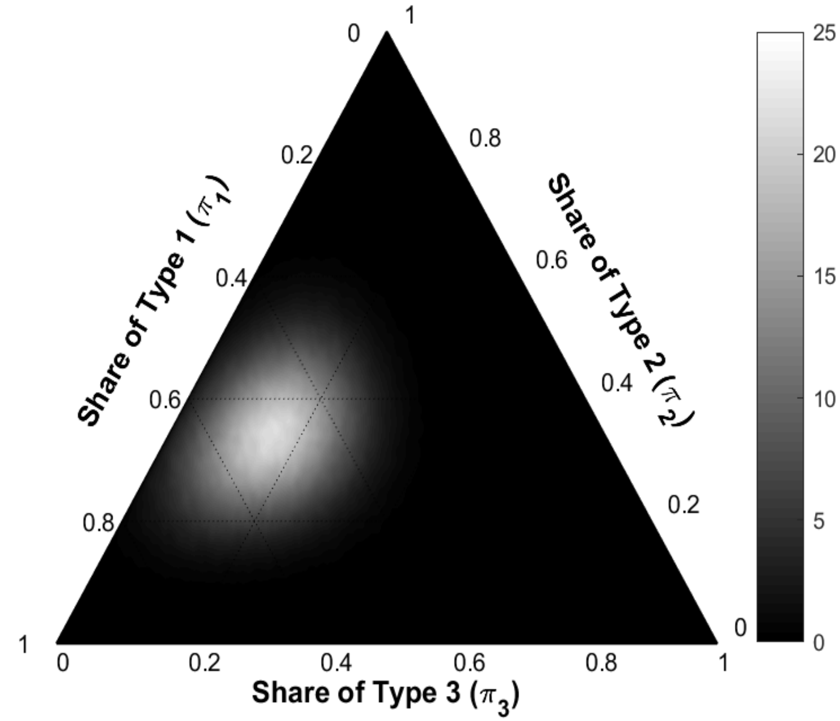
Notes: Figure shows the marginal posterior density of δ (solid line) as well as the associated prior (dashed line) in a variant of our extended model with a Dirichlet distribution with parameters (1,1,2) as alternative prior on (π_1,π_2,π_3). The thick line indicates the 90% highest posterior density region. Estimates are based on 1,000,000 MCMC draws and smoothed using a Gaussian kernel with a bandwidth of .1.

Figure A.36: Joint Distribution of Type Shares for Alternative Prior (V/V)

