LEGITIMIZING POLICY

Daniel L. Chen, Moti Michaeli, and Daniel Spiro[*]

Abstract: In many settings of political bargaining over policy, agents care not only about getting their will but also about having others approve the chosen policy thus giving it more weight. What is the effect on the bargaining outcome when agents care about such legitimacy of the policy? We study this question theoretically and empirically. We show that the median-voter theorem holds in groups that are ideologically very cohesive and in groups with extreme ideological disagreement. However, in groups with intermediate ideological disagreement, the median-voter theorem does not hold. This is since, on the individual level, ideological disagreement with the median has a non-monotonic effect on the policy. We test our model in a natural experimental setting—U.S. appeals courts—where causal identification is based on random assignment of judges into judicial panels, each consisting of three judges who rule on a case. Here judges care about legitimacy of the policy they write because a norm of consensus prevails and because increased legitimacy reduces the likelihood of the judicial case to be heard by the Supreme Court. The predicted pattern of how policies depend on the participants' ideologies are corroborated by our empirical tests.

Keywords: Legitimacy; group decision making; judicial decision making; bargaining; ideology.

JEL codes: D7, K0, Z1.

1 Introduction

We consider, theoretically and empirically, bargaining within groups of agents motivated by ideology. In particular, we study situations in which the choice of whether to support a policy affects its impact. Consider an agent with some ideological bliss point who needs to decide whether to support a policy. Suppose further that this decision will not directly affect whether the policy goes through, but rather its weight: if the agent legitimizes the policy by supporting it, the value of the policy is amplified. This can capture, for instance, that the success of the implementation is more likely if there are many signatories or that the population subject to the policy will take it more seriously. We ask three questions: 1) What policies would the agent support? 2) How does this affect the policies that are negotiated by the group? 3) Given this, who gets the right to propose the policy and does it give the proposer more influence?

We develop a model providing theoretical predictions for these questions. It is kept as simple as possible containing only the minimal necessary ingredients. The model is based on multiparty bargaining between three agents.[1] The agents have different ideologies. What they are set to decide is the ideology of a policy, i.e., a Hotelling-style model. The further away the policy is from an agent, the lower his/her payoff is (in line with models by Cardona and Ponsati, 2011; Diermeier and Merlo, 2000; Predtetchinski, 2011; Banks and Duggan, 2006; Duggan et al., 2000). This is important, since we don't want to limit ourselves to situations where the piece of the pie is fixed (as in Baron and Ferejohn, 1989) but instead to allow for circumstances where if one agent is better off, not all others are necessarily worse off. This could capture, for instance, decisions such as which tax rate to set, how strict to be on abortion, a country's position in foreign policy, where to place a job-generating industry etc. We use a very simple protocol of bargaining.[2] The agents first vote on who gets to propose. Then the chosen agent proposes. Finally, the non-proposers decide whether to sign

---

[1] In the previous literature this is often referred to as legislative bargaining. It was pioneered by Baron and Ferejohn (1989). See Eraslan and Evdokimov (2019) for a literature review.

[2] In particular, it is much simpler than the models by Cardona and Ponsati (2011); Diermeier and Merlo (2000); Predtetchinski (2011); Banks and Duggan (2006); Duggan et al. (2000).

the proposal.[3] The main innovations of our model compared to the previous literature are two seemingly small, but crucial, components that embed the idea of legitimizing a policy. First, an agent gets positive payoff if the policy is near her bliss point, but as the policy moves away from the agent, her payoff eventually becomes negative. One can think of this as capturing whether the agent thinks the policy does more good (if equaling or close to her bliss point) or more harm (if far from the bliss point). Second, payoffs are amplified when more agents support the policy. These two assumptions are jointly important. When a policy comes with a positive payoff, endorsing it (thus increasing its legitimacy) increases the gain from it. Conversely, endorsing a negative policy increases the *loss* attached to it.[4]

How ideology of "group" members affects outcomes is hard to study empirically due to unobservability of individual ideology and since most groups are formed endogenously (e.g., in a party system, the formation of the parties is endogenous). In the U.S. Courts of Appeals (U.S. Federal Circuit Courts) we find a setting where groups make ideologically contentious decisions (Epstein et al. 2013, Chen et al. 2019) and where the two problems are resolved. First, there exist commonly used and exogenous measures of individual ideology. Second, assignment to groups is random: for each judicial case, three judges are randomly assigned to sit together on a panel. Apart from the binary verdict, the panel composes an "opinion" motivating the verdict. The opinion forms a policy, in the sense that it guides how to rule on future cases in same-level and lower courts.[5]

Our theoretical predictions are supported by the empirical analysis. On question

---

[3]Our model bears resemblance to a much richer model by Cameron and Kornhauser (2009) that is applied specifically to collegial courts. In their model there are more agents (to reflect the size of the US Supreme Court), and beyond choosing policy from a continuum, the agents face a binary choice of verdict. Thus, many more things can happen in equilibrium. One of the main differences is that in their model agents do not care about the ideology of the policy unless they sign it, thus they do not have instrumental motives like in our model, where agents care about the ideology of the policy per se. In that sense, their model is similar to the expressive motives modeled in Chen et al. (2019). Furthermore, Cameron and Kornhauser (2009) do not test their predictions empirically. Despite our empirical setting being similar to the setting they are modeling, we choose to abstract away from many of their complexities to make the model parsimonious and applicable to other settings too.

[4]This aspect relates to Gratton et al. (2020) with the main difference being that in their model legitimacy is exogenous and refers to the leader's political capital. Their focus is on how the leader uses this legitimacy. Our paper is concerned with the step before, in endogenously determining the legitimacy of the chosen policy (or of the leader for that matter). But we abstract from how the legitimacy is used after that.

[5]Ideology (or personal preferences) has also been documented to play a role in other courts. See, for instance, Cohen et al. (2015) and Anwar et al. (2018) for recent evidence.

1, the model predicts that agents will be more inclined to support policies close to their own bliss point. In our empirical setting, this has already been established by previous research (Epstein et al., 2013; Chen et al., 2019). On question 3 our model predicts that the median agent will have the strongest, but not the sole, influence on policy. In many group constellations the median can play the other agents against each other and get her will fully. In other constellations, the median will need to compromise. The median's particular effect on policy has been corroborated in our empirical setting by previous research (see Cross, 2007; Chen et al., 2019 and, e.g., Ambrus et al. 2015 for a different setting). The novelty of the current paper, both empirically and theoretically, regards question 2 – how policy is affected by the quest for legitimacy.. Here our model predicts that the policy will deviate from the median's preference in the direction of another agent's preference when, roughly speaking, the group's level of ideological cohesion is intermediate. When all ideologies closely align (high cohesion), the median gets her will fully since all sign onto her preferred policy. When the group has strong ideological disagreement (low cohesion), there is no common ground for a policy to be signed by all, hence the median (endogenously having the most influence) gets the policy in line with her ideology but attains only few signatures thus low legitimacy. When the group is at an intermediate level of cohesion, essentially such that there exist policies that many would sign and such that compromise in return for legitimacy is in the interest of the median, the policy deviates from the median's bliss point. The empirical prediction generated by this pattern is that, as a function of, say, the "rightist" agent's ideological distance to the median, the policy is first independent and then increasing (and finally independent of it again). A similar pattern is predicted for the leftist agent's ideology. We test this prediction both non-parametrically (a local polynomial regression) and using structural-break tests (to find a shift in the policy from first being independent of a non-median ideology to later being affected by it). The results align with our theoretical predictions.

