NON-CONFRONTATIONAL EXTREMISTS

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

Abstract: In many contexts individuals are subject to norms and decisions they disagree with ideologically. What is the effect of regularly being in an ideological minority on the propensity to confront majority norms and decisions? We study this in an ideologically-salient field setting—US appeals courts—using exogenous predictors of ideology and random assignment of judges. We find that ideological interaction silences extremists: judges who are ideologically extreme relative to their peers are less confrontational—dissent less often—than other judges despite shaping case outcomes the least. Considering many mechanisms, we find that a model of peer pressure can explain the observations.

Keywords: Judicial decision making, group decision making, ideology, peer pressure.

JEL codes: D7, K0, Z1.

*Daniel L. Chen, daniel.chen@iast.fr, Toulouse School of Economics, Institute for Advanced Study in Toulouse, University of Toulouse Capitole, Toulouse, France; Moti Michaeli, the University of Haifa, motimich@econ.haifa.ac.il; Daniel Spiro, Dept. of Economics at Uppsala University, daniel.spiro.ec@gmail.com.
We are grateful for comments from Zhijun Chen, Gabrielle Gratton, Bård Harstad, Zvike Neeman, Fabrizio Panebianco and seminar participants at Cattolica Milan, Monash, UCSD, UNSW and the SAET conference. The authors gratefully acknowledge funding from Handelsbanken Research Foundations. Work on this project is conducted while Daniel Chen receives financial support from the European Research Council (Grant No. 614708), Swiss National Science Foundation (Grant Nos. 100018-152678 and 106014-150820), and Agence Nationale de la Recherche.

1 Introduction

In many settings, individuals holding convictions can dramatically shift behavior when they cannot follow these convictions fully. Overeating when induced to give up a diet, succumbing to full-blown cheating if induced to cheat, or lying the maximum amount once deciding to lie (Baumeister and Heatherton 1996; Hurkens and Kartik 2009; Gino et al. 2010b; Gneezy et al. 2013). Preference falsification is taking an action that differs from one's true preference. The public frequently conveys, especially to researchers or pollsters, preferences that differ from what they truly want, often because they believe the conveyed preference is more acceptable socially. These phenomena have been popularly labeled the "what-the-hell-effect" (Ariely 2012) and "Private Truths, Public Lies" (Kuran (1997)Kuran 1995). In this paper, we provide quasi-experimental evidence of this happening also in a high-stakes ideological setting – judges in ideological minority seemingly give up on challenging court decisions.

The question of whether holding non-mainstream views makes a person prone to challenge norms and decisions has remained empirically open due to unobservability of individual ideology and due to endogeneity of the choice of whom to interact with. These two problems are resolved at the U.S. Courts of Appeals (U.S. Federal Circuit Courts): There exist commonly used and exogenous measures of individual ideology (Section 2), and assignment of whom an individual interacts with is determined exogenously (Section 2.1).

Each Circuit Court consists of a pool of judges from which, for each judicial case, three judges are randomly assigned to sit together on a panel. The panel decides on a verdict (affirming or overturning the lower court verdict) and composes an "opinion" (i.e., a text) motivating the verdict. It is well documented that ideology plays an important role in setting a case's outcome (e.g., Epstein et al. 2013; Sunstein et al. 2006; Berdejó and Chen 2014). The ideologically-charged decisions of the Circuit Courts received increased attention for instance, following the decision that halted President Trump's proposed "Muslim-immigration moratorium" (see, e.g., NYTimes 2017).[1]

[1]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.

Roughly speaking, the way by which an individual judge's action can affect the judicial outcome is twofold: 1) the judge may have a direct effect on the case outcome; 2) if the judge does not manage to directly affect it, she may write a minority opinion (dissent) thus confronting the majority opinion.[2] Should we expect judges in ideological minority ("extremists") to have an impact on case outcomes? If not, should we expect them to behave more or less confrontationally (i.e., dissent more or less) than others? We answer these questions by presenting robust empirical results showing that I) extremists rarely affect the verdict (Section 3.1) and II) a judge's ideological disagreement with the majority is a strong predictor of when she will dissent (Section 3.2), yet III) extremists, i.e. those who are most frequently in ideological disagreement, are on aggregate the least confrontational against majority opinions (Section 3.3). This final observation is surprising given that (I) and (II) imply that these extremists have the most reasons to be confrontational.

Our empirical analysis indicates that this finding—that judges who are in strong ideological disagreement with their peers on the greater judge pool remain silent—is not about people with extreme ideology per se. Rather, we emphasize that it occurs when a judge is ideologically extreme *relative* to the current ideological composition of her pool. This suggests that the interaction between peers who are in ideological disagreement is silencing those with views far from the mainstream. The implications of this finding are that it may hide undercurrents of dissatisfaction, distort the perception of the distribution of views and create false impressions of consensus where it is absent.

To understand what may be the reason for this silencing, i.e., to rationalize the empirical observations, we consider a large number of explanations that touch upon various aspects of judicial and group decision making (Section 4 and Appendix D.1). Out of these various explanations, only two are able to account for all the empirical observations presented: one where confrontation (dissent) has an expressive purpose and one where it has an instrumental purpose (trying to reverse the verdict using the Supreme Court). We develop formal models for both explanations and derive auxiliary predictions for these two models

[2]Dissenting means stating a disagreement with both the verdict and the majority opinion. The judge may also concur which is a milder form of disagreement that is used by a judge who agrees with the verdict but disagrees with the motivation, i.e., disagrees with the opinion. See further details in Section 2.1.

along dimensions by which they differ. Testing these predictions refutes the instrumental model (which is hence relegated to the appendix, D.1.15 and D.2) and lends support to the expressive-dissent model which is therefore described in detail in the main text (Section 5).[3] This model, which is inspired by Beim et al. (2014) and Cameron and Kornhauser (2009), has the following features.

Three judges are randomly assigned to a panel where they negotiate the ideological flavor of the verdict, with the median judge succeeding to set the verdict to match her own ideology (in line with our empirical findings). Each judge then decides whether to confront this verdict by formally dissenting. Doing so is costly in terms of time and collegiality.[4] But not dissenting entails a personal cost: the more often a judge signs verdicts she disagrees with, the worse she feels. We show that the trade-off for if and when to dissent may drive judges with non-mainstream ideology to be less confrontational than judges with more mainstream ideologies. However, this result holds only provided that the individual's ideological cost is concave. A concave ideological cost captures that once a person has deviated from what she considers ideologically or morally right, she becomes almost numb to additional deviations. It is embodied in the notion of the "What the Hell Effect" (Ariely 2012; Baumeister and Heatherton 1996), which has been shown to apply to behaviors ranging from diet-keeping (Baumeister and Heatherton 1996) through truth-telling (Kajackaite and Gneezy 2017; Gino et al. 2010a; Hurkens and Kartik 2009) to voting (Kendall et al. 2015).[5]

The intuition for the result is as follows. A judge who is in ideological minority in the greater group of peers (the court's pool) will rarely be the median judge in a panel and will mostly have to decide whether or not to sign verdicts she disagrees with. However, always

[3] The conclusion that dissents are not meant to affect the Supreme Court are in line with that of Kim (2009).

[4] A very large literature, including judges' writings about their own experience, documents a norm of consensus (see e.g. Edwards and Livermore 2008). Epstein et al. (2011) refer to this as "dissent aversion". This literature attributes the peer pressure largely to collegiality concerns (Fischman 2011; Hettinger et al. 2007, Sunstein et al. 2006).