A. Prior

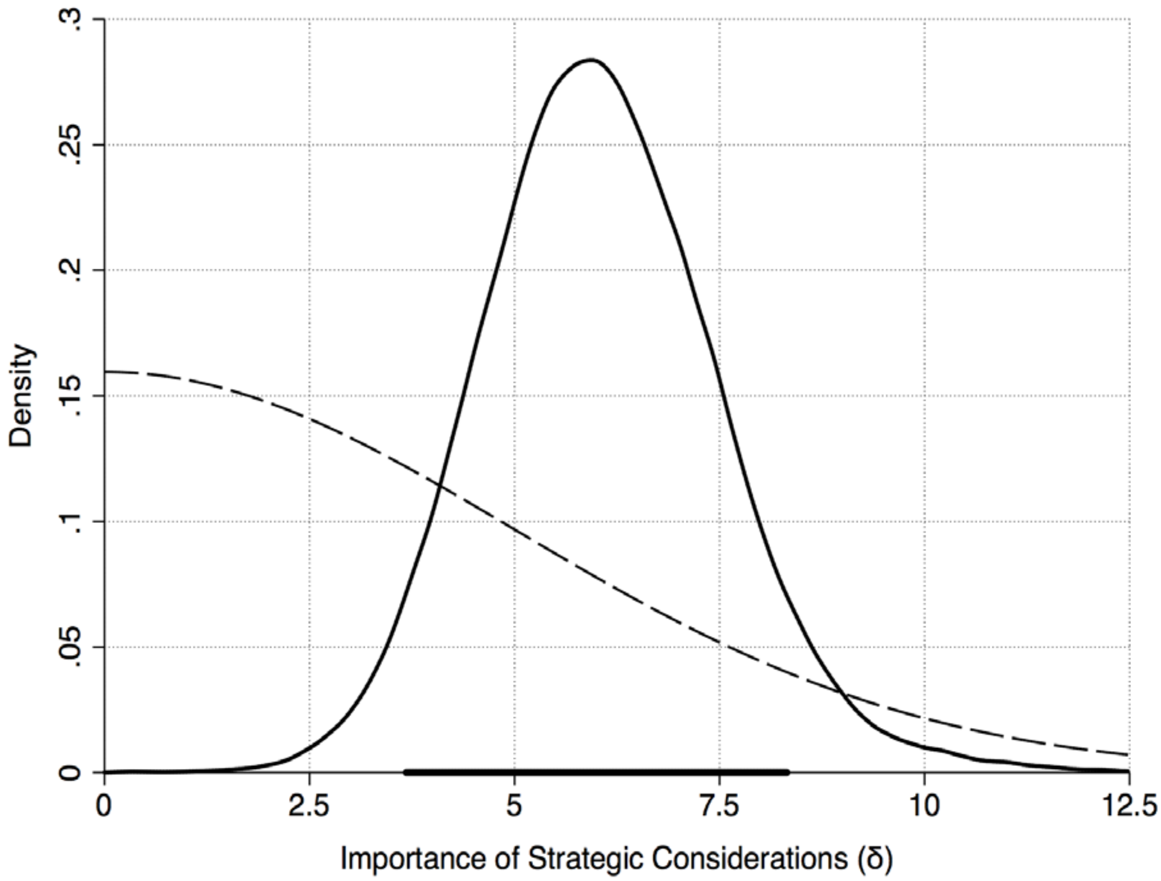


B. Posterior



Notes: Figure shows the prior (left) and posterior (right) joint distribution for the type shares in our extended model, using a Dirichlet distribution with parameters (1,2,1) as an alternative prior for (π_1,π_2,π_3). Estimates of the posterior are based on 1,000,000 MCMC draws and smoothed using a Gaussian kernel with a bandwidth of .005.

Figure A.37: Posterior Distribution of δ, Extended Model with Alternative Prior on (π_1,π_2,π_3) (V/V)



Notes: Figure shows the marginal posterior density of δ (solid line) as well as the associated prior (dashed line) in a variant of our extended model with a Dirichlet distribution with parameters (1,2,1) as alternative prior on (π_1,π_2,π_3). The thick line indicates the 90% highest posterior density region. Estimates are based on 1,000,000 MCMC draws and smoothed using a Gaussian kernel with a bandwidth of .1.

Table A.1: Deviations from the Party Line, by Voting Position

Position	Raw Mean	Standard Error
1	.185	.002
2	.193	.002
3	.170	.002
4	.182	.002
5	.193	.002
6	.187	.002
7	.189	.002
8	.197	.002
9	.186	.002
10	.194	.002
11	.193	.002
12	.204	.002
13	.190	.002
14	.192	.002
15	.191	.002
16	.198	.002
17	.197	.002
18	.194	.002
19	.199	.002
20	.189	.002
21	.189	.002
22	.184	.002
23	.187	.002
24	.191	.002
25	.193	.002
26	.193	.002
27	.187	.002
28	.198	.002
29	.189	.002
30	.195	.002
31	.190	.002
32	.192	.002
33	.191	.002
34	.182	.002
35	.183	.002
36	.181	.002
37	.178	.002
38	.179	.002
39	.184	.002
40	.184	.002
41	.191	.002
42	.188	.002
43	.186	.002
44	.191	.002
45	.186	.002
46	.191	.002
47	.189	.002
48	.184	.002
49	.182	.002
50	.179	.002
51	.181	.002
52	.181	.002
53	.185	.002
54	.177	.002
55	.183	.002
56	.180	.002
57	.172	.002
58	.167	.002
59	.174	.002
60	.176	.002
61	.163	.002
62	.168	.002
63	.156	.002
64	.160	.002
65	.168	.002
66	.168	.002
67	.163	.002
68	.169	.002
69	.172	.002
70	.166	.002
71	.171	.002
72	.183	.002
73	.174	.002
74	.184	.003
75	.175	.002
76	.170	.002
77	.171	.002
78	.177	.003
79	.179	.003
80	.177	.003
81	.178	.003
82	.189	.003
83	.194	.003
84	.187	.003
85	.177	.003
86	.184	.003
87	.195	.003
88	.186	.003
89	.202	.003
90	.197	.003
91	.176	.003
92	.181	.003
93	.190	.003
94	.179	.003
95	.169	.003
96	.160	.003
97	.155	.004
98	.141	.004
99	.163	.005
100	.166	.007

Notes: Entries are mean defection rates at each position in the vote order together with the associated standard errors.

Table A.2: Deviatings from the Party Line, Controlling for Bill-Party Fixed Effects

	A. Roll Call Specific Order				*B. Order Among All Senators*			
	Deviate							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Alphabetical Rank	-.172***	-.030***	-.051**	-.040**	-.221***	-.114*	-.098	-.059
	(.055)	(.011)	(.021)	(.019)	(.071)	(.065)	(.109)	(.102)
Controls:								
Senator Fixed Effects	Yes	No	No	No	Yes	No	No	No
Senator × Congress Fixed Effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Fraction of Copartisans Deviating	No	Yes	No	No	No	Yes	No	No
Bill Fixed Effects	No	No	Yes	No	No	No	Yes	No
Bill × Party Fixed Effects	No	No	No	Yes	No	No	No	Yes
R-Squared	.055	.241	.192	.279	.055	.241	.192	.279
Number of Observations	2,897,879	2,897,879	2,897,879	2,897,879	2,897,879	2,897,879	2,897,879	2,897,879

Notes: Entries are coefficients and standard errors from estimating equation (1) by ordinary least squares. Heteroskedasticity robust standard errors are clustered by Congress and reported in parentheses. As explained in the main text, Alphabetical Rank corresponds to senators' alphabetical percentile ranking among their colleagues. The four left-most columns are based on senators' alphabetical rank among those who participated in a given roll call, whereas the four right-most columns construct rank based on the entire set of senators who served in Congress at the time a given roll call was held. ***, **, and * denote statistical significance at the 1%-, 5%-, and 10%-levels, respectively.