2 Model

Each agent has an ideology score t which is public information. The *pool* of agents consists of a continuum of infinitesimal agents with a unidimensional continuous distribution of ideology scores $F(t)$ over a range T. For each case, three agents (called L, M and R, where ideology goes from lowest to highest score) are randomly and independently drawn from the pool of agents to sit together in a *group*. With some abuse of notation we let $L \leq M \leq R$ also denote the ideology score of agents.

The timing of the actions within a group is as follows. First, agents vote on who will get to propose a policy. Second, the proposer proposes a policy $v \in \mathbb{R}$. Third, the two other agents decide simultaneously whether to sign the policy or not. The proposer automatically signs the policy upon proposing it.

An agent cares about the ideology of the policy and about its *legitimacy*, as reflected by the number of agents signing it. To capture this, we let the payoff of agent t be given by

$$
(1) \qquad u(t,v) = \begin{cases} [c - |t - v|]\,\lambda^{n-2} & \text{if } n \geq 2 \\[2mm] -K & \text{if } n < 2 \end{cases} \quad ,
$$

where $n \in \{1, 2, 3\}$ is the number of agents signing policy v and $\lambda > 1$ is the legitimacy attributed to the policy. The payoffs are depicted in Figure 1. On top of preferring to minimize $|t - v|$, an agent wishes to increase the legitimacy of policies she is happy with ($|t - v| \leq c$) and decrease the legitimacy of policies she is unhappy with ($|t - v| > c$). The legitimacy attributed to opinion v will be denoted by $\lambda(v)$. The term $-K$ represents the outside option in case no agreement is reached. We assume here that the panel has to reach majority agreement, hence that $-K$ is sufficiently negative so that it is lower than the lowest possible payoff under agreement (like in many legislative bargaining models such as Cardona and Ponsati, 2011; Predtetchinski, 2011).

5

FIGURE 1.— Payoffs
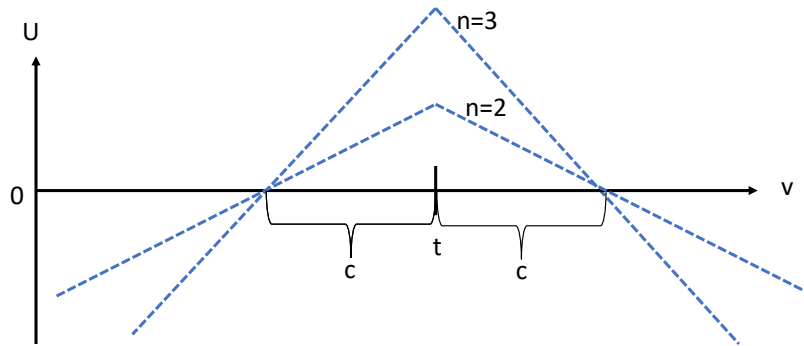
For convenience in presentation we adopt the convention that in the voting stage, in case an agent is indifferent between letting M or some other agent propose, she votes for M, and that if an agent is indifferent between signing or not, she signs. We solve the game with backward induction.

2.1 Period 3: When does an agent sign?

From the payoff function it immediately follows that, in a subgame starting in period 3, an agent will sign policies that are sufficiently close to her bliss point ($|t - v| \leq c$). If the policy is far ($|t - v| > c$), the agent will sign it only if precisely one other agent is signing in the equilibrium of the subgame (since not signing would yield $-K$). Thus trivially follows Lemma 1 (all the formal proofs can be found in the appendix).

LEMMA 1   *The best response of an agent t in the third period is to sign a policy if $|t - v| \leq c$ or if only one other agent is signing. In a pure-strategy SPE, 1) two agents sign if, given v, $max_{t \in \{L,M,R\}} |t - v| > c$; 2) three agents sign if $|t - v| \leq c$ for all agents.*

The empirical prediction following this lemma is simple: agents are more likely to sign and endorse policies that are closer to their ideal points.[6]

---

[6]This is even more likely to be the case if we refine the equilibrium concept to be correlated, i.e., in case the two non-proposing agents have $|t - v| > c$, the one with the larger $|t - v|$ is not signing (which is sensible). This refinement is not necessary for any of our results.

## 2.2 Period 2: Which policies do agents propose?

Which policy an agent would propose, given the chance, depends on the number of signatures she expects to get. Following Lemma 1, the proposer can ensure at least two signatures for any policy. Of course, if only two signatures are attainable, proposing $v = t$ is the best option and yields payoff $u(t, v) = c$. Improving upon this payoff can only be done by ensuring more signatures. The proposer will be willing to propose $v \neq t$ if and only if such a proposal yields

(2) $\qquad [c - |t - v|]\lambda \geq c \Leftrightarrow |t - v| \leq c - c/\lambda.$

Denote the set of policies that fulfill this inequality when the proposer is t by V_t, and denote by V_3 the set of proposals that all three agents would sign, i.e. $v \in V_3$ if and only if $max_{t \in \{L,M,R\}} |t - v| \leq c.$

LEMMA 2 *A proposer t will propose the $v \in V_t \cap V_3$ that minimizes $|t - v|$. If $V_t \cap V_3 = \emptyset$, t will propose $v = t$ and only two agents will sign.*

FIGURE 2.— Compromise range
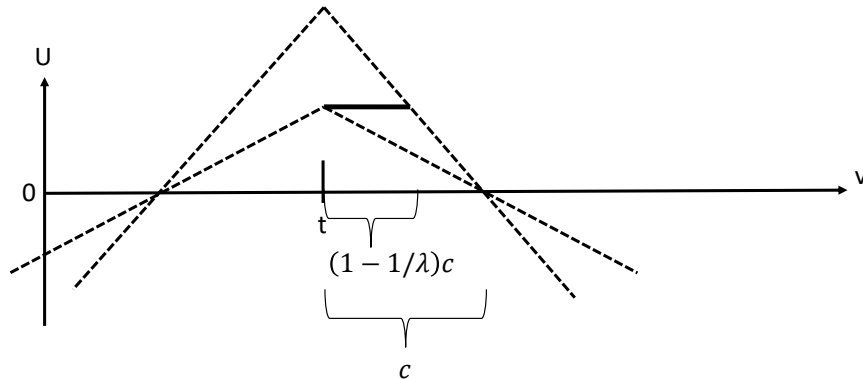
Comparing the range of policies over which the proposer is willing to compromise with the range over which other agents are willing to sign (Lemma 1 and Figure 2), we get

a first notion of the bargaining power of the proposer: the proposer is willing to deviate at most $c - c/\lambda$ from the bliss point while non-proposers are willing to sign policies at a distance c. From these ranges of compromise also follows that the proposed policy will depend on the combination of bliss points in the group, as will be outlined in the next two subsections. The results also depend on who the proposer is. There are two qualitatively different cases: when M proposes and when someone else proposes.

2.2.1 Median proposer

Let v_t^* denote the equilibrium proposal of agent t. The following proposition outlines the different cases under a median proposer.

PROPOSITION 1   *Suppose M is the proposer. Then:*

1. If $R - M \leq c$ and $M - L \leq c$ then $v_M^* = M$ and all agents sign.
2. If $c < R - M \leq 2c - c/\lambda$ and $R - L \leq 2c$ then $v_M^* = R - c$ and all agents sign.
3. If $c < M - L \leq 2c - c/\lambda$ and $R - L \leq 2c$ then $v_M^* = L + c$ and all agents sign.
4. Otherwise, $v_M^* = M$ and only two agents sign.