[5] A concave cost has been argued by Osborne (1995) to be realistic in ideological settings as well. Lim (2013) shows that the best cost function to fit data coming from judicial courts is one that is first convex (for small ideological deviations) and then concave (for large deviations). We do not use such an S-shaped function because, for tractability, our model utilizes power functions, whose curvature is fixed. However, in a numerical simulation (omitted to conserve space) we find that an S-shaped function as considered by Lim (2013) can replicate our empirical findings. Hence, our theoretical finding is very much in line with Lim (2013).

dissenting on verdicts she does not like would imply a very high collegial pressure as she will be facing such verdicts virtually all the time, and signing only some verdicts while dissenting against others helps little when the perceived cost of signing few unfavorable verdicts instead of many is almost the same. Hence, facing a sufficiently high collegial pressure, such an extreme judge will tend to sign virtually all verdicts, thus being non-confrontational. In comparison, judges with more mainstream ideology will more often be the median of their panels, or agree with the decision of the median of the panel, hence will less often need to decide whether to sign unfavorable verdicts. So when they do face this problem they dissent, since the cost of deviating from their ideological bliss point (due to the concavity) is high even if it is only rarely done. Such judges will therefore dissent from time to time. Thus, overall, judges far from the mainstream ideology will be the least confrontational despite having the most reasons to be just that (rarely affecting the verdict and often disagreeing with the panel's verdict). More precisely, our theoretical model predicts and the empirical results show (see visualization in Figure 3) a hill-shaped relationship between a judge's dissent rate and how extreme she is relative to her peers in the pool of judges she interacts with over time. That is, centrist judges rarely dissent, moderately ideological judges often dissent and extremely ideological judges rarely dissent. The above theoretical logic does not hold if the personal cost of bliss-point deviation is linear or convex. In the conclusions (Section 6) we discuss the broader empirical and theoretical implications of our finding of a concave ideological cost, should it indeed be true and hold more generally.

2 Identification and data

2.1 Institutional background and empirical strategy

The U.S. Federal Courts are a system of local level (District Court), intermediate level (Circuit Court), and national level (Supreme Court) councils. Members of these are appointed by the U.S. President and confirmed by the U.S. Senate. They are responsible for the adjudication of disputes involving federal law. Their decisions establish precedent for adjudication in future cases in the same court and in lower courts within its geographic boundaries. Each state has 1–4 District Courts. The 94 U.S. District Courts serve as trial

courts with juries. The 12 U.S. Circuit Courts (Courts of Appeals), which are the empirical focus of this paper, take cases appealed from the District Courts. The Circuit Courts have no juries. Each Circuit Court presides over 3–9 states and consists of a **pool** of 8–40 judges (depending on the circuit).

Circuit Courts rule on the application of federal law, such as the constitutional validity of state laws, among other things. 98% of their decisions are final. Hence, they have a substantial impact on precedence, decision making and policy in the US.

In each of the 12 Circuit Courts, every judicial case gets three randomly assigned judges who are drawn from the court's pool of judges (see details below). We refer to these three judges as the **panel**. The three judges in the panel decide a binary verdict (affirming or overturning the lower court verdict), where a majority of two judges is needed to set the verdict (formerly called "disposition"). They also compose an **opinion** (i.e., a text) motivating the verdict. The opinion serves as precedent for future cases and as such has a large impact on society and policy. A judge has to write a separate (minority) opinion if she either **dissents** (votes against the binary verdict) or **concurs** (votes for the verdict but for a different reason, as manifested in her minority opinion).[6] Both dissents and concurrences are costly in terms of time and collegiality and they cannot be cited as binding precedent. Note that, for a judge, dissenting and concurring are two mutually exclusive actions that both imply expressing dissatisfaction with the court's decision—a form of confrontation.

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 motivated below and Point 2 in the next subsection.

Who a judge interacts with is determined by which panels a judge is assigned to and by the assignment of judges to the pool from which the panels are drawn. The assignments to the pool and to panels are both plausibly exogenous for a single judge. Starting with the assignment to the pool, appointment of judges to Circuit Courts is done by the President and confirmed by the U.S. Senate. Strictly speaking, a judge may of course turn down such

[6] Given the binary nature of the verdict, at most one judge will dissent, but it can be the case that one judge dissents and another one concurs (where the third judge writes the "majority" opinion).

a job offer and there is scarcely any way of knowing the extent to which this happens. But vacancies are rare and it is considered a great honor to serve in a Circuit Court—for most judges this would be the peak of their career. Hence, we do not find it likely that a judge will turn down an offer awaiting a better one.[7] Importantly, given the life tenure of judges, once a judge is appointed she has no control over her peer group in practice. This is since 96% of all judges serve until they retire or pass away and since a vacant position on a circuit appears only when a judge retires (61% of vacancies in our sample), passes away (12%) or the number of seats is expanded (24%)[8] while resignations are rare (3% of vacant positions). So, while it cannot be entirely ruled out that judges within a circuit pressure each other to quit, this is not a quantitatively important problem for what we do.

The assignment of judges into panels within a Circuit Court is "as if" random. While assignment rules vary by circuit, there is no evidence that cases are systematically assigned to judges based on their or the cases' characteristics. True, some circuits take into account concerns such as workload, leading some scholars to argue that case assignment is not fully random but only quasi-random (see e.g. Chilton and Levy 2015; Hall 2010). However, we follow previous research in treating case assignments in these courts as being as-if-random (for a recent example see Beim et al., forthcoming).[9]

2.2 Data and Measurement of Ideology

The data on dissents and concurrences come from Openjurist, which contains all published cases from 1950 to 2007 (all summary statistics are presented in Appendix Table 1). The data was first digitized by Berdejó and Chen (2014) for whether there was a dissenting opinion and whether there was a concurring opinion. The current paper extracts the judge names and merges each judge with his/her ideology score.

[7] On average there is a vacancy once every 1.5 years for any single circuit (naturally, less often for smaller circuits and more often for larger ones). By then, the window of opportunity may have changed due to a change of President or senators or due to the arrival of competitors that better fit these politicians' tastes.

[8] New seats appeared, for instance, when the 5th Circuit was split into two in 1981.

[9] Indeed, Chen and Sethi (2011) use data from Boyd et al. (2010) and Sunstein et al. (2006), who code 19 case characteristics as determined by the lower court for 415 gender-discrimination Circuit Court cases, and find that case characteristics are uncorrelated with judicial-panel composition; Berdejo and Chen (2016) report omnibus tests of whether case and litigant characteristics vary over 4-year cycles; and Chen (2016) does the same for the caseloads and characteristics of judges authoring or sitting on the panel.

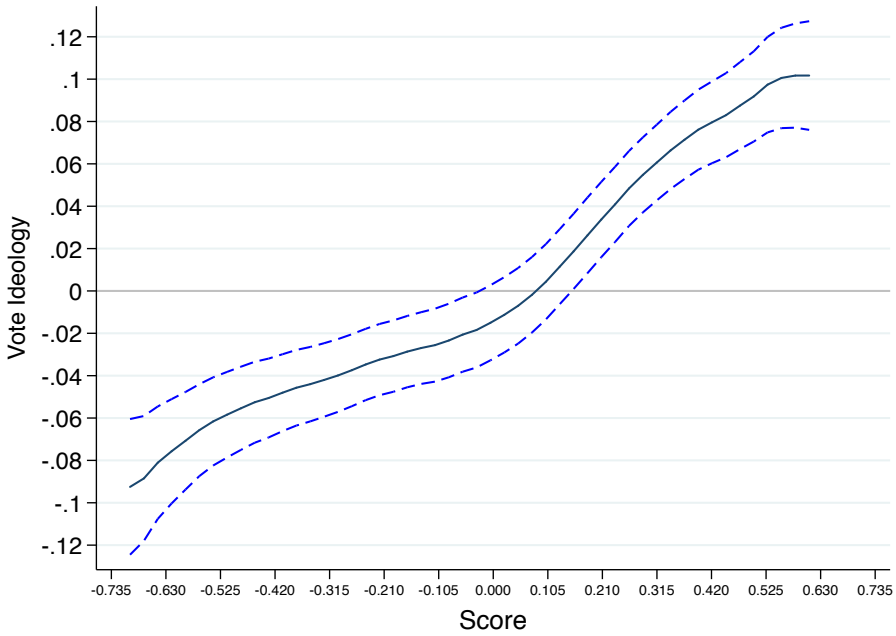
The ideology score we use is a standard summary measure coming from the Judicial Common Space (JCS) database (Epstein et al. 2007) that was first coded by Giles et al. (2001). Many papers have used this score (e.g. Peresie 2005 and Kim 2009), which was referred to by Cross (2007, p. 19) as the "best currently available measure for circuit court judicial ideology". This is a unidimensional score, which assumes that various ideological dimensions can be effectively collapsed to only one that goes from "very liberal" to "very conservative". Most importantly, unlike common measures of Supreme Court judges' ideology, this ideology score is *exogenous* (see details below). A further discussion of existing ideological measures and their pros and cons appears in Appendix A.1.