Table A.3: Deviations from the Party Line Controlling for (Un)Certainty in Outcomes, Close Roll Calls

	A. Roll Call Specific Order			*B. Order Among All Senators*		
	Deviate					
	(1)	(2)	(3)	(4)	(5)	(6)
Alphabetical Rank	-.316***	-.240***	-.057***	-.385***	-.686*	-.095
	(.091)	(.064)	(.015)	(.126)	(.359)	(.091)
Outcome Already Determined	.014**	.002	.036***	.013**	.002	.036***
	(.005)	(.005)	(.004)	(.005)	(.005)	(.004)
Controls:						
Senator Fixed Effects	Yes	No	No	Yes	No	No
Senator × Congress Fixed Effects	No	Yes	Yes	No	Yes	Yes
Fraction of Copartisans Deviating	No	No	Yes	No	No	Yes
R-Squared	.114	.152	.278	.114	.152	.276
Number of Observations	1,400,492	1,400,492	1,400,492	1,400,492	1,400,492	1,400,492

Notes: Entries are coefficients and standard errors from estimating equation (1), controlling for whether the outcome of the roll call had already been determined by the time a particular senator voted. Heteroskedasticity robust standard errors are clustered by Congress and reported in parentheses. As explained in the main text, "Alphabetical Rank" corresponds to senators' alphabetical percentile ranking among their colleagues. "Outcome Already Determined" is an indicator variable equal to one if and only if the number of yeas (nays) required for passage (failure) of the bill had already been reached when it was a particular senator's turn to vote (based on her alphabetical rank). The three left-most columns are based on senators' alphabetical rank among those who participated in a given roll call, whereas the three right-most columns construct rank based on the entire set of senators who served in Congress at the time a given roll call was held. ***, **, and * denote statistical significance at the 1%-, 5%-, and 10%-levels, respectively.

Table A.4: Estimated Order Effects for Alternative Definitions of the Party Line

	λ		
	Roll Call Specific Order	Order Among All Senators	*N*
Based on Majority of Fellow Party Members	-.172***	-.221***	2,897,879
	(.055)	(.071)	
Based on Vote of Party Leader	-.367***	-.422***	2,176,089
	(.087)	(.108)	
Based on Vote of Party Whip	-.237**	-.276**	2,251,000
	(.095)	(.112)	

Notes: Entries in the center columns are point estimates and standard errors for λ, given different definitions of the party line. The respective definition is indicated on the left of each row. Estimates are based on equation (1) and either the roll call–specific order or the order of all senators who served in the chamber when a given vote was held. Heteroskedasticity robust standard errors are clustered by Congress and reported in parentheses. The column on the right shows the number of valid observations associated with each definition. Sample sizes vary because parties did not adopt today's leadership system until the early twentieth century. ***, **, and * denote statistical significance at the 1%-, 5%-, and 10%-levels, respectively.

Table A.5: Estimated Order Effects for Roll Calls with Little Abstention

	λ: Roll Call Specific Order	λ: Order Among All Senators	*N*
All Roll Calls	-.172***	-.221***	2,897,879
	(.055)	(.071)	
10% Abstention or Less	-.193**	-.284***	2,123,525
	(.075)	(.093)	
5% Abstention or Less	-.295***	-.395***	1,834,738
	(.087)	(.101)	
1% Abstention or Less	-.416	-.435	205,258
	(.237)	(.249)	

Notes: Entries in the center columns are point estimates and standard errors for λ, by level of abstention. Estimates are based on equation (1) and either the roll call–specific order or the order of all senators who served in the chamber when a given vote was held. Heteroskedasticity robust standard errors are clustered by Congress and reported in parentheses. The column on the right shows the number of valid observations within each subsample. ***, **, and * denote statistical significance at the 1%-, 5%-, and 10%-levels, respectively.

Table A.6: Deviations from the Party Line as a Function of Alphabetical Rank among All Members of the Chamber

A. Senate

	Deviate					
	(1)	(2)	(3)	(4)	(5)	(6)
Alphabetical Rank	-.221***	-.511**	-.102*	-.362*	-.382***	-.685*
	(.071)	(.255)	(.054)	(.200)	(.126)	(.359)
Sample	All Roll Calls	All Roll Calls	Lopsided Roll Calls	Lopsided Roll Calls	Close Roll Calls	Close Roll Calls
Controls:						
Senator Fixed Effects	Yes	No	Yes	No	Yes	No
Senator × Congress Fixed Effects	No	Yes	No	Yes	No	Yes
R-Squared	.055	.074	.036	.056	.114	.152
Number of Observations	2,897,879	2,897,879	1,497,387	1,497,387	1,400,492	1,400,492

B. House of Representatives

	Deviate					
	(7)	(8)	(9)	(10)	(11)	(12)
Alphabetical Rank	-.290	.298	-.408	-.352	-.104	.969
	(.205)	(.929)	(.248)	(1.439)	(.262)	(.975)
Sample	All Roll Calls	All Roll Calls	Lopsided Roll Calls	Lopsided Roll Calls	Close Roll Calls	Close Roll Calls
Controls:						
Senator Fixed Effects	Yes	No	Yes	No	Yes	No
Senator × Congress Fixed Effects	No	Yes	No	Yes	No	Yes
R-Squared	.047	.060	.029	.047	.144	.164
Number of Observations	9,618,470	9,618,470	5,230,871	5,230,871	4,387,599	4,387,599