The results of the proposition are shown graphically in Figure 3. The intuition behind it is as follows. The median, being the proposer, has bargaining power. She knows she will attain at least one more signature, except for her own. The question then is which compromise is needed to attain the last signature and whether it is worth it. There are essentially three main cases.

FIGURE 3.— Parameter space, M as proposer
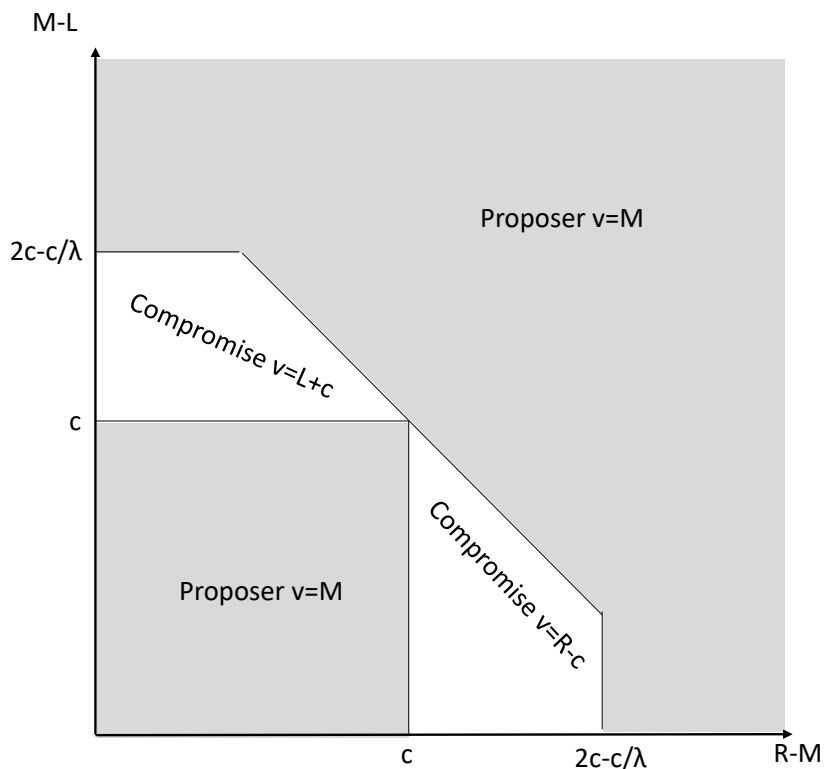
Either both the other agents are close enough to be willing to sign the median's bliss point. In this case, represented by point 1 in the proposition, the median does not need to compromise. This is the grey box on the lower left in Figure 3. Alternatively, the other two agents are at an intermediate distance from the median and each other, represented by points 2 and 3 in the proposition. Here one of them is not close enough for both to sign the median's bliss point. The median then has to compromise to achieve a unanimously endorsed policy. How much to compromise? Just enough to make the further-away agent sign (hence $v_M^* = R-c$ if this agent is R and $v_M^* = L+c$ if this agent is L) without losing the signature of

9

the other agent (hence the condition $R - L \leq 2c$ in points 2 and 3).[7] This case is represented by the white regions in Figure 3. Finally, if three signatures are not attainable, or the cost of getting the third signature is too high for M, then M just settles for two signatures by proposing her bliss point. This is represented by the grey region in the upper right in Figure 3.

Towards presenting the result that we will use for empirically testing our model, we present the following corollary that follows from Proposition 1 (we pose the corollary in terms of how v_M^* depends on R, equivalent statements can be made on how it depends on L).

COROLLARY 1 v_M^* *as a function of R has the following properties:*

1. If $M - L > 2c - c/\lambda$, $v_M^* = M$ is constant and independent of R.

2. If $c < M - L \leq 2c - c/\lambda$, $v_M^* = \begin{cases} L + c & \text{for } R \leq L + 2c \\ M & \text{for } R > L + 2c \end{cases}$, i.e., $v_M^*(R)$ is constant
   for a non-zero range, then discontinuously jumps up and is constant thereafter.

3. If $M - L \leq c$, $v_M^* = \begin{cases} M & \text{for } R \leq M + c \\ R - c & \text{for } M + c < R \leq min\{M + 2c - c/\lambda, L + 2c\} \\ M & \text{for } R > min\{M + 2c - c/\lambda, L + 2c\} \end{cases}$, i.e.,
   $v_M^*(R)$ is constant for a non-zero range, then increasing and then discontinuously drops and is constant thereafter.

While Corollary 1 makes it clear that the pattern of $v_M^*(R)$ depends on the value of $M - L$, one might observe that the different cases listed in the corollary share some clear similarities: when $R - M$ is small, $v_M^*(R)$ is constant; then, for larger values of $R - M$,

---

[7]That is, there are two constraints guiding how far the median is willing to compromise. To see this more in detail, suppose L is the closer one to M, while R is further away from M. Then compromise is needed in the direction of R. The first constraint is that the policy cannot be so far right that L is no longer willing to sign (this is represented by $R - L \leq 2c \Rightarrow M - L \leq 2c - (R - M)$ in point 2). The second is that it has to be worth it for M to get the third signature, that is, the ideological loss cannot be so large that the added strength is not enough compensation. This is the constraint $R - M \leq 2c - c/\lambda$ in point 2. If this constraint is not met, the leftmost policy that R is willing to sign ($v = R - c$) is not in V_M. Naturally, once the third signature is attainable, the median just has to go as far right as needed to make R indifferent between signing and not, hence choosing $v_M^* = R - c$.

$v_M^*(R)$ might be increasing; and finally, for even larger values of $R - M$, $v_M^*(R)$ might drop (and stay constant afterwords). This similarity of the different patterns suggests that, when taking the model to the data, it is be possible to distill clear unified characteristics, at least for sufficiently small values of $R - M$. The next proposition presents these unified characteristics. The thought experiment the proposition performs is the following. Suppose there is a large number of groups. Now suppose we observe an R at a given distance r from M in these groups (of three agents each). The question the proposition answers is: how will the equilibrium policy depend on $R - M$?

PROPOSITION 2   *Define $r \equiv R - M \leq T$. Then $E\left[v_M^*(r)\right]$ has one of the following patterns:*

   (i) If $T < c$ then $E\left[v_M^*(r)\right]$ is constant in r.

   (ii) If $c < T \leq 2c - c/\lambda$ then $E\left[v_M^*(r)\right]$ is first constant and then increases in r.

   (iii) If $T > 2c - c/\lambda$ then $E\left[v_M^*(r)\right]$ is first constant, then increases in r and afterwards
it is ambiguous.

Proposition 2 has important implications for our empirical investigation. It basically implies that the range of possible views (or, if you wish, the size of c relative to this range, which captures the ideological flexibility of the agents) determines how much of the full picture will be revealed.[8] We emphasize in the proposition the range of types since, when we go to the data, we do not know how broad the type space is relative to c. Essentially, if the pool of agents is ideologically very broad (large T relative to c), enabling anything from very cohesive to very non-cohesive groups, then we are in case (iii). Here, as a function of r, we should observe policies that are first flat (representing the median getting her will), then increasing (representing the need to compromise to get the signature of R) and then ambiguous. If the pool of agents is somewhat more narrow then we are in case (ii). Here we are predicted to see the same pattern of flat and then increasing policy, but not the ambiguous part, as very non-cohesive groups cannot be formed. Finally, if the pool is narrow (point (i)), then all groups are necessarily cohesive, implying that we will only observe the flat part. An equivalent prediction can be stated by instead increasing the distance from L

---

[8]Note that, theoretically, the function $E\left[v_M^*(r)\right]$ is not defined for $r > T$.

to M. Then the pattern is flat, then falling then ambiguous.