The general idea behind the JCS score is leveraging the political appointment process. That is, it assumes 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. It further exploits the norm of senatorial courtesy by the President and is 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).[10] 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 it assigns the ideology of the judge *before* her behavior at the court is observed, which of course is key since we are interested in how a judge's behavior at the court is affected by her ideology. The second main advantage of this score is its high ability to predict judges' voting patterns in court, as clearly visualized in Figure 1 (see below how the voting variable is constructed). The ideology score takes values in between roughly ± 0.7 (see Appendix Figure 1 for a histogram of the distribution

[10] 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.

of ideology scores in our data).

FIGURE 1.— Vote Ideology and the Judicial Common Space database Ideology Score - local polynomial



Notes: x-axis: (Non-demeaned) Judicial Common Space database Ideology score of a judge, where more conservative scores are along the right on the x-axis. y-axis: Vote ideology, demeaned to be centered at zero. The figure presents a local polynomial regression with an Epanechnikov kernel, where the dependent variable is Vote Ideology, which is coded as 1 for conservative, 0 for mixed or not applicable, and -1 for liberal. The dashed lines depict the 95% confidence interval. Data come from the U.S. Courts of Appeals Database Project (1925-2002 5% Sample).

For robustness, we use the party of appointing President as an alternative ideology score, and find qualitatively the same results: see Section 3.3.

To examine the ideological flavor of verdicts, we employ the U.S. Courts of Appeals Database Project, a random sample of roughly 5% of appeals-courts decisions from 1925 to 2002.[11] This database includes hand-coded information on the ideological content of each coded verdict (liberal = -1, conservative = 1, and mixed or unable to code = 0), to which we will refer as **Verdict Ideology**.[12] This database also reports dissents, which enables us to measure the average ideology of each judge's votes. This measure is set to equal the coded ideology of the verdict if the judge did not dissent and gets the opposite coding if the judge

[11] Documentation and data available at http://www.songerproject.org/us-courts-of-appeals-databases.html.

[12] 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.

dissented. We call this variable **Vote Ideology** and it reflects the ideology of verdicts a judge signs and composes (this variable is used, for instance, in the y-axis of Figure 1).

Our sample contains 293,868 decisions in Openjurist and 18,686 decisions in the Courts of Appeals database. Overall, 8.5% of opinions in Openjurist have dissents (6.4% have concurrences) while in the Courts of Appeals database 7.9% of opinions have dissents (3.6% have concurrences).

2.3 Main variables

The ideological score of a judge, based on the JCS scoring method, is captured by the variable **Score.** To construct a measure of ideological disagreement between a judge and her peers on the panel of a specific case, we use this JCS score to calculate the **Score Relative to Panel Median**, which is positive when the judge is more conservative than the median and negative when the judge is more liberal than the panel median. We also take the absolute value of this variable and refer to it as **Distance to Panel Median**.[13]

Next we construct a measure of ideological disagreement between a judge and her peers in the pool, whom she can expect to meet regularly. To do this we begin by calculating the average ideology score of the pool of judges for each Circuit and each year.[14] This average score represents the center of the pool of judges available to be assigned at that Circuit-year, to which we refer as **Center of Judge Pool**. We then demean the ideology score of a judge by the Center of Judge Pool and get a measure we refer to as **Score Relative to Center of Judge Pool**. This measure can take both positive values (if the judge is more conservative than her peers) and negative ones (if the judge is more liberal than her peers) and serves as our main ideological measure. The value of this measure for a given judge may change during her service as the composition of her pool changes. A histogram of Score Relative to Center of Judge Pool is displayed in the center part of Appendix Figure 1. We refer to the absolute value of this measure as **Distance to Center of Judge Pool**.

[13] 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.

[14] We use Circuit and year since this represents the ideology of the other judges a specific judge expects to sit with in a panel.

Using the data from Openjurist we calculate the **Dissent Rate** and **Concurrence Rate** for each judge in each Circuit-year. These are essentially measures of a judge's tendency to confront her peers.

3 Empirical Results

In this section we present our three main empirical observations. The first two observations essentially provide evidence that our "experimental societies" (different Circuit Courts) reflect the mainstream ideology of their members and that ideology is salient in individual decisions. We do this by (1) showing that the verdict reflects (mainly) the ideology of the median of a panel (Section 3.1) and that extremists have a small impact on the verdict (Appendix A.2.1); and (2) showing that ideological disagreement with the median (Section 3.2) triggers dissent. We then move to our main observation that extremists are non-confrontational by showing that (3) the frequency of dissent—the judge's main tool of confrontation—is first increasing and then decreasing as a function of a judge's Distance to Center of Judge Pool (Section 3.3).

3.1 Who affects the ideological flavor of verdicts?

We check here whose ideology is reflected in the decisions made in our experimental society, i.e. in the verdicts of Circuit Courts. Given the randomization of judges into panels, we have a setting that enables us to causally query, for each panel, who affects the ideological flavor of the judicial outcome. We do this by measuring the correlation between the ideology of each judge on the panel and the ideology of the verdict when the judge is the median of the panel and when not. Our identifying assumption is that unobserved factors are not driving at the same time the ideology of any specific judge and the ideology of the judicial outcome. This assumption holds due to our usage of exogenous measures of ideology (the JCS scores, see Section 2.2) which are determined *prior* to the trials hence are independent of the verdicts, and due to the as-if-random assignment of judges into panels and of panels to cases. To further ensure our identifying assumption holds, we use circuit and year fixed effects. In the appendix we further ascertain that extremists in particular do not affect verdicts, by showing that the verdicts' ideology is the least correlated with these judges'

ideology (Appendix A.2.1).

To conduct the analysis, we run an OLS regression of the Verdict Ideology (from the U.S. Courts of Appeals Database Project) on a judge's Score and its interaction with whether the judge is the median of the panel in terms of this (ideology) score, controlling for the Center of Judge Pool and including Circuit (C_c) and year (T_t) fixed effects:

$$(1)\quad \textbf{Verdict Ideology}_{pict} = \alpha + \gamma_1 \textbf{Score}_i + \gamma_2 1\,(i\,is\,median_{pict}) + \gamma_3 \textbf{Score}_i * 1\,(i\,is\,median_{pict}) + C_c + T_t + \nu_{pict}$$

for judge i on panel p in Circuit c and year t.[15] If the ideology of a judge influences the court's verdict, we should expect a positive relationship between the judge's ideology score (where a high value means a very conservative judge) and the likelihood of a conservative verdict. Table I shows that the effect of a judge on the verdict is more than threefold when the judge is the median of the panel (compare 0.0345 to $0.0345 + 0.0783$).

TABLE I

Ideology of Verdict and Ideology Scores of Panel Members

	(1)
Score	0.0345***
	(0.0124)
Panel Median	0.0025**
	(0.0012)
Score * Panel Median	0.0783***
	(0.0278)
Center of Judge Pool	0.159
Ideology Score	(0.135)
Circuit Fixed Effects	Y
Year Fixed Effects	Y
N	23031
R-sq	0.028

Notes: Robust standard errors clustered at the circuit-year level in parentheses (* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$). Data 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. The dependent variable is verdict ideology, which is coded as 1 for conservative, 0 for mixed or not applicable, and -1 for liberal. Circuit and Year fixed effects are included.

Our first empirical result is thus:

[15] The regression includes only three-judge panels where there are no tied or missing scores. Appendix Section A.2.1 presents robustness checks of this result. Standard errors are clustered at the Circuit-year level since we think of the Circuit at a given year as our "experimental society".

Finding 1: *The ideology of the panel median has the strongest effect on the verdict.*

This result is not very surprising given the pivotal role of the median in a three-judge panel and is consistent with many conventional bargaining models. It also aligns with a recent experimental study of decision making within groups (Ambrus et al. 2015), finding that the median has the greatest impact on the group's joint decision. Ambrus et al. (2015) further find that the most extreme individuals relative to the median have a particularly small effect on the joint outcome, a result that is largely echoed by our further investigation, see Appendix Figure 4. For robustness checks and descriptions of other papers' related results see Appendix A.2.1. In total this shows that in individual interactions within our experimental society (i.e., a Circuit Court), ideologically extreme individuals have a small effect on joint decisions.