Notes: Entries are coefficients and standard errors from estimating equation (1) by ordinary least squares. The upper panel does so for the U.S. Senate, while the entries in the lower panel refer to the House of Representatives after the introduction of electronic voting machines. Estimates are based on legislators' alphabetical rank among everyone who served in the chamber at the time the vote was held. Heteroskedasticity robust standard errors are clustered by Congress and reported in parentheses. As explained in the main text, roll calls are classified as "close" or "lopsided" according to the cutoffs in Snyder and Groseclose (2000). ***, **, and * denote statistical significance at the 1%-, 5%-, and 10%-levels, respectively.

Table A.7: Reduced-Form Comparative Statics

	λ: Roll Call Specific Order	λ: Order Among All Senators	*N*
Baseline	-.172***	-.221***	2,897,879
	(.055)	(.071)	
Split Minority	.013	.041	1,414,507
	(.046)	(.065)	
By Majority Party's Seat Advanatge:			
1 or 2 Seats	.075	.216	251,203
	(.079)	(.180)	
3 to 5 Seats	-.115	-.113	202,566
	(.108)	(.137)	
6 to 10 Seats	-.180*	-.280*	877,082
	(.104)	(.141)	
11 to 20 Seats	-.182	-.372***	720,202
	(.114)	(.128)	
> 20 Seats	.007	.094	772,785
	(.078)	(.120)	

Notes: Entries in the center columns are point estimates and standard errors for λ in different subsamples of the data. The respective restriction is indicated on the left of each row. Estimates are based on equation (1) and either the roll call–specific order or the order of all senators who served in the chamber when a given vote was held. Heteroskedasticity robust standard errors are clustered by Congress and reported in parentheses. The column on the right lists the number of observations in each subsample. ***, **, and * denote statistical significance at the 1%-, 5%-, and 10%-levels, respectively.

Table A.8: Order Effects, by Senators' Characteristics

	A. Roll Call Specific Order		*B. Order Among All Senators*	
	λ	F-Test [*p* -value]	λ	F-Test [*p* -value]
Baseline	-.172***	--	-.221***	--
	(.055)		(.071)	
By Party Memebrship:				
Majority Party	-.180***	.006	-.231***	.007
	(.054)		(.072)	
Minority Party	-.163***		-.214***	
	(.056)		(.073)	
By Age:				
< 50 Years	-.161***	.987	-.206***	.978
	(.053)		(.069)	
50 to 65 Years	-.160***		-.204***	
	(.053)		(.069)	
> 65 Years	-.160***		-.204***	
	(.055)		(.071)	
By Gender:				
Male	-.171***	.620	-.220***	.702
	(.055)		(.072)	
Female	-.235*		-.277*	
	(.133)		(.151)	
By Educational Attainment:				
Less than College	-.267***	.098	-.340***	.135
	(.077)		(.107)	
College Educated	-.139**		-.180**	
	(.059)		(.075)	
By Closeness to Next Election:				
End of Term	-.179***	.117	-.229***	.109
	(.056)		(.072)	
Not End of Term	-.170***		-.220***	
	(.054)		(.071)	

Notes: Entries are point estimates and standard errors for λ in different subsamples of the data. The respective restriction is indicated on the left of each row. Heteroskedasticity robust standard errors are clustered by Congress and reported in parentheses. Within each panel, the rightmost column displays *p* -values from an F-test for equality of coefficients.

Table A.9: Analysis of Bills (Not) Mentioned in CQ Weekly, 91st–111th Congress

A. Roll Call Specific Order

	Deviate			
	(1)	(2)	(3)	(4)
Alphabetical Rank	-.385***	-.381***	-.365***	-.399***
	(.104)	(.127)	(.097)	(.114)
Senator Fixed Effects	Yes	No	No	No
Sample	All Roll Calls	Bills *Not* Covered in CQ Almanac	Bills Covered in CQ Almanac	Coverage Above Sample Median
R-Squared	.050	.060	.048	.054
Number of Observations	1,530,770	358,698	1,172,072	573,259

B. Order Among All Senators

	Deviate			
	(5)	(6)	(7)	(8)
Alphabetical Rank	-.515***	-.566**	-.478**	-.520***
	(.109)	(.153)	(.100)	(.121)
Senator Fixed Effects	Yes	No	No	No
Sample	All Roll Calls	Bills *Not* Covered in CQ Almanac	Bills Covered in CQ Almanac	Coverage Above Sample Median
R-Squared	.050	.060	.048	.055
Number of Observations	1,530,770	358,698	1,172,072	573,259

Notes: Entries are coefficients and standard errors from estimating equation (1) by ordinary least squares, restricting attention to roll calls held in the 91st to 111th Congresses. The leftmost column considers all roll calls during this period, whereas the remaining three columns differentiate between rolls calls that received no mention in the Congressional Quarterly Almanac (cols. 2 and 6), bills that were mentioned (cols. 3 and 7), and bills for which the number of lines in the Almanac exceeded the Congress-specific sample median (cols. 4 and 8), i.e., bills that were covered particularly intensely. Heteroskedasticity robust standard errors are clustered by Congress and reported in parentheses. As explained in the text, Alphabetical Rank corresponds to senators' alphabetical percentile ranking among their colleagues. The regressions in the upper panel are based on senators' alphabetical rank among those who participated in a given roll call, whereas the ones in the lower panel construct rank based on the entire set of senators who served in Congress at the time a given roll call was held. ***, **, and * denote statistical significance at the 1%-, 5%-, and 10%-levels, respectively.