2.2.2 Non-median proposer

We move now to analyze which policy a non-median would propose. For brevity we will focus the description and intuition on L as proposer (corresponding results can be obtained for R as proposer). The following proposition outlines the different cases.

PROPOSITION 3 *Suppose L is the proposer.*

1. If $R - L \leq c$ then $v_L^* = L$ and all agents sign.
2. If $c < R - L \leq 2c - c/\lambda$ then $v_L^* = R - c$ and all agents sign.
3. If $R - L > 2c - c/\lambda$ then $v_L^* = L$ and only two agents sign.

FIGURE 4.— Parameter space, L proposes
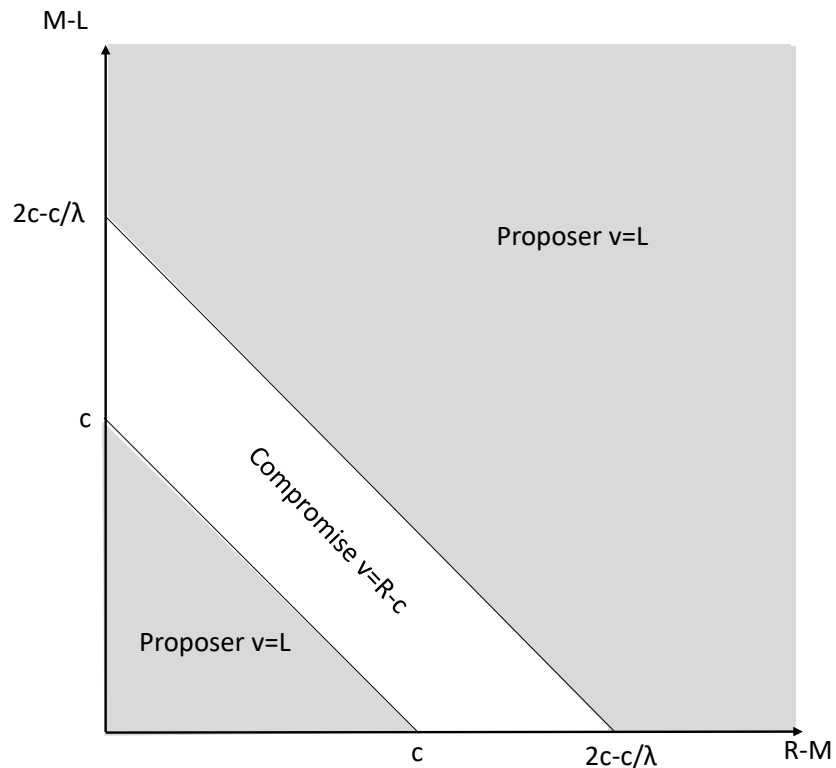
The results of the proposition are depicted in Figure 4 and its intuition rests on a similar, but simpler, intuition as that for when M is the proposer. L is willing to compromise with the policy if it is needed to get all three signatures and if it is worth it in the sense that

reduction in the value of the policy is small enough to be worth the extra signature. When all agents are close to each other (point 1 in the proposition) then all would sign L's bliss point so no compromise is needed. When R is at an intermediate distance (point 2) then it is necessary to compromise to get R's signature and it is worth it for L since the compromise is not too large. Finally when R is far from L (point 3) then a compromise is necessary to get a signature but it is not worth it for L.

2.3   Period 1: Who gets to propose?

We move now to the final step of the analysis: determining who gets chosen to be proposer in the first period of the game. We emphasize that one can think of several other ways than the one assumed here for how a proposer is chosen, but the qualitative results for how the policy is determined by the proposer (as outlined in the results thus far) is the same. In the simple method we examine here, we assume that in Period 1 the agents vote, in pairwise comparisons, for who gets to propose. In case of a draw, one of the agents is picked at random to be the proposer. To determine who wins that vote we need to analyze the payoffs for the agents under the different proposers.

LEMMA 3   *For any distribution of L, M and R, there always exist two agents who have a weakly higher payoff with M as proposer than with any other proposer.*

The intuition for this lemma is very simple. First, naturally, M is always better off with herself as proposer. This is straightforward since when M gets to propose, she can always propose a policy that mimics that of another potential proposer or propose something even better.[9] Second, R always gets a higher payoff with M as proposer than with L. The reason is that M is more willing to compromise in the direction of R than L is, and since the fallback of the proposer simply proposing her bliss point is better for R under M. Given the lemma, under pairwise comparisons (with a convention of voting for M when indifferent) the following trivially follows.

---

[9]Note however that under some ideological constellations M would propose the same ideology as L, making M (and in fact everyone else as well) indifferent between the two. This is why the lemma states a weak preference. Who they vote for in such a situation, M or L, obviously does not have an effect on the predicted policy. Likewise there are ideological constellations where M and R would choose the same policy if they get to propose.

PROPOSITION 4 *Under pairwise comparisons, M wins the right to propose.*

This voting procedure thus gives M the right to propose, implying that the empirical predictions for the questions we outline in the introduction are: 1) agents sign policies close to their bliss point (Lemma 1); 2) the policy as a function of R should follow the shape outlined in Proposition 2; 3) M has stronger bargaining power than the others (Lemma 2).

3   Identification and data

3.1   Institutional background and empirical strategy

We use as our laboratory the U.S. federal courts, where it is frequently seen that judges have the power to stymie executive orders from the U.S. president from taking place. The courts operate in a hierarchy. At the lowest level are the 94 District Courts (there are up to 4 Districts in a state). Here, cases first come to trial and there is a jury deciding issues of fact and a judge deciding issues of law. If there is an appeal, it goes up to one of the 12 Circuit Courts. There are up to 9 states in a Circuit (the name comes from the fact the judges used to ride on a horse in a "circuit" to reach the entire jurisdiction). If there is an appeal from here, the case goes to the U.S. Supreme Court, which handles very few cases, so the Circuit Courts determine the majority of what sets precedent in this common law system.[10]

Judges in the Circuit Courts have life-tenure and are appointed by the U.S. President and confirmed by the U.S. Senate. Each Circuit Court consists of a **pool** of 8-40 judges sitting in different duty stations in different states across the Circuit. For each case, three judges are randomly chosen to form the **panel** that rules on the case.[11] This panel is the equivalent of the group in our model. The three judges in the panel decide a binary verdict (affirming

---

[10]The federal courts get almost 400,000 cases per year, but only 100 reach the Supreme Court, so the Circuit Courts (taking roughly 67,000 cases per year) are responsible for the majority of precedent-setting cases, cases which law students are reading, as they learn about this high-stakes common-law space, where judges can introduce theories, shift standards or thresholds, and rule on the constitutionality of states' laws.

[11]The assignment of judges in Circuit Courts fall into two categories: 1) Once a case arrives, three randomly chosen judges are assigned to the case; 2) Once a year, the calendar is randomly set up in advance determining which judges will sit in which panels on which days in the upcoming year, and when a case comes up it gets assigned to the next panel. It is well established and has been thoroughly tested that both procedures are indeed random (see e.g. Chen and Sethi (2011), Berdejo and Chen (2016) and Chen (2016)).

or overturning the lower court verdict), where a majority of two judges is needed to set the verdict. They also compose an **opinion** (i.e., a text) motivating the verdict. The opinion (the equivalent of the policy in our theory) serves as precedent for future cases and as such has a large impact on society and policy. Furthermore, being a text, the opinion can reflect the assertiveness of the panel and its ideological composition. Our empirical strategy rests on two key ingredients: 1) judges do not choose whom to interact with; and 2) our measure of their ideology is (reasonably) correct. Point 1 is guaranteed by the random assignment (as explained above) and Point 2 is motivated in the next subsection.