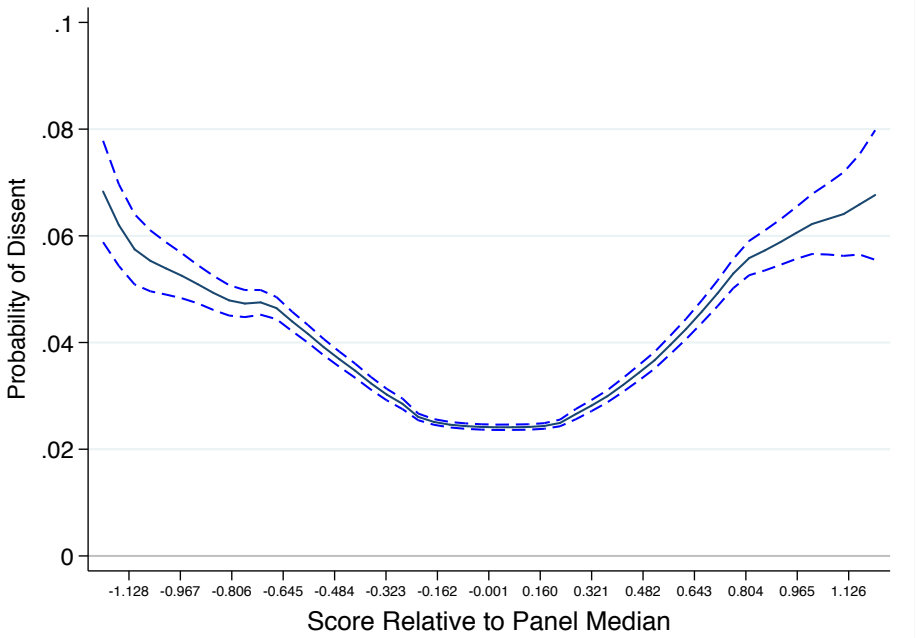
3.2 Is ideological disagreement decisive for when a judge dissents?

We can think of our experimental society—a Circuit Court at a particular period in time—as consisting of many interactions. For a single agent (a judge), each of these interactions (cases) will be with different constellations of individuals (different panels), and in particular with a different judge being at the pivotal role of the panel's median. We now wish to test whether, for the single agent, ideological disagreement is salient in deciding *when* to confront. What we therefore need is exogenous variation in the panel composition (for each judge). This is provided by the judiciary's as-if-random assignment into panels. Then, we can test whether the ideological distance from the judge to the panel's median, who was just shown to have the largest effect on the verdict, positively affects that judge's decision to dissent. If the result of this test is affirmative, it captures that an individual judge concentrates her fighting in situations of ideological disagreement. Our identifying assumption is that unobserved factors are not associated at the same time with the distance from a judge to the median of the panel and with the general tendency to dissent. Apart from the as-if randomization of judges into panels and the use of an ideology measure determined prior to circuit-court interaction, we address this concern by the inclusion of Circuit and

year fixed effects. Since we are interested in studying the variation at the single-judge level, we also include judge fixed effects.

We proceed now to investigating this within-judge property regarding when a judge will choose to dissent. Figure 2 presents a non-parametric visualization of the Dissent Rate by ideology Score Relative to Panel Median using a local polynomial regression (see Altman 1992 and Fan and Gijbels 1996). This figure reveals a clear pattern: the more a judge is distant from the panel median, the more likely she is to dissent. This holds both on the left and on the right. We present similar non-parametric visualizations for concurrences in Appendix Figure 5.

FIGURE 2.— Dissent and Ideology Score Relative to Panel Median – local polynomial



Notes: x-axis: Ideology score of a judge demeaned by the median of the panel of judges assigned on the case. y-axis: Rate of dissent. The figure presents a local polynomial regression with an Epanechnikov kernel, where the dependent variable is rate of dissent. The dashed lines depict the 95% confidence interval. Data come from OpenJurist (1950-2007). Ideology scores come from the Judicial Common Space database.

To test it in a regression specification, we regress (using a linear probability model) the Dissent and Concurrence rates of each judge-case combination on polynomials of the judge's Distance to the Panel Median—an absolute value that captures the strength of ideological disagreement of the judge with the most influential member on the panel. We also add judge fixed effects (I_i) to ensure that the result is not driven by the ideology scores of judges per

se. We further include Circuit (C_c) and year (T_t) fixed effects and cluster the standard errors at the judge level. The basic regression specification is:

$$\text{(2)} \quad \begin{aligned} &\textbf{Dissent Rate}_{pcit} = \gamma_1 \textbf{Distance to Panel Median}_{pcit} + \\ &\gamma_2 \textbf{Distance to Panel Median}^2{}_{pcit} + I_i + C_c + T_t + \nu_{pcit} \end{aligned}$$

for judge i on panel p in Circuit c at year t. Table II indicates that the frequency of both dissents and concurrences increases in the distance to the panel median, implying that ideological disagreement is a driver of dissent. Table A.4 in the appendix shows the result is robust to using a logistic specification instead.

TABLE II

Dissent and Ideological Distance to Median of Panel

	(1) Dissent	(2) Concur
Distance to Median of Panel	0.0260***	0.00793**
	(0.00542)	(0.00318)
Distance^2	0.0111	0.00878*
	(0.00737)	(0.00459)
Judge Fixed Effects	Y	Y
Circuit Fixed Effects	Y	Y
Year Fixed Effects	Y	Y
N	541163	541163
R-sq	0.019	0.018

Notes: Robust standard errors clustered at the judge level in parentheses (* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$). Data on cases comes from OpenJurist (1950-2007). Ideology scores come from the Judicial Common Space database. Ideology scores are demeaned by the actual center of the panel of judges assigned on a case. The dependent variable is a dummy for whether a judge dissented (column 1) or concurred (column 2) in the panel. Fixed effects include year, circuit, and judge.

Based on these results our second empirical finding is:

Finding 2: *A judge is more likely to dissent when the panel median is ideologically far from her.*

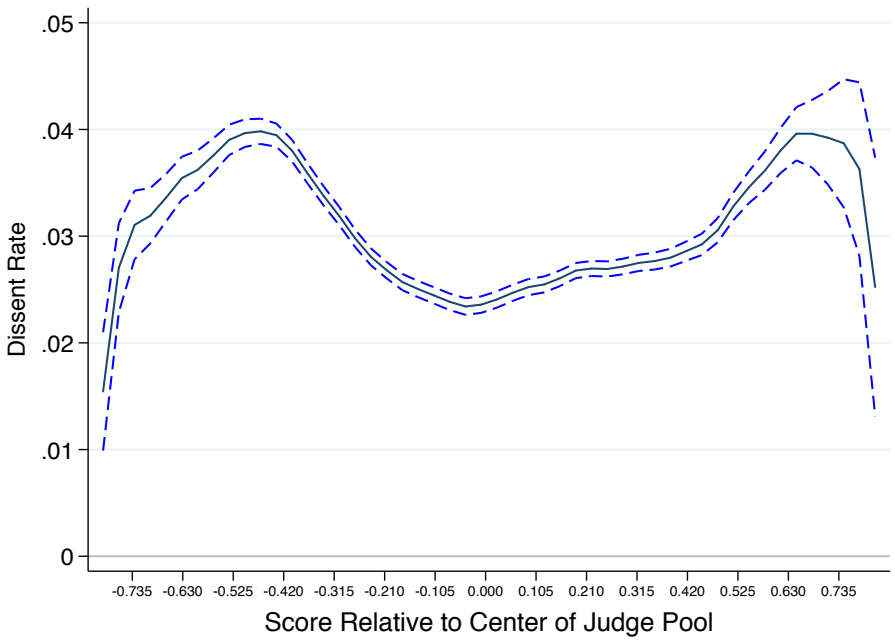
This result, that for the individual judge ideological disagreement with her panel peers is an important determinant for when she will dissent, is in line with previous observations in the literature (e.g., Wahlbeck et al. 1999, Spriggs et al. 1999 and Hettinger et al. 2004).