Table A.10: Monte Carlo Rejection Rates

A. 500 Roll Calls

	Herding Effect (θ)			
Valence (α)	0	.05σ	.1σ	.15σ
0	.055	.255	.657	.894
.1σ	.050	.272	.586	.895
.2σ	.061	.209	.484	.750
.3σ	.052	.150	.343	.526

B. 1,000 Roll Calls

	Herding Effect (θ)			
Valence (α)	0	.05σ	.1σ	.15σ
0	.047	.406	.877	.998
.1σ	.044	.384	.830	.994
.2σ	.047	.311	.710	.966
.3σ	.041	.232	.536	.817

C. 10,000 Roll Calls

	Herding Effect (θ)			
Valence (α)	0	.05σ	.1σ	.15σ
0	.042	.996	1	1
.1σ	.049	.994	1	1
.2σ	.036	.977	1	1
.3σ	.050	.896	1	1

Notes: Entries denote the fraction of Monte Carlo simulations for which the null hypothesis of no herding is rejected at the 5%-significance level. Each entry is based on a 1,000 simulations. For details on the data generating process and the estimation procedure, see Appendix H.

Table A.11: Do Senators Learn from Observing the Votes of the Leadership?

	Deviate from Choice of Leader	Deviate from Choice of Whip
	(1)	(2)
Rank Before Leader	.210***	
	(.006)	
Rank After Leader	.211***	
	(.006)	
Rank Before Whip		.222***
		(.005)
Rank After Whip		.214***
		(.006)
H_0: After = Before [*p* -value]	.948	.327
R-Squared	.047	.044
Number of Observations	2,039,313	2,111,534

Notes: Entries are coefficients and standard errors from regressing an indicator whether a senator votes the same ways as either the party leader or whip on indicators for their relative positions in the vote order and senator fixed effects. Heteroskedasticity robust standard errors are clustered by Congress and reported in parentheses. ***, **, and * denote statistical significance at the 1%-, 5%-, and 10%-levels, respectively.

Table A.12: Do Senators React to Opponents Who Come Later in the Vote Order?

A. Based On Opponents' Actual Behavior

	Deviate					
	(1)	(2)	(3)	(4)	(5)	(6)
Opponent Votes Next × Deviates from	.042***	.008***	.006***	.039***	.008***	.006***
His Party's Position	(.004)	(.002)	(.002)	(.004)	(.002)	(.002)
Opponent Votes Two Positions Removed				.035***	.004***	.003*
× Deviates from His Party's Position				(.004)	(.001)	(.001)
Controls:						
Senator Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Fraction of Copartisans Deviating	No	Yes	No	No	Yes	No
Bill × Party Fixed Effects	No	No	Yes	No	No	Yes
R-Squared	.056	.228	.268	.057	.228	.268
Number of Observations	2,859,570	2,859,570	2,859,570	2,821,537	2,821,537	2,821,537

B. Based On Opponents' Predicted Behavior

	Deviate					
	(1)	(2)	(3)	(4)	(5)	(6)
Opponent Votes Next × Pred. Prob. He	.104***	.020***	.016***	.091***	.020***	.017***
Deviates from His Party's Position	(.008)	(.003)	(.004)	(.008)	(.004)	(.004)
Opponent Votes Two Positions Removed ×				.075***	.007***	.006**
Pred. Prob. He Deviates from His Party's Position				(.007)	(.002)	(.003)
Controls:						
Senator Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Fraction of Copartisans Deviating	No	Yes	No	No	Yes	No
Bill × Party Fixed Effects	No	No	Yes	No	No	Yes
R-Squared	.057	.228	.268	.059	.228	.268
Number of Observations	2,859,570	2,859,570	2,859,570	2,821,537	2,821,537	2,821,537

C. Based On Both

	Deviate					
	(1)	(2)	(3)	(4)	(5)	(6)
Opponent Votes Next × Deviates from	.002	.000	.000	.001	.000	.000
His Party's Position	(.002)	(.001)	(.001)	(.002)	(.001)	(.001)
Opponent Votes Next × Pred. Prob. He	.103***	.020***	.015***	.090***	.020***	.017***
Deviates from His Party's Position	(.009)	(.004)	(.004)	(.008)	(.004)	(.004)
Opponent Votes Two Positions Removed				.003	.001	.001
× Deviates from His Party's Position				(.002)	(.002)	(.002)
Opponent Votes Two Positions Removed ×				.072***	.006**	.004*
Pred. Prob. He Deviates from His Party's Position				(.072)	(.002)	(.003)
Controls:						
Senator Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Fraction of Copartisans Deviating	No	Yes	No	No	Yes	No
Bill × Party Fixed Effects	No	No	Yes	No	No	Yes
R-Squared	.057	.228	.268	.059	.228	.268
Number of Observations	2,859,570	2,859,570	2,859,570	2,821,537	2,821,537	2,821,537