3.2   Data and main variables

We use an ideology score that leverages the nature of the judicial appointment process. It is a standard summary measure coming from the Judicial Common Space database (Epstein et al. 2007) that was first coded by Giles et al. (2001). Many papers have used this score (e.g. Peresie 2005 and Kim 2009). The general idea is that—given that vacancies are rare and that Circuit Courts have a substantial impact on policy—the appointing politicians take the opportunities they get to assign judges of their ideological liking. Moreover, there is a norm of senatorial courtesy by the President. The score is therefore constructed as follows. If a judge is appointed from a state where the President and at least one home-state Senator are of the same party, the nominee is assigned the score of the home-state Senator (or the average of the home-state Senators if both members of the delegation are from the President's party).[12] If neither home-state Senator is of the President's party, the judge receives the score of the appointing President. The score thus assumes that the President does favors to senators from the same party while ignoring the preferences of senators from the other party. The score has two additional main advantages. First, it is exogenous since (unlike common measures of Supreme Court judges' ideology based on their votes or the clerks they hire) it assigns the ideology of the judge *before* her behavior at the court is observed, thus enabling us to identify how panels' decisions are affected by the ideology of their members.

---

[12]The scores of the Senators are located in a two-dimensional space on the basis of the positions that they take in roll-call votes, but only the first of the two dimensions is salient for most purposes. The ideology scores of Presidents are then estimated along this same dimension based on the public positions that they take on bills before Congress.

The second main advantage of this score is its high ability to predict judges' voting patterns in court, as established by Chen et al. (2019). The ideology score takes values in between roughly ±0.8 (see Appendix Figure 7 for a histogram of the distribution of ideology scores in our data).

To measure the ideological color of the opinions, we also employ the U.S. Courts of Appeals Database Project, a random sample of roughly 5% of appeals-courts decisions from 1925 to 2002.[13] This database includes hand-coded information on the ideological content of each coded opinion (liberal = -1, conservative = 1, and mixed or unable to code = 0), to which we will refer as **Opinion Ideology**.[14] To enable testing the predictions in Proposition 2 we calculate the **Score Relative to Center of Panel**, which is positive when the judge is more conservative than the median and negative when the judge is more liberal than the panel median.[15]

4 Empirical Results

Our first empirical prediction is Lemma 1, which predicts that agents will be more inclined to sign policies close to their own bliss point. In our empirical setting, this has been established by previous research (Epstein et al., 2013; Chen et al., 2019; see also Wahlbeck et al. 1999, Spriggs et al. 1999 and Hettinger et al. 2004). Our second empirical prediction is Proposition 4 and Lemma 2, which jointly predict that the median agent will have the strongest—though not the sole—influence on policy. This holds also in our empirical setting, as reported in previous research (see Cross, 2007; Chen et al., 2019 and, e.g., Ambrus et al. 2015 for a different setting). The main and novel prediction—both empirically and theoretically—is Proposition 2. We turn now to examine the empirical evidence for this prediction.

---

[13]Documentation and data available at http://www.cas.sc.edu/poli/juri/appctdata.htm.

[14]The Appeals Court Database Project states that for most issue categories, these will correspond to conventional notions of "liberal" and "conservative". The directionality codes parallel closely the directionality codes in the Spaeth Supreme Court database.

[15]Any analysis requiring the panel median includes only panels where there are no tied or missing scores (panels with tied scores are excluded because the identity of the median judge is not uniquely determined). All results presented in the paper are robust to including also tied scores.

4.1 Moving from Median: Flat, then Increasing

According to Proposition 2, we should see that holding constant the ideology of M and L while increasing R, there is no movement of the opinion ideology for moderate distances between R and M. This is so because M sees no reason to compromise to gain R's vote. However, as the distance grows to an intermediate distance, the other judges will compromise a bit in order to maintain the legitimacy of the opinion (e.g. the opinion will move in the direction of R when R moves further away from M).[16] We again emphasize that since the range of judge ideologies relative to their flexibility (c) is not observable, we do not know whether we should see only the flat part or also the increasing part predicted. What we do know is that we should not (if the model is correct) observe any other pattern, e.g., it should not be increasing right away.

We analyze these predictions in the data in two ways. In this sub-section, Figure 5 presents a separate analysis for judges on the left of the other two panelists and an analysis of judges on the right of the other two panelists. The left figure presents a local polynomial estimator for the relation between two variables. One variable is whether the outcome is a conservative opinion. The second variable is the left judge's score relative to the panel center, controlling for the scores of the center judge and the right judge. Both variables are residualized on circuit-by-year fixed effects. We see that as the left judge becomes more distant and reaches an intermediate distance away from the other two judges, the opinion is more reflective of the left judge's ideology. The right figure presents the same pattern for the right judge. These observed patterns align with case (ii) of Proposition 2, implying an intermediate cohesiveness (or, alternatively, intermediate flexibility) of the judges in the pool.

---

[16]Eventually, when R becomes large the effect on the opinion is ambiguous.

FIGURE 5.— Ideology of Opinion and Ideology Scores of Panel Members
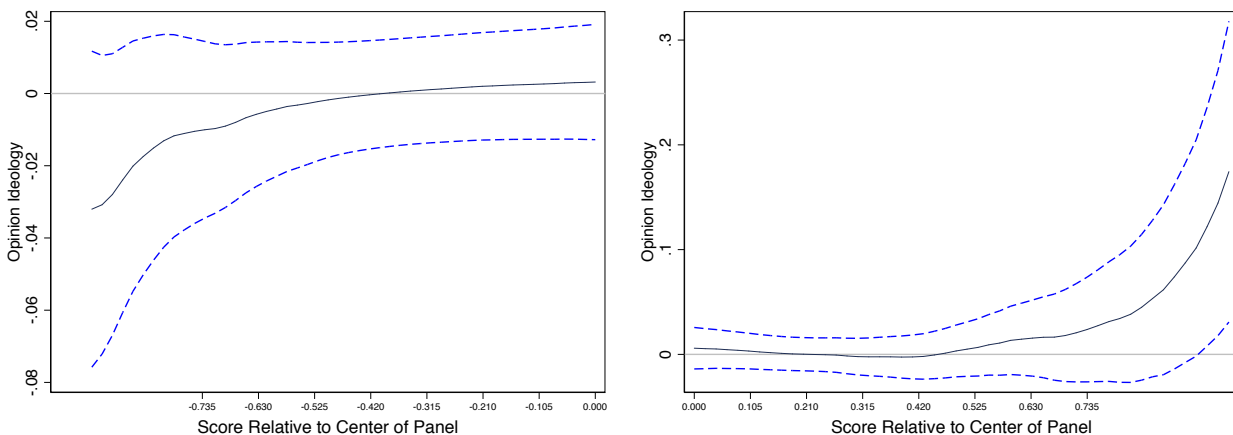
Notes: x-axis: Ideology score of a judge demeaned by the median of the *panel* of judges assigned on the case, where relatively more conservative scores are along the right on the x-axis. y-axis: Opinion ideology, which is coded as 1 for conservative, 0 for mixed or not applicable, and -1 for liberal. The figure presents a local polynomial regression with an Epanechnikov kernel, where the dependent variable is opinion ideology. The left sub-figure represents judges on the left of the other two panelists, and the right sub-figure represents the same for the judges on the right of the other two panelists. Both the x- and y-axis are residualized by Circuit-by-year fixed effects. The dashed lines depict the 95% confidence interval. Data come from the U.S. Courts of Appeals Database Project (1925-2002 5% Sample). Ideology scores come from the Judicial Common Space database. Sample includes three-judge panels where there are no tied or missing scores.