3.3 Which judges are dissenting the most?

We move now to our main research question: are extremists more confrontational than others? To answer this question, instead of looking at single interactions (as in the previous section) we now look at choices made over many interactions and how these are affected by the society one is operating in. The ideal setting for testing whether operating in a society where one is ideologically different from its mainstream would be a randomized control trial that randomizes individuals into long-term cultural norms. The setting we have is a judge who is appointed to a Circuit where the cultural norm is represented in ideological space by the average ideology of the judges in the Circuit. It differs from the ideal in that judges are not randomized into a Circuit: Appointments are possible whenever previous judges retire. Our identifying assumption is thus that the decision to retire or accept an appointment is exogenous to the average ideology in the Circuit (see further description in Section 2.1). Our main test consists of examining how a judge's ideological distance to the mainstream ideology in the pool in a given year (the societal norm) affects her choice of how often to dissent. One concern would be that unobserved factors are associated at the same time with distance to the group norm and with dissent rate. We address this by including a large number of controls for judge characteristics. We further address this concern by adding a test where we track specifically the pattern of dissent of extreme judges, while exploiting the shift in group norms over the years. That is, we augment our main analysis with one that looks only at the group of judges who are at times throughout their career extreme with respect to the Circuit and at other times not. To this test we add judge fixed effects. Ideally, all judges would have been placed in different societies so that each one would be in ideological minority at some point. In practice this does not happen, as the composition of judges within a circuit shifts rather slowly. We are thus not able to use fixed effects for all judges, only for those who are at some point extreme. Another concern, given our focus on extreme judges, is measurement error in ideology. We address this concern by checking the robustness of our results when using another measure of being in ideological minority – the share of judges in a circuit that were appointed by a President from a different party than the judge's appointing President.

Starting with our main test, Figure 3 presents a non-parametric visualization of Dissent Rate by the ideology Score Relative to Center of Judge Pool using a local polynomial regression. Note that this is mainly a between-judge comparison. As explained earlier, for each Circuit at each year, each judge's score was recalculated to reflect her location relative to the average ideology in the pool of judges in that circuit and year. That is, Score Relative to Center of Judge Pool measures the ideological disagreement between a judge and all the peers she regularly interacts with.

FIGURE 3.— Dissent and Ideology Score Relative to Center of Judge Pool - local polynomial



Notes: x-axis: Ideology score of a judge demeaned by the center of the *pool* of judges available to be assigned in a Circuit-year. y-axis: Rate of dissent. The figure presents a local polynomial regression with an Epanechnikov kernel, where the dependent variable is rate of dissent. The dashed lines depict the 95% confidence interval. Data come from OpenJurist (1950-2007). Ideology scores come from the Judicial Common Space database.

Intuitively, given Finding 2, one might expect that the more distant a judge is from the pool center, the more she will be inclined to write separate minority opinions (dissent or concur). But Figure 3 reveals a surprising pattern: starting from the left, the most extreme judges rarely dissent, then there is a marked increase in dissent as judges become more moderate, followed by a decrease in dissent rates towards the center of the judge pool. A similar pattern appears on the right. We will refer to this pattern as a *spider pattern*, due

to the figure's resemblance of the body and legs of a spider. Appendix Figure 6 shows that the spider pattern is robust to residualizing by circuit and year fixed effects and presents a similar graph when grouping judges into 15 separate bins according to Score Relative to Center of Judge Pool.[16]

As further evidence for the spider pattern, Appendix Figure 6 presents the Concurrence Rate of judges according to Score Relative to Center of Judge Pool (local polynomial and bins).[17] Notably, the pattern of the spider is robust: concurrence rates are surprisingly uncommon for the most extreme judges, the rate is higher for judges at moderate distance to the center and lower again for the judges at the very center (though here the pattern is less clear on the right). Finding these patterns for dissents and concurrences separately of course strengthens our confidence in these results.

To test it in a regression specification, we regress the dissent and concurrence rate of each judge on polynomials of her distance to the center of the pool of judges in her Circuit-year:

$$\begin{aligned}\textbf{Dissent Rate}_{cit} = \alpha + \gamma_1 \textbf{Distance to Center of Judge Pool}_{cit} + \\ \gamma_2 \textbf{Distance to Center of Judge Pool}^2{}_{cit} + C_c + T_t + \nu_{cit}\end{aligned} \tag{3}$$

for judge i in Circuit c and year t. Circuit and year fixed effects are represented by C_c and T_t and standard errors are clustered by Circuit-year. In other words, this model specification asks if a judge's yearly dissent rate is associated with the judge's distance to the center of the judge's Circuit pool in that year and whether the nature of this association is quadratic (we also confirm robustness to longer time periods, see below). We use a quadratic relationship because this pattern is suggested by the raw data, and we control for Circuit and year fixed effects because our treatment of interest varies at the Circuit-year level. Table III indicates that the spider pattern is robust: according to the estimated linear and quadratic

[16] Note that the share of judges having the same score (or range of scores) has no effect on the pattern of dissent rate. The pattern depicts the average dissent rate for each score, regardless of how many judges have this score, hence there is no distorting effect of weighting by a variable number of judges.

[17] We analyze concurrences separately from dissents in order to show the robustness of our results and because they are legally distinct. However, as both require writing a separate minority opinion, we bind them together in the further robustness checks reported in the appendix, thus treating them as two alternative manifestations of the same thing: a judge's decision to confront her panel's majority opinion.

coefficients in the table, the maximum dissent rate is obtained for Distance to Center of Judge Pool of 0.6 (for dissents) and 0.46 (for concurrences) which are clearly within the bounds of our distribution, which goes from around -0.8 to +0.8.[18] Plugging these values of the locations of the peaks into the regression equation, we get that the rate of dissent is estimated to increase by 1.2 percentage points when moving from the center to the peak, and then to drop back by 0.1 percentage points,[19] while the rate of concurrences is estimated to increase by 0.65 percentage points when moving from the center to the peak, and then to drop again by 0.62 percentage points. Given the likely underestimation that results from using a proxy score like the one we use (see Fischman and Law 2009) and the underlying small rate of dissents and concurrences, these drops in the rates are not negligible.[20] The visualization in Figure 3, where the dissent rate estimated at the peaks is twice the size of that estimated at the extremes, is another indication of the sizable drop of the dissent rate at the edges.

TABLE III

Dissent and Ideological Distance to Center of Judge Pool

	(1)	(2)
	Dissent	Concur
Distance to Center of Judge Pool	0.0404***	0.0285***
	(0.00756)	(0.00570)
Distance^2	-0.0334***	-0.0313***
	(0.0118)	(0.00862)
Circuit Fixed Effects	Y	Y
Year Fixed Effects	Y	Y
N	10043	10043
R-sq	0.109	0.086

Notes: Robust standard errors clustered at the circuit-year level in parentheses (* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$). Data on cases comes from OpenJurist (1950-2007). Ideology scores come from the Judicial Common Space database. The main independent variable is (the absolute value of) ideology score demeaned by the center of the pool of judges available to be assigned in a Circuit-year. The dependent variable is the judge's dissent rate (column 1) or concurrence rate (column 2) in a Circuit-year. Fixed effects include year and circuit. Observations are weighted by the number of votes cast by the judge in the Circuit-year.

[18] See Appendix Figure 1 for the distribution of Score Relative to Center of Judge Pool.

[19] The number 1.2 comes from using the Distance to Center of Judge Pool at the peak (0.6) and the coefficients from Table III, which yields $0.0404 * 0.6 - 0.0334\,(0.6)^2 \simeq 0.012$. The number 0.1 is then the difference (in percentage points) between this value of 0.012 at the peak and the value of $0.0404 * 0.8 - 0.0334\,(0.8)^2 \simeq 0.011$ at the end of the distribution, where Distance to Center of Judge Pool equals 0.8.

[20] We remind the reader that 8.5% of opinions in Openjurist have dissents and 6.4% have concurrences. After dividing by three (the number of judges in a panel), we get that the average rates of dissent and concurrence for a judge are both in the range of 2%-3%.

To verify that the spider pattern is not driven by some distortion in the ideology score, we ran the same regression using an alternative ideology score—the party of appointing President. This score, in its raw form, does not distinguish between judges nominated by different Republican Presidents (all have a score of 1). Likewise it does not distinguish between judges nominated by different Democrat Presidents (all have a score of 0). To create a relative score based on this raw score we calculate the average ideology in the pool in a circuit-year and calculate a judge's distance to this average.[21] Hence, the relative score reflects whether a judges is in minority (a large absolute score) or in majority (a small absolute score). Table IV shows the spider pattern is robust to using this alternative score. Furthermore, this score provides a very straightforward reinterpretation of the spider pattern. An "extremist" in this scoring system is simply a judge placed in a pool of judges in which a large majority of the judges were nominated by a President from the other party. Hence, the spider pattern when using this score clearly captures the tendency of judges to stay silent when they are in ideological minority, while behaving the most confrontationally when having a similar number of peers from their own party and from the other party.