Notes: Entries in the three leftmost columns are coefficients and standard errors from regression an indicator for whether a senator deviates from the party line on an indicator for whether someone from the opposing party comes next in the vote order and the same indicator interacted with with whether that other senator defects (upper panel) or is predicted to defect (middle panel) from his party's position. The "predicted probability of defection" is that implied by a senator's DW-Nominate score (Poole and Rosenthal 1997). Entries in the three rightmost columns additionally consider opponents who vote two positions after the senator in question. The lower panel runs a horse race between actual and predicted defection. All estimates are based on legislators' alphabetical rank among everyone who served in the chamber at the time the vote was held. Heteroskedasticity robust standard errors are clustered by Congress and reported in parentheses. ***, **, and * denote statistical significance at the 1%-, 5%-, and 10%-levels, respectively.

Table A.13: Preference Configurations and Deviations with S=5

Preference Configuration					Equilibrium Votes					Deviation				
S1	S2	S3	S4	S5	S1	S2	S3	S4	S5	S1	S2	S3	S4	S5
Y	Y	Ỹ	Ỹ	Ỹ	y	y	n	n	y	0	0	1	1	0
Y	Y	Ỹ	Ỹ	Ñ	y	y	n	y	y	0	0	1	0	1
Y	Y	Ỹ	Ỹ	N	y	y	n	y	n	0	0	1	0	0
Y	Y	Ỹ	Ñ	Ỹ	y	y	n	y	n	0	0	1	1	1
Y	Y	Ỹ	N	Ỹ	y	y	n	n	y	0	0	1	0	0
Y	Y	Ñ	Ỹ	Ỹ	y	y	y	n	n	0	0	1	1	1
Y	Y	N	Ỹ	Ỹ	y	y	n	n	y	0	0	0	1	0
Y	Ỹ	Y	Ỹ	Ỹ	y	n	y	n	y	0	1	0	1	0
Y	Ỹ	Y	Ỹ	Ñ	y	n	y	y	y	0	1	0	0	1
Y	Ỹ	Y	Ỹ	N	y	n	y	y	n	0	1	0	0	0
Y	Ỹ	Y	Ñ	Ỹ	y	n	y	y	n	0	1	0	1	1
Y	Ỹ	Y	N	Ỹ	y	n	y	n	y	0	1	0	0	0
Y	Ỹ	Ỹ	Y	Ỹ	y	n	n	y	y	0	1	1	0	0