4.2 Structural Break Test

In this sub-section, we ask whether the regression coefficient is stable over the range of judge ideologies. We test whether the data abruptly changes in ways not predicted by the model by testing for structural breaks in the residuals. Under the null hypothesis of no structural break, the cumulative sum of residuals will have mean zero. Inference is based on a sequence of sums of recursive residuals (in time series, this would be a "one-step-ahead forecast error") computed iteratively from nested subsamples of the data. Values of the sequence outside an expected range suggest a structural change in the model.

Figure 6 presents the raw data in bins for all the judges pooled together (a single observation is a judge in a panel, where each dot represents the average of all opinions in a bin of judges with similar ideology scores relative to the panel median). A slope becomes visible to the left and right of the vertical lines. Table I presents a formal test for structural breaks at these vertical lines using the cumulative sum of the recursive residuals. Panel A reports tests where the x-axis is divided into 20 evenly spaced units, and Panel B reports

18

the same for 40 units. In each panel, the first row reports a statistically significant structural break that exists somewhere over the entire range of the x-axis. The second and third rows investigate parameter stability for negative and positive values of the x-axis, respectively. In line with the visualization provided in Figure 6, the structural break is estimated at around 20% from the end on both sides: for the analysis with 20 bins, a trend break statistically significant at the 1% level is observed on both the left and on the right, at 7 bins distance from the center. For the analysis with 40 bins, a trend break is statistically significant at the 1% level for the right but only the 5% level for the left, at 16 bins to the right and 10 to the left from the center.[17]

FIGURE 6.— Ideology of Opinion and Ideology Scores of Panel Members
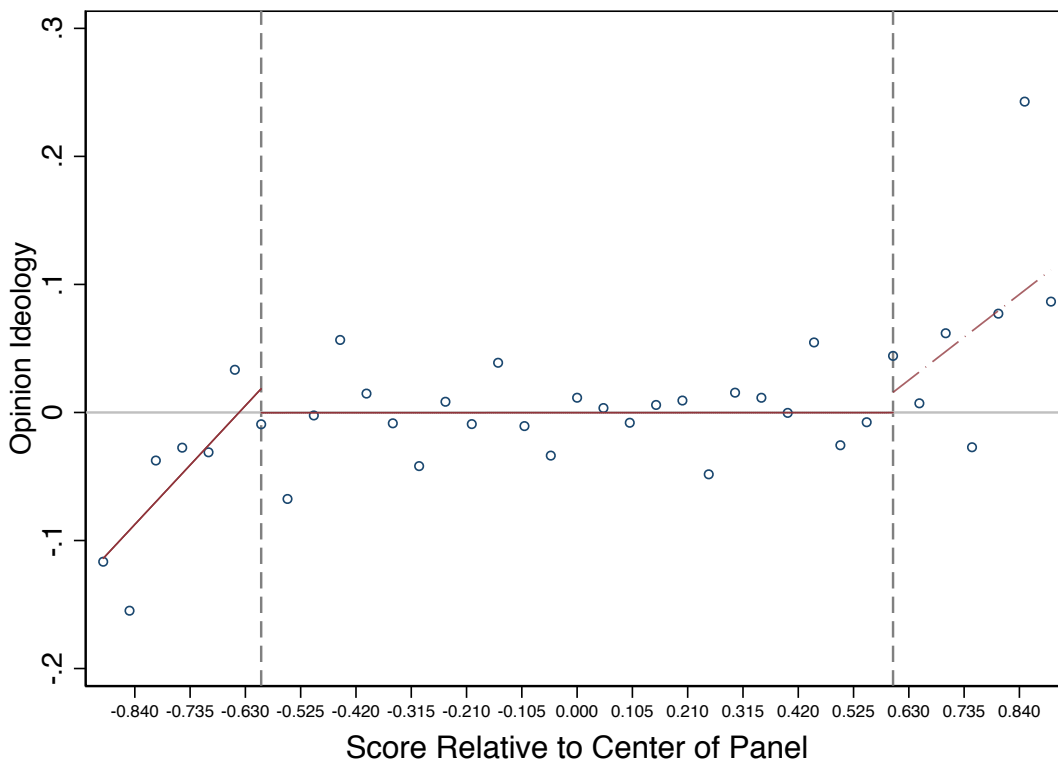
Notes: x-axis: Ideology score of a judge demeaned by the median of the *panel* of judges assigned on the case, where relatively more conservative scores are along the right on the x-axis. y-axis: Opinion ideology, which is coded as 1 for conservative, 0 for mixed or not applicable, and -1 for liberal. Each dot represents the average of all opinions in a bin of judges with similar ideology scores. The lines represent the linear fit for a section of the data. Data come from the U.S. Courts of Appeals Database Project (1925-2002 5% Sample). Ideology scores come from the Judicial Common Space database. Sample includes three-judge panels where there are no tied or missing scores.

[17]The parameter stability test for positive values of the x-axis use the cumulative sum of the OLS residuals, which is more useful for detecting structural breaks on the right-side of the sample.

TABLE I

Tests for Structural Breaks (Cumulative sum test for parameter stability)

Panel A: 20 Bins	Test Statistic	1% Critical Value	5% Critical Value	10% Critical Value
Entire Range	1.08	1.14	0.95	0.85
< 0	1.40	1.14	0.95	0.85
> 0	1.94	1.63	1.36	1.22
Estimated Breaks: 7 and -7 relative to center				

Panel B: 40 Bins	Test Statistic	1% Critical Value	5% Critical Value	10% Critical Value
Entire Range	1.50	1.14	0.95	0.85
< 0	1.07	1.14	0.95	0.85
> 0	2.88	1.63	1.36	1.22
Estimated Breaks: 16 and -10 relative to center				

Notes: This table reports a cumulative sum test for parameter stability over the x-axis. The test statistic is constructed from the cumulative sum of the recursive residuals. Panel A investigates reports tests where the x-axis is divided into 20 evenly spaced units, while Panel B reports the same for 40 units. In each panel, the first row reports a statistically significant structural break. The second and third rows investigate parameter stability for negative and positive values of the x-axis, respectively, and the parameter stability test for positive values of the x-axis use the cumulative sum of the OLS residuals, which is more useful for detecting structural breaks on the right-side of the sample.

5   Conclusions

We analyze situations in which the choice of whether to support a policy affects its impact. Our overarching research question is: What is the effect on the bargaining outcome when agents care about such legitimacy of the decision? The group's ideological cohesion has a non-monotonic effect on the policy. When all ideologies closely align (high cohesion), the median gets her will fully since all agents sign onto such a policy. When the group has extreme ideological disagreement, there is no common ground for a policy to be signed by all, hence the median (endogenously having the most influence) gets the policy in line with her ideology but attains few signatures. Finally, when the group is at an intermediate level of cohesion, the policy deviates from the median's bliss point since the median can, and is willing to, compromise to gain higher legitimacy. This creates a novel empirical prediction whereby the policy is independent of non-median-members' ideology in cohesive groups, then there is a structural break so that at intermediate agreement the policy is increasing in a non-median's ideology. We test this prediction in the natural-experimental setting of the US

Circuit Courts and find support for it and other predictions of the model.

Naturally, in our ambition to develop a parsimonious and widely useful model, we have abstracted from many interesting features that have been considered by the theoretical literature on legislative bargaining that we relate to and by the law literature on the specific setting of collegial courts. Still, we believe that our model delivers an important lesson that pertains to many political settings and to the wide topic of group decision making.