[21] For instance, a judge nominated by a Democrat President (that is, score 0) in a circuit-year consisting of a total of 8 Republican-nominated and 2 Democrat-nominated judges will get a relative score of $0 - \frac{(8*(+1)+2*0)}{10} = -0.8$. A Republican-nominated judge in that same pool will get a relative score of $1 - \frac{(8*(+1)+2*0)}{10} = 0.2$.

TABLE IV

Dissent and Alternative Ideology Score

	(1)	(2)
	Dissent	Concur
Distance to Center of Judge Pool	0.0501***	0.0284***
Score Based on Party of Appointment	(0.0157)	(0.0107)
Distance^2	-0.0367**	-0.0222*
	(0.0164)	(0.0115)
Circuit Fixed Effects	Y	Y
Year Fixed Effects	Y	Y
N	10033	10033
R-sq	0.106	0.085

Notes: Robust standard errors clustered at the circuit-year level in parentheses (* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$). Data on cases comes from OpenJurist (1950-2007). Ideology scores are simply the party of appointment (Republican or Democrat, coded as 1 and 0). The main independent variable is (the absolute value of) ideology score demeaned by the center of the pool of judges available to be assigned in a Circuit-year. The dependent variable is the judge's dissent rate (column 1) or concurrence rate (column 2) in a Circuit-year. Fixed effects include year and circuit. Observations are weighted by the number of votes cast by the judge in the Circuit-year.

This easy-to-interpret score further enables us to quantify and interpret the effect we find. Starting with the dissent rate (column 1), we can compute that the peak is at 0.68. This means that, for a conservative judge, dissent rate increases by 1.7 percentage points when moving from an all-conservative pool to a pool that has 68% liberals, and then drops by 0.36 percentage points in an all-liberal pool (i.e., drops by roughly one-fifth of the increase).[22] Proceeding to the concurrence rate (column 2), we can compute that the peak is at 0.64. This means that, for a conservative judge, concurrence rate increases by 1.8 percentage points when moving from an all-conservative pool to a pool that has 64% liberals, and then drops by 1.18 percentage points in an all-liberal pool (i.e., drops by about two-thirds of the increase).

Additional robustness checks are reported in Appendix Tables A.5 to A.7. These address concerns of unobserved factors driving the results. Appendix Table A.5 shows that the results are robust to weighting each judge equally (rather than according to the number of votes); dropping visiting judges; and using higher levels of aggregation (e.g., two-year or lifetime rates, where in the latter there is one observation per judge). They are also robust to

[22] The number 1.7 comes from using the Distance to Center of Judge Pool at the peak (0.68) and the coefficients from Table IV, which yields $0.0501 * 0.68 - 0.0367\,(0.68)^2 \simeq 0.017$. The number 0.37 is then the difference (in percentage points) between this value of 0.017 at the peak and the value of $0.0501 * 1 - 0.0367\,(1)^2 \simeq 0.0134$ at the end, where Distance to Center of Judge Pool equals 1 (an all-liberal pool).

including a cubic term of Distance to Center of Judge Pool, running Logit, including time-varying judge tenure interacted with a judge's ideology score and controlling for Distance to Supreme Court or Distance to Median of Panel (i.e., the pattern is not driven by interactions within particular panels). In Appendix Table A.5, we also show that when the Distance to Center of Judge Pool is randomly re-assigned to another observation there is no relationship, which further mitigates the concern that the documented relationship is a statistical artifact or is driven by the level of clustering. Appendix Table A.6 shows that the results are robust to dropping one Circuit at a time, thus mitigating the concern that the pattern is driven by a particular court or by outliers. Appendix Table A.7 (columns 1 and 2) controls for a large set of personal characteristics of judges (which can be seen as a second-best alternative to using judge fixed effects—see more on judge fixed effects below). Appendix Table A.7 (columns 3 and 4) re-runs the main quadratic specification in equation (3) but using two forms of bootstrapping for extreme judges (since there are fewer such observations) and corroborates that the results (and their statistical significance) do not change. Finally, Appendix Table A.7 (columns 5 and 6) shows that the pattern is robust to splitting the sample according to whether the case affirmed or reversed the lower court decision. Based on these results we formulate the following empirical finding:

Finding 3: *There is a hill-shaped relationship between a judge's dissent rate and the absolute distance to her pool center.*

This result expresses that judges who are extreme relative to their greater group of peers are less confrontational than more moderately-distanced judges. It is novel to this paper and, to our knowledge, no paper in any domain has ever examined how being in ideological minority affects one's tendency to confront decisions of the majority. The result is particularly surprising given the earlier result that a judge who is in ideological minority rarely affects the verdict (Finding 1) and given that, at the level of a single judge, confrontation is driven by ideological disagreement with this verdict (Finding 2). Yet extreme judges remain silent. In Appendix A.1.3 we show an important apparent implication of Finding 3: the silencing

of extreme judges implies they end up having a less ideologically biased pattern of signed verdicts than moderates do.[23]

We wish to emphasize the following. Finding 3 is driven by a judge's ideology relative to her peers—the same judges who exhibit the low dissent rate when they are extremists, exhibit a high dissent rate when they are moderates relative to their pool. To further investigate the validity of this assertion, we proceed now to examine how the behavior of a given extremist judge changes when mainstream ideology of the circuit changes, i.e. by including judge fixed effects. Table V reports the coefficients of a regression of dissent rate on polynomials of Distance to Center of Judge Pool with judge fixed effects, using a subsample that contains all the judges who, at a certain point of their career, have Distance to Center of Judge Pool greater than 0.6 (the location of the hump of regression specification 3, see Table III). Table V clearly shows that the spider pattern holds for that subsample of "occasionally extreme" judges, indicating that these judges do not remain silent when they are placed in a pool that is ideologically closer to their bliss point (as reflected by the increasing part of the spider).[24] Moreover, as can been seen in Appendix Figure 7, the spider pattern is profoundly attenuated when considering the raw (i.e., not relative) Ideology Score. Hence, we can conclude that the non-confrontational behavior of extreme judges is not driven by their ideology per se or by some related personal characteristics but rather by regularly being in strong ideological disagreement with most of their peers—ideological interaction silences extremists.[25]

[23] Our main model (Section 5) can account for this finding too—see Appendix B.1.

[24] For judges who were never "extreme" at some point of their career, using judge fixed effects renders more measurement error (see Greene, 2018) and less meaningful variance to estimate the spider pattern.

[25] Recall also the reinterpretation of extremism as belonging to a partisan minority when using the alternative party-of-appointing-President score.

TABLE V

Dissent and Ideological Distance to Center of Judge Pool for "Extreme" Judges, Including Judge Fixed Effects

	(1)	(2)
	Dissent	Concur
Distance to Center of Judge Pool	0.0971***	0.102***
	(0.0334)	(0.0291)
Distance^2	-0.102***	-0.105***
	(0.0388)	(0.0303)
Circuit Fixed Effects	Y	Y
Year Fixed Effects	Y	Y
Judge Fixed Effects	Y	Y
N	1519	1519
R-sq	0.425	0.343

Notes: Robust standard errors clustered at the circuit-year level in parentheses (* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$). Data on cases comes from OpenJurist (1950-2007). Sample of judges who are extreme relative to their pool at some point of their career (Distance to Center of Judge Pool greater than 0.6 (the location of the hump of regression specification). Ideology scores come from the Judicial Common Space database. The main independent variable is (the absolute value of) ideology score demeaned by the center of the pool of judges available to be assigned in a Circuit-year. The dependent variable is the judge's dissent rate (column 1) or concurrence rate (column 2) in a Circuit-year. Fixed effects include year, circuit, and judge. Observations are weighted by the number of votes cast by the judge in the Circuit-year.

4 Explaining the empirical observations

How can these puzzling empirical observations be explained? In our quest to understand the observations we have considered 16 separate explanations representing the "usual suspects" within research on judicial and group decision making.[26] After analyzing each explanation separately we are left with two separate explanations (or models) that can explain all three findings and do not violate some simple auxiliary empirical observations. One model is about judges using dissents instrumentally to awake the interest of the Supreme Court who may then reverse the verdict, and the other model is about judges using dissents to morally or emotionally distance themselves from verdicts they do not agree with ideologically. We take these two models and derive additional empirical predictions by which they differ and test them empirically (see Appendix D.1.15).[27] These tests favor one of the models (the moral/emotional distancing), which is hence presented below in the main text, while the other model (instrumental dissents) is presented in great detail in Appendix D.2. We

[26] We are grateful to the many researchers, seminar audiences and practitioners who have suggested explanations to consider.