Y	Ỹ	Ỹ	Y	Ñ	y	n	y	y	y	0	1	0	0	1
Y	Ỹ	Ỹ	Y	N	y	n	y	y	n	0	1	0	0	0
Y	Ỹ	Ỹ	Ỹ	Y	y	n	n	y	y	0	1	1	0	0
Y	Ỹ	Ỹ	Ỹ	Ỹ	y	n	n	y	y	0	1	1	0	0
Y	Ỹ	Ỹ	Ỹ	Ñ	y	n	y	y	y	0	1	0	0	1
Y	Ỹ	Ỹ	Ỹ	N	y	n	y	y	n	0	1	0	0	0
Y	Ỹ	Ỹ	Ñ	Y	y	n	y	y	y	0	1	0	1	0
Y	Ỹ	Ỹ	Ñ	Ỹ	y	n	y	y	n	0	1	0	1	1
Y	Ỹ	Ỹ	N	Y	y	n	y	n	y	0	1	0	0	0
Y	Ỹ	Ỹ	N	Ỹ	y	n	y	n	y	0	1	0	0	0
Y	Ỹ	Ñ	Y	Ỹ	y	n	y	y	n	0	1	1	0	1
Y	Ỹ	Ñ	Ỹ	Y	y	n	y	n	y	0	1	1	1	0
Y	Ỹ	Ñ	Ỹ	Ỹ	y	n	y	n	y	0	1	1	1	0
Y	Ỹ	N	Y	Ỹ	y	n	n	y	y	0	1	0	0	0
Y	Ỹ	N	Ỹ	Y	y	n	n	y	y	0	1	0	0	0
Y	Ỹ	N	Ỹ	Ỹ	y	n	n	y	y	0	1	0	0	0
Y	Ñ	Y	Ỹ	Ỹ	y	y	y	n	n	0	1	0	1	1
Y	Ñ	Ỹ	Y	Ỹ	y	y	n	y	n	0	1	1	0	1
Y	Ñ	Ỹ	Ỹ	Y	y	y	n	n	y	0	1	1	1	0
Y	Ñ	Ỹ	Ỹ	Ỹ	y	y	n	n	y	0	1	1	1	0
Y	N	Y	Ỹ	Ỹ	y	n	y	n	y	0	0	0	1	0
Y	N	Ỹ	Y	Ỹ	y	n	n	y	y	0	0	1	0	0
Y	N	Ỹ	Ỹ	Y	y	n	n	y	y	0	0	1	0	0
Y	N	Ỹ	Ỹ	Ỹ	y	n	n	y	y	0	0	1	0	0
Ỹ	Y	Y	Ỹ	Ỹ	n	y	y	n	y	1	0	0	1	0
Ỹ	Y	Y	Ỹ	Ñ	n	y	y	y	y	1	0	0	0	1
Ỹ	Y	Y	Ỹ	N	n	y	y	y	n	1	0	0	0	0
Ỹ	Y	Y	Ñ	Ỹ	n	y	y	y	n	1	0	0	1	1
Ỹ	Y	Y	N	Ỹ	n	y	y	n	y	1	0	0	0	0
Ỹ	Y	Ỹ	Y	Ỹ	n	y	n	y	y	1	0	1	0	0
Ỹ	Y	Ỹ	Y	Ñ	n	y	y	y	y	1	0	0	0	1
Ỹ	Y	Ỹ	Y	N	n	y	y	y	n	1	0	0	0	0
Ỹ	Y	Ỹ	Ỹ	Y	n	y	n	y	y	1	0	1	0	0
Ỹ	Y	Ỹ	Ỹ	Ỹ	n	y	n	y	y	1	0	1	0	0
Ỹ	Y	Ỹ	Ỹ	Ñ	n	y	y	y	y	1	0	0	0	1
Ỹ	Y	Ỹ	Ỹ	N	n	y	y	y	n	1	0	0	0	0
Ỹ	Y	Ỹ	Ñ	Y	n	y	y	y	y	1	0	0	1	0
Ỹ	Y	Ỹ	Ñ	Ỹ	n	y	y	y	n	1	0	0	1	1
Ỹ	Y	Ỹ	N	Y	n	y	y	n	y	1	0	0	0	0
Ỹ	Y	Ỹ	N	Ỹ	n	y	y	n	y	1	0	0	0	0
Ỹ	Y	Ñ	Y	Ỹ	n	y	y	y	n	1	0	1	0	1
Ỹ	Y	Ñ	Ỹ	Y	n	y	y	n	y	1	0	1	1	0
Ỹ	Y	Ñ	Ỹ	Ỹ	n	y	y	n	y	1	0	1	1	0
Ỹ	Y	N	Y	Ỹ	n	y	n	y	y	1	0	0	0	0
Ỹ	Y	N	Ỹ	Y	n	y	n	y	y	1	0	0	0	0
Ỹ	Y	N	Ỹ	Ỹ	n	y	n	y	y	1	0	0	0	0
Ỹ	Ỹ	Y	Y	Ỹ	n	n	y	y	y	1	1	0	0	0
Ỹ	Ỹ	Y	Y	Ñ	n	y	y	y	y	1	0	0	0	1
Ỹ	Ỹ	Y	Y	N	n	y	y	y	n	1	0	0	0	0
Ỹ	Ỹ	Y	Ỹ	Y	n	n	y	y	y	1	1	0	0	0
Ỹ	Ỹ	Y	Ỹ	Ỹ	n	n	y	y	y	1	1	0	0	0
Ỹ	Ỹ	Y	Ỹ	Ñ	n	y	y	y	y	1	0	0	0	1
Ỹ	Ỹ	Y	Ỹ	N	n	y	y	y	n	1	0	0	0	0