References

Ambrus, A., B. Greiner, and P. A. Pathak (2015). How individual preferences are aggregated in groups: An experimental study. *Journal of Public Economics 129*, 1 – 13.

Anwar, S., P. Bayer, and R. Hjalmarsson (2018). Politics in the courtroom: Political ideology and jury decision making. *Journal of the European Economic Association 17*(3), 834–875.

Banks, J. S. and J. Duggan (2006). A general bargaining model of legislative policy-making. *Quarterly Journal of Political Science 1*(1), 49–85.

Baron, D. P. and J. A. Ferejohn (1989). Bargaining in legislatures. *American political science review 83*(4), 1181–1206.

Berdejo, C. and D. L. Chen (2016). Electoral Cycles Among U.S. Courts of Appeals Judges. Technical report.

Cameron, C. M. and L. A. Kornhauser (2009). Modeling collegial courts (3): Adjudication equilibria. *NYU School of Law, Public Law Research Paper* (09-39), 09–29.

Cardona, D. and C. Ponsati (2011). Uniqueness of stationary equilibria in bargaining one-dimensional policies under (super) majority rules. *Games and Economic Behavior 73*(1), 65–75.

Chen, D. L. (2016). Priming Ideology: Why Presidential Elections Affect U.S. Courts of Appeals Judges. Technical report.

Chen, D. L., M. Michaeli, and D. Spiro (2019). Non-confrontational extremists.

Chen, D. L. and J. Sethi (2011, October). Insiders and outsiders: Does forbidding sexual harassment exacerbate gender inequality? Working paper, University of Chicago.

Cohen, A., A. Klement, and Z. Neeman (2015). Judicial decision making: a dynamic reputation approach. *The Journal of Legal Studies 44*(S1), S133–S159.

Cross, F. B. (2007). *Decision making in the US Courts of Appeals*. Stanford University Press.

Diermeier, D. and A. Merlo (2000). Government turnover in parliamentary democracies. *Journal of Economic Theory 94*(1), 46–79.

Duggan, J. et al. (2000). A bargaining model of collective choice. *American Political Science Review 94*(1), 73–88.

Epstein, L., W. M. Landes, and R. A. Posner (2013). *The Behavior of Federal Judges: A Theoretical and Empirical Study of Rational Choice*. Harvard University Press.

Epstein, L., A. D. Martin, J. A. Segal, and C. Westerland (2007). The judicial common space. *Journal of Law, Economics, and Organization 23*(2), 303–325.

Eraslan, H. and K. S. Evdokimov (2019). Legislative and multilateral bargaining. *Annual Review of Economics 11*, 443–472.

Giles, M. W., V. A. Hettinger, and T. C. Peppers (2001). Picking Federal Judges: A Note on Policy and Partisan Selection Agendas. *Political Research Quarterly 54*(3), 623–641.

Gratton, G., R. Holden, and B. Lee (2020). Political capital. Technical report.

Hettinger, V. A., S. A. Lindquist, and W. L. Martinek (2004). Comparing attitudinal and strategic accounts of dissenting behavior on the us courts of appeals. *American Journal of Political Science 48*(1), 123–137.

Kim, P. T. (2009). Deliberation and strategy on the united states courts of appeals: An empirical exploration of panel effects. *University of Pennsylvania Law Review 157*, 1319–1381.

Peresie, J. L. (2005). Female judges matter: Gender and collegial decisionmaking in the federal appellate courts. *The Yale Law Journal 114*(7), 1759–1790.

Predtetchinski, A. (2011). One-dimensional bargaining. *Games and Economic Behavior 72*(2), 526–543.

Spriggs, J. F., F. Maltzman, and P. J. Wahlbeck (1999). Bargaining on the us supreme court: Justices' responses to majority opinion drafts. *The Journal of Politics 61*(2), 485–506.

Wahlbeck, P. J., J. F. Spriggs, and F. Maltzman (1999). The politics of dissents and concurrences on the us supreme court. *American Politics Quarterly 27*(4), 488–514.

A   Proofs

A.1   Proof of Lemma 1

PROOF: First, if $|t - v| \leq c$, then clearly signing the policy is better for agent t than not signing, giving the utility function in (1). Hence, since signing is a dominant strategy whenever $|t - v| \leq c$, we get that all three agents sign if $|t - v| \leq c$ for all three. Second, regardless of the value of t, if only one other agent signs the policy, then t prefers to sign as well, by the assumption that $-K$ is lower than the lowest possible payoff under agreement. Hence, if $max_t |t - v| > c$, a pure SPE is one where only two agents sign (there is no SPE where all three sign because the agent with $max_t |t - v|$ has a profitable deviation to not signing).   Q.E.D.

A.2   Proof of Lemma 2

PROOF: Any $v \in V_t \cap V_3$ is guaranteed to be signed by all three. The utility of the proposer when proposing such v is then given by $[c - |t - v|] \lambda$, which monotonically decreases in $|t - v|$. By the definition of V_t, this utility is larger than the proposer's utility from any $v \notin V_t$. If however $V_t \cap V_3 = \emptyset$, then, by the definition of V_t and V_3 and by Lemma 1, the proposal of t would be signed only by two agents, in which case t's utility is given by $c - |t - v|$, which is maximized when $v = t$.   Q.E.D.

A.3   Proof of Proposition 1

PROOF: 1) When $R - M \leq c$ and $M - L \leq c$ then, following Lemma 1, both R and L would sign $v = M$ implying, by Lemma 2, that $v_M^* = M$. 2) When $c < R - M \leq 2c - c/\lambda$ and $M - L \leq 2c - (R - M)$ then $R - M > c \geq M - L$ hence, by Lemma 1, $v = M$ would yield only two signatures. At the same time, since $M - L \leq 2c - (R - M) \leftrightarrow R - L \leq 2c$ and $R - M \leq 2c - c/\lambda$ there exist, by Lemmas 1 and 2, another v s.t. $v \in V_M \cap V_3$. Since $R - M > c$, by Lemma 1, such v necessarily has $v > M$. Minimizing, by Lemma 2, $|M - v|$ among $v \in V_M \cap V_3$ then implies choosing $v_M^* = R - c$. 3) Follows the same steps as for point 2. 4) Otherwise $V_M \cap V_3 = \emptyset$, either because $R - L > 2c$ (implying $V_3 = \emptyset$), or because $max\{R - M, M - L\} > 2c - c/\lambda$. The fact that $V_M \cap V_3 = \emptyset$ implies, by Lemma 2, that $v_M^* = M$ and two agents sign.   Q.E.D.