[27] We use wartime, retired judges and newly appointed judges for these auxiliary checks.

wish to emphasize that, while we have considered a large number of explanations and mechanisms, we cannot conclude that our main contender is the only or the true explanation for the empirical patterns observed. Just like when matching any theory to data, regardless of how many theories one has considered, one cannot exclude the possibility that yet another theory that is consistent with the investigated empirical pattern exists.

Before presenting what we consider to be the main contender for explaining the findings, let us briefly describe some of the other 14 explanations (for more details see Appendix D.1). These explanations can be divided into a few categories (many explanations fit under more than one category and indeed most fail for two, three or even four reasons). One category consists of mechanisms that seem reasonable but ultimately fail to explain the main Finding 3 , part of this finding or some of the other findings. It should be noted that our aim here is not just to explain why extremists dissent seldom but the full hill-shaped pattern of dissent when going from centrists to extremists and also to fit the other two findings. Mechanisms that fit into this category are, for instance, that extremists wish to signal that they are not extreme (D.1.1), or that moderates wish to signal that they are extreme to please some extreme electorate (D.1.2).

Another category consists of explanations that are inconsistent with other empirical observations. In particular, any explanation that suggests that extreme judges do not dissent because they manage to pull the decision in their direction is inconsistent with auxiliary data. This includes, for instance, mechanisms saying that extremists, by bargaining well within the panel (D.1.3) or by getting co-panelists to agree with them in later cases (D.1.4) or by including facts that tie the hands of the Supreme Court (D.1.5), manage to affect the verdict and the majority opinion disproportionally. Such explanations are inconsistent with observations presented in Appendix Figure 4 showing that, at both ends of the ideological spectrum, there is a negative correlation between a judge's ideology and the verdict. Another refuted mechanism is that moderates dissent more than the others since they are spokesmen representing the party line hence are more vocal (D.1.6).

Yet another category consists of mechanisms that, upon closer inspection, either resemble or necessitate the mechanism of our main-contender theory in order to be able to

explain the empirical findings. These mechanisms include that the peer pressure when dissenting is higher if the minority opinion is extreme (D.1.7), that extreme judges face a higher risk of impeachment (D.1.8) or that extreme minority opinions get fewer citations hence are not worth the effort of writing (D.1.9).

Finally, we also consider potential threats to our empirical variables or results. Many of these are also described as part of the robustness checks in the previous (empirical) section, but in the appendix (D.1.10-D.1.14) we discuss further grounds for why some empirical concerns are not likely to be the reason for our empirical patterns.

5 Theoretical Model

We now outline a theoretical model that can qualitatively explain Findings 1-3. We keep it as simple as possible in order to highlight the mechanism.[28] The model is inspired by, but is a simplified version of, Beim et al. (2014) and Cameron and Kornhauser (2009).

There is a large *pool of judges* (formally, a continuum of infinitesimal judges with a unit mass). Each judge has an ideology score t which is public information. The judges' ideologies are drawn from a symmetric distribution F in $[-1, 1]$. There is a long sequence (a continuum) of judicial cases. For each case, three judges (t_1, t_2 and t_3) are randomly and independently drawn from the pool of judges to sit together on a *panel*.

The timing of the actions within each case is as follows. First, a case with characteristic y arrives, where y is drawn from the same distribution F.[29] Second, the three judges in the panel vote, in random order with majority rule, on whether the verdict $v(y)$ should be liberal ($v(y) = 0$) or conservative ($v(y) = 1$).[30] Letting $v_t(y)$ denote judge t's ideologically preferred verdict in a case with characteristic y, it is conservative ($v_t(y) = 1$) if $y < t$ and otherwise liberal ($v_t(y) = 0$). Third, each judge decides whether to sign the verdict or not, where not signing implies dissenting. Let $s(v(y); t)$ be an indicator function that equals 1 if judge t

[28] We can change it in a multitude of ways without changing the main results qualitatively. For instance, we abstract from the possibility that some cases are non-ideological. Adding such a component would generally smooth out the judges' actions so that the dissent rate of judges with different ideologies would be more similar and (to the extent that judges agree on such factual cases) lower the dissent rate.

[29] A simple special case is where both t and y are uniformly distributed in $[-1, 1]$.

[30] We use the word *vote* in the theory to represent deliberation within a panel. In practice this action is empirically unobservable.

chooses to sign verdict $v(y)$ and equals 0 if she dissents. We assume (for simplicity) that the set of panel peers faced by judge t during her career is known in advance and is identical to the continuous distribution of two random judges drawn from the distribution of ideologies $F(t)$.

Whenever a judge disagrees with the verdict, she has to trade off the inner discomfort associated with actively approving the verdict by signing it, and the collegial pressure when choosing instead not to sign. The objective of the judge is to minimize the sum of these two costs. The inner discomfort is captured by the following cost function

$$D(x) = x^{\alpha}, \tag{4}$$

where

$$x \equiv \int_{-1}^{1} I(v_t(y) \neq v(y)) s(v(y); t) f(y)\, dy, \tag{5}$$

and $I(v_t(y) \neq v(y))$ is an indicator function taking the value of 1 when the judge disagrees with the verdict. We assume $\alpha > 0$ so that D increases in x, which means the ideological cost is increasing in the number of verdicts a judge signs yet disagrees with.[31] D can be interpreted in two main ways: either as a judge's own perceived cost of being part of a ruling she disagrees with; or as a loss when not standing up for the ideology of one's supporters or ideological faction. Letting this disutility depend on all cases captures the idea that judges are optimizing over their whole career. Put differently, the judge's reference point (or ambition) is signing only verdicts she agrees with. Then she perceives a cost based on the number of

[31] The inner discomfort function does not take into account how much the judge disagrees with the verdict, just whether she disagrees. We have also solved a more complex version of the model, available upon request, where judges care not only about whether the binary verdict misaligns with their preferences but also about by how much it misaligns. The main qualitative results stay the same, in particular the conclusion that the inner discomfort has to be concave in order to produce the spider pattern of dissent. Since that alternative model is more elaborate, it allows for a smoother spider-pattern of dissent, in particular when the D function is mildly concave ($\alpha \in [1/3, 2/3]$).

deviations from this reference point.[32] In total, a judge t has the loss function

$$L = D(x) + WP(t), \tag{6}$$

where

$$P(t) \equiv \int_{-1}^{1} (1 - s(v(y); t)) f(y) dy$$

is the judge's rate of dissent, and the last term in the loss function represents the cost arising from collegial pressure. W can also be interpreted as capturing the effort of writing a separate dissenting opinion. We let this cost be linear in the rate of dissent for tractability, thus W is the constant marginal cost of dissent. This simplifying assumption does not drive the qualitative results of the model.[33]

5.1 Model results

We will now present the model results and how they relate to the empirical findings. We focus on stationary strategies, i.e., where judges' actions are not history dependent. Quite clearly, a judge is going to sign all verdicts she agrees with, because doing so inflicts no cost while not signing results in collegial pressure. The crux of the analysis therefore lies in investigating which portion of the unfavorable verdicts $v(y) \neq v_t(y)$ a judge with ideology t chooses to sign and how this translates to the total dissent rate.

Findings 1 and 2 are, nearly trivially, predicted by the model. First, denoting the bliss point of the median judge of the panel by t_m and her ideologically preferred verdict for a case with characteristic y by $v_m(y)$, Finding 1 (the median's ideology has the strongest effect on the verdict) follows from the striving of each judge to minimize the number of verdicts she disagrees with.

[32] It should be noted that if a judge would optimize case by case (that is, the power would be applied to each deviation) then our main theoretical conclusions would still hold. However, an additional assumption would be needed, namely, that the collegial pressure (see below) is increasing in the extremeness of the judge. We prefer to avoid this extra assumption and we also find the assumption of case-by-case optimization less realistic, but we briefly refer to this alternative in Appendix D.1.7.

[33] That is, the results hold qualitatively also for a convex or a slightly concave collegial pressure.