Table A.13 continued

Ỹ	Ỹ	Y	Ñ	Y	n	y	y	y	y	1	0	0	1	0
Ỹ	Ỹ	Y	Ñ	Ỹ	n	y	y	y	n	1	0	0	1	1
Ỹ	Ỹ	Y	N	Y	n	y	y	n	y	1	0	0	0	0
Ỹ	Ỹ	Y	N	Ỹ	n	y	y	n	y	1	0	0	0	0
Ỹ	Ỹ	Ỹ	Y	Y	n	n	y	y	y	1	1	0	0	0
Ỹ	Ỹ	Ỹ	Y	Ỹ	n	n	y	y	y	1	1	0	0	0
Ỹ	Ỹ	Ỹ	Y	Ñ	n	y	y	y	y	1	0	0	0	1
Ỹ	Ỹ	Ỹ	Y	N	n	y	y	y	n	1	0	0	0	0
Ỹ	Ỹ	Ỹ	Ỹ	Y	n	n	y	y	y	1	1	0	0	0
Ỹ	Ỹ	Ỹ	Ỹ	Ỹ	n	n	y	y	y	1	1	0	0	0
Ỹ	Ỹ	Ỹ	Ỹ	Ñ	n	y	y	y	y	1	0	0	0	1
Ỹ	Ỹ	Ỹ	Ỹ	N	n	y	y	y	n	1	0	0	0	0
Ỹ	Ỹ	Ỹ	Ñ	Y	n	y	y	y	y	1	0	0	1	0
Ỹ	Ỹ	Ỹ	Ñ	Ỹ	n	y	y	y	n	1	0	0	1	1
Ỹ	Ỹ	Ỹ	N	Y	n	y	y	n	y	1	0	0	0	0
Ỹ	Ỹ	Ỹ	N	Ỹ	n	y	y	n	y	1	0	0	0	0
Ỹ	Ỹ	Ñ	Y	Y	n	y	y	y	y	1	0	1	0	0
Ỹ	Ỹ	Ñ	Y	Ỹ	n	y	y	y	n	1	0	1	0	1
Ỹ	Ỹ	Ñ	Ỹ	Y	n	y	y	n	y	1	0	1	1	0
Ỹ	Ỹ	Ñ	Ỹ	Ỹ	n	y	y	n	y	1	0	1	1	0
Ỹ	Ỹ	N	Y	Y	n	y	n	y	y	1	0	0	0	0
Ỹ	Ỹ	N	Y	Ỹ	n	y	n	y	y	1	0	0	0	0
Ỹ	Ỹ	N	Ỹ	Y	n	y	n	y	y	1	0	0	0	0
Ỹ	Ỹ	N	Ỹ	Ỹ	n	y	n	y	y	1	0	0	0	0
Ỹ	Ñ	Y	Y	Ỹ	n	y	y	y	n	1	1	0	0	1
Ỹ	Ñ	Y	Ỹ	Y	n	y	y	n	y	1	1	0	1	0
Ỹ	Ñ	Y	Ỹ	Ỹ	n	y	y	n	y	1	1	0	1	0
Ỹ	Ñ	Ỹ	Y	Y	n	y	n	y	y	1	1	1	0	0
Ỹ	Ñ	Ỹ	Y	Ỹ	n	y	n	y	y	1	1	1	0	0
Ỹ	Ñ	Ỹ	Ỹ	Y	n	y	n	y	y	1	1	1	0	0
Ỹ	Ñ	Ỹ	Ỹ	Ỹ	n	y	n	y	y	1	1	1	0	0
Ỹ	N	Y	Y	Ỹ	n	n	y	y	y	1	0	0	0	0
Ỹ	N	Y	Ỹ	Y	n	n	y	y	y	1	0	0	0	0
Ỹ	N	Y	Ỹ	Ỹ	n	n	y	y	y	1	0	0	0	0
Ỹ	N	Ỹ	Y	Y	n	n	y	y	y	1	0	0	0	0
Ỹ	N	Ỹ	Y	Ỹ	n	n	y	y	y	1	0	0	0	0
Ỹ	N	Ỹ	Ỹ	Y	n	n	y	y	y	1	0	0	0	0
Ỹ	N	Ỹ	Ỹ	Ỹ	n	n	y	y	y	1	0	0	0	0
Ñ	Y	Y	Ỹ	Ỹ	y	y	y	n	n	1	0	0	1	1
Ñ	Y	Ỹ	Y	Ỹ	y	y	n	y	n	1	0	1	0	1
Ñ	Y	Ỹ	Ỹ	Y	y	y	n	n	y	1	0	1	1	0
Ñ	Y	Ỹ	Ỹ	Ỹ	y	y	n	n	y	1	0	1	1	0
Ñ	Ỹ	Y	Y	Ỹ	y	n	y	y	n	1	1	0	0	1
Ñ	Ỹ	Y	Ỹ	Y	y	n	y	n	y	1	1	0	1	0
Ñ	Ỹ	Y	Ỹ	Ỹ	y	n	y	n	y	1	1	0	1	0
Ñ	Ỹ	Ỹ	Y	Y	y	n	n	y	y	1	1	1	0	0
Ñ	Ỹ	Ỹ	Y	Ỹ	y	n	n	y	y	1	1	1	0	0
Ñ	Ỹ	Ỹ	Ỹ	Y	y	n	n	y	y	1	1	1	0	0
Ñ	Ỹ	Ỹ	Ỹ	Ỹ	y	n	n	y	y	1	1	1	0	0
N	Y	Y	Ỹ	Ỹ	n	y	y	n	y	0	0	0	1	0
N	Y	Ỹ	Y	Ỹ	n	y	n	y	y	0	0	1	0	0
N	Y	Ỹ	Ỹ	Y	n	y	n	y	y	0	0	1	0	0
N	Y	Ỹ	Ỹ	Ỹ	n	y	n	y	y	0	0	1	0	0
N	Ỹ	Y	Y	Ỹ	n	n	y	y	y	0	1	0	0	0
N	Ỹ	Y	Ỹ	Y	n	n	y	y	y	0	1	0	0	0
N	Ỹ	Y	Ỹ	Ỹ	n	n	y	y	y	0	1	0	0	0
N	Ỹ	Ỹ	Y	Y	n	n	y	y	y	0	1	0	0	0
N	Ỹ	Ỹ	Y	Ỹ	n	n	y	y	y	0	1	0	0	0
N	Ỹ	Ỹ	Ỹ	Y	n	n	y	y	y	0	1	0	0	0
N	Ỹ	Ỹ	Ỹ	Ỹ	n	n	y	y	y	0	1	0	0	0
									Average Deviation Rate:	.619	.429	.333	.270	.222

Notes: The three left-most columns list all possible preference configurations for S =5 in which $|Y| < \bar{y} < |Y| + |\tilde{Y}|$. As in the main text, Y (N) indicates that a senator always vote "yea" ("nay"). $\tilde{Y}$ ($\tilde{N}$) indicates that a senator would vote "yea" ("nay") if and only if doing so would change the outcome of the roll call. The three columns in the middle show agents' equilibrium choices, with y (n) implying that a senator votes "yea" ("nay"). The three right-most columns indicate whether an agent deviates from her instrumental preferences. As explained in Appendix A.5, we say that a senator deviates if either $i \in \tilde{Y}$ and she votes "nay" or $i \in \tilde{N}$ and she says "yea."