A.4   Proof of Proposition 2

PROOF: (i) If $max|R - L| = T \leq c$, then, for any t and any v in the range T, we have $[c - |t - v|] \geq 0$, implying that all three agents sign the proposal. Consequently, it is a dominant strategy for M to propose $v = M$, in which case $v_M^* (R)$ is constant. (ii) If c is such that $c < max|R - L| = T \leq 2c - c/\lambda$,[18] then the third zone of Corollary 1(3), where v_M^* drops from $R - c$ to M, is unreachable (because, in that region, $c \geq M - L$ and $R > min\{M + 2c - c/\lambda, L + 2c\}$ should both hold, but this requires that either $R - L > 2c$—which is

---

[18]This part is best understood by looking at Figure 3 and focusing on the simplex whose edges are $(0, 0) , (2c - c/\lambda, 0) , (0, 2c - c/\lambda)$.

impossible given that $T \leq 2c - c/\lambda < 2c$—or $max\,|R - L| \leq 2c - c/\lambda < R - M$, which is impossible as well). Similarly, the second zone of Corollary 1(2) is unreachable too, because it requires $R - L > 2c$. Hence, in any of the three possible regions of $M - L$ (as defined in Corollary 1), v_M^* is either constant (at M in the first region and at $L + c$ in the second), or is weakly increasing (discontinuously for every pair of $\{L_i, M_i\}$ in the third region for which r crosses c, and continuously for every such pair when r continues to increase beyond c and $v_M^* = R - c$). (iii) From a theoretical point of view there are two subcases here. The first subcase is when c is such that $2c - c/\lambda \leq max\,|R - L| = T \leq 2c$.[19] In this case—like in case (2) above—the second zone of Corollary 1(2) is unreachable. Furthermore, the only way to reach the third zone of Corollary 1(3), where $v_M^* = M$ because $R > min\{M + 2c - c/\lambda, L + 2c\}$, is that $R - M > 2c - c/\lambda$ (see also Figure 3). Hence, for pairs of $\{L_i, M_i\}$ in the first and second regions v_M^* is constant (at M in the first region and at $L + c$ in the second), while in the third region v_M^* is constant until $r = c$. When r reaches c, the only relevant pairs of $\{L_i, M_i\}$ are in the third region, and for all of them v_M^* is increasing discontinuously when r crosses c and then continuously when r continues to increase beyond c (where $v_M^* = R - c$). This continues till r reaches the point where $r = R - M = 2c - c/\lambda$. At that point, v_M^* drops (for all remaining pairs of $\{L_i, M_i\}$) from $R - c = M + c - c/\lambda$ to M, and stays there. The second subcase is when $T > 2c$, hence any zone in Figure 3 is reachable. As long as $r = R - M < c/\lambda$, $E_i\left[v_{M_i}^*(r)\right]$ is constant, because in each of the three possible regions v_M^* stays constant. Then, when r is in the range $[c/\lambda, c]$, v_M^* weakly increases (due to discontinuous "jumps" in the second region, where v_M^* switches from $L - c$ to M). At $r = c$ there is another discontinuous increase, due to v_M^* switching from M to $R - c$ for all pairs of $\{L_i, M_i\}$ that are in the third region and for which $r = R - M$ can exceed c. Once r crosses the value of c, $v_M^* = R - c$ for all remaining pairs hence $E_i\left[v_{M_i}^*(r)\right]$ continues to increase, but there are pairs of $\{L_i, M_i\}$ (also in the third region) for which v_M^* discontinuously drops from $R - c$ to M, which explains why, overall, the pattern is ambiguous at this stage. Finally, after r crosses the point where it equals $2c - c/\lambda$, v_M^* drops (for all remaining pairs of $\{L_i, M_i\}$) from $R - c = M + c - c/\lambda$ to M, and stays there. *Q.E.D.*

A.5 Proof of Proposition 3

PROOF: 1) When $R - L \leq c$ (implying also $M - L \leq c$) then, by Lemma 1 all agents sign $v = L$, hence by Lemma 2 this is the proposed policy.

2) When $c < R - L$ then, by Lemma 1, only two agents would sign $v = L$. At the same time there exists $v \in V_L$ that all would sign since $R - L \leq 2c - c/\lambda$. Hence, by Lemma 2, L chooses the v that is closest to herself in V_L, this is $v_L^* = R - c$.

3) When $R - L > 2c - c/\lambda$ then, by Lemma 1, there exist no $v \in V_L$ that would be signed by all, hence by Lemma 2, L proposes $v_L^* = L$. *Q.E.D.*

---

[19]This part is best understood by looking at Figure 3 and focusing on the simplex whose edges are $(0,0)\,, (2c,0)\,, (0,2c)$.

A.6 Proof of Lemma 3

PROOF: Proposition 3 states that when L proposes, either (1) $v_L^* = L$ or (2) $v_L^* = R - c$. In case (1), given Proposition 1, we have $v_L^* = L \leq v_M^* \leq R$, implying that v_M^* is closer to R than v_L^* is.[20] This also guarantees a higher payoff for R when M rather than L proposes, unless one of the following two (soon-to-be-ruled-out) scenarios holds: (i) $R - c < v_L^* = L < v_M^* \leq R$ (so that R has a positive payoff from both v_L^* and v_M^*) and the signatures are such that $v_L^* = L$ gets three signatures while v_M^* gets only two, or (ii) $v_L^* = L < v_M^* < R - c$ (so that R has a negative payoff from both v_L^* and v_M^*) and the signatures are such that $v_L^* = L$ gets two signatures while v_M^* gets three. However, (i) is impossible because $R - c < L < v_M^* \leq R$ implies further that $v_M^* < L + c$, hence would be signed by all three agents and not only by two as wrongly asserted; and (ii) is impossible because $v_M^* < R - c$ implies that R might sign v_M^* only if L does not, implying it cannot get three signatures, again as wrongly asserted. Finally, if case (2) holds, so that $v_L^* = R - c$ (and so R gets a zero payoff), then it must be—by Proposition 3—that $R - L \leq 2c - c/\lambda$, implying also that $R - M \leq 2c - c/\lambda$ and $M - L \leq 2c - c/\lambda$. Hence, by Proposition 1, we know that all three agents sign v_M^*, implying R gets a positive payoff.   *Q.E.D.*

A.7 Proof of Proposition 4

PROOF: The proposition trivially follows from Lemma 3.   *Q.E.D.*

---

[20]Recall that $v_M^* \in \{L + c, M, R - c\}$. We know that if $v_M^* = R - c$ then it must be that $L < R - c$, because Proposition 1 states that $v_M^* = R - c$ only if $M < R - c$, in which case also $L < R - c$; and we know that if $v_M^* = L + c$ then it must be that $L + c < R$ because Proposition 1 states that $v_M^* = L + c$ only if $L + c < M$, in which case also $L + c < R$.

# B Additional Empirical Results

B.1 Summary statistics

TABLE II

SUMMARY STATISTICS

Vote-Level	Mean	Standard Deviation
Dissent	0.03	0.17
Concur	0.02	0.13
Ideology Score	0.01	0.34
N	541182	
Case-Level (Songer-Auburn sample)		
Opinion Ideology (1 = Conservative)	0.19	0.90
Panel Median	0.33	0.47
Ideology Score	-0.03	0.34
N	7677	
Number of Judges per Circuit-Year	16.95	9.65
Number of Circuit-Years	667	

Notes: Data on dissents and concurrences comes from OpenJurist (1950-2007). Data on opinion ideology come from the U.S. Courts of Appeals Database Project (1925-2002 5% Sample). Sample includes three-judge panels where there are no tied or missing scores. Opinion ideology is coded as 1 for conservative, 0 for mixed or not applicable, and -1 for liberal. Ideology scores come from the Judicial Common Space database (Epstein et al. 2007), which provides a summary measure using the voting patterns of the appointing President and home-state Senators.

B.2 Distribution of ideological scores

FIGURE 7.— Distribution of Ideology Scores
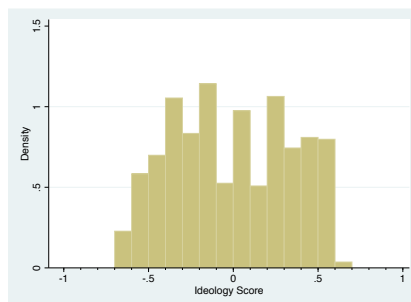
Notes: Ideology scores from the Judicial Common Space database (Epstein et al. 2007).