PROPOSITION 1 ***(Finding 1)*** *There exists an equilibrium with stationary strategies. In any such equilibrium,* $v(y) = v_m(y)$ *in all panels.*

PROOF: See Appendix C.1. *Q.E.D.*

Second, given the voting outcome, it is clear that signing is always optimal for the median judge in the panel. As for the other judges in the panel, who might choose to dissent, the following proposition characterizes the relation between their choice of when to dissent and the ideology of the median judge in the panel.

PROPOSITION 2 ***(Finding 2)*** *In any equilibrium with stationary strategies and for any single judge* t*, the probability that* t *dissents in a case weakly increases in* $|t - t_m|$.

PROOF: See Appendix C.2. *Q.E.D.*

The model thus predicts that a judge is more likely to dissent the farther away her bliss point is from that of the median in the panel (Finding 2). This is unsurprising given that the chance of disagreeing with the verdict increases when the judge is far from the median.

Denote the actual probability of dissent of judge t by $P^*(t)$. To study the shape of $P^*(t)$ as a function of $|t|$, we first have to know how the frequency of the events that trigger dissent, $v(y) \neq v_t(y)$, changes with $|t|$. The following lemma states the intuitive result that this frequency increases in the extremeness of the judge's ideology.

LEMMA 1 *In any equilibrium with stationary strategies,* $Prob(v(y) \neq v_t(y))$ *increases in* $|t|$.

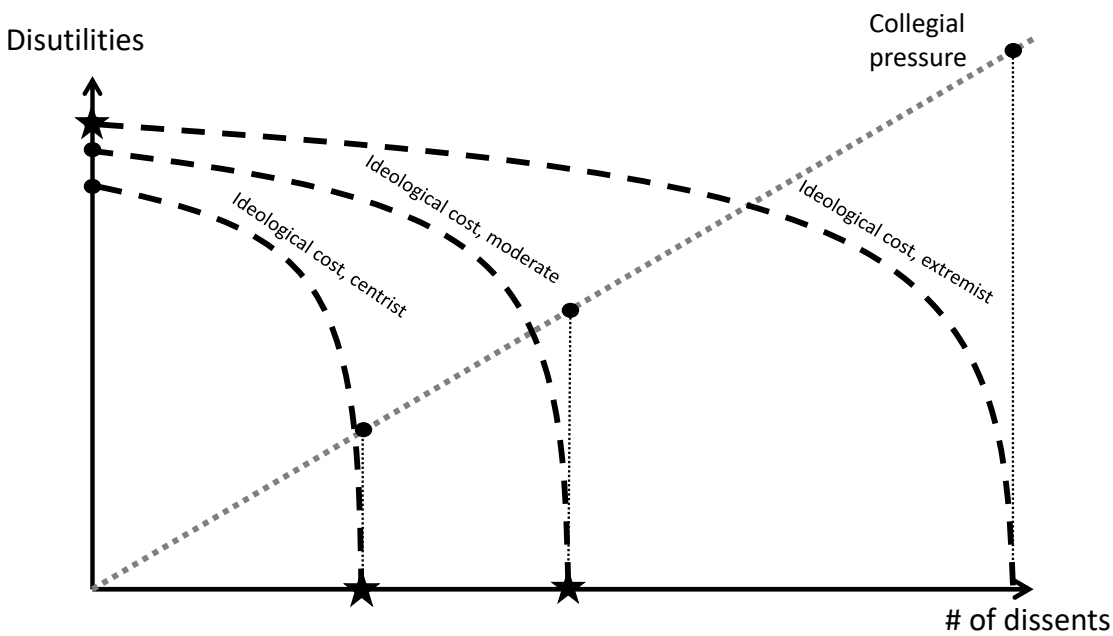
PROOF: See Appendix C.3. *Q.E.D.*

We are now ready to show our main result, namely that the model generates the spider pattern of Finding 3—a hill-shaped relationship between the Dissent Rate and the Distance to Center of Judge Pool.

PROPOSITION 3 ***(Finding 3)*** *Consider an equilibrium with stationary strategies. For any $\alpha < 1$ there exists a range of W for which $P^*(t)$ is hill-shaped in $[0, 1]$ (and has a spider pattern in $[-1, 1]$). If $\alpha \geq 1$, a spider pattern cannot exist in equilibrium.*

PROOF: See Appendix C.4. *Q.E.D.*

FIGURE 4.— Why a concave cost can generate a spider pattern



The proposition states that the ideological cost has to be concave for the spider pattern to hold. We will now explain why.

Figure 4 depicts the optimal choices of each type of judge. On the horizontal axis we see the choice variable (number of dissents) and on the vertical axis the resulting utility loss, which is comprised of two parts. First, there is disutility from the collegial pressure, which is linearly increasing in the number of dissents (the gray dotted line). Second, there is ideological disutility, which equals zero if a judge dissents whenever she is not the median and then increases concavely (when $\alpha < 1$) as the judge lowers her rate of dissent by signing unfavorable verdicts (the black dashed lines). Since judges differ in how often they are median in a panel—where more centrist judges are more likely to be median—they will differ in the number of times they have to dissent in order to fully stick to their ideology. In the figure this implies that the ideological cost (the dashed line) of an extremist is further to the right

than that of a moderate, which is further to the right than that of a centrist – the more extreme a judge is the more opportunities she has to dissent. Consequently, the ideological disutility associated with *never* dissenting (the point where the dashed lines meet the vertical axis) is the largest for an extremist and the smallest for a centrist, but the concavity of the ideological cost means they are not vastly different.[34]

The optimal choices of the three types of judges are depicted in Figure 4 with stars. First note that when D (the dashed line) is concave, there is no point in signing only few unfavorable verdicts. This is since once a judge has signed a few such verdicts the ideological cost of signing more is very low. In order to reduce the ideological cost to any meaningful degree the judge has to never sign any verdict she disagrees with. We call this feature of the concave ideological cost "perfectionism" (as in Michaeli and Spiro 2015). This perfectionism drives judges to either always dissent when they disagree with the verdict (the point where the dashed line meets the horizontal axis), or to never dissent (the point where the dashed line meets the vertical axis).

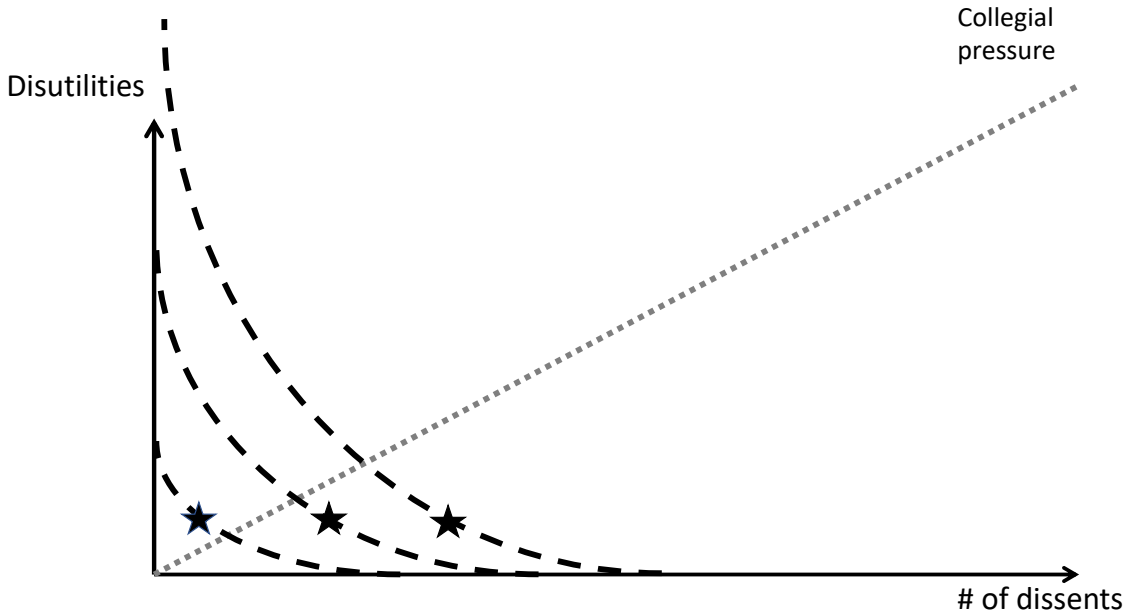
Starting with the centrist judge, she optimizes when choosing to always stick to her ideology because this implies a small collegial pressure and no ideological cost, whereas any compromise she would make (by decreasing her number of dissents) would entail a large ideological cost due to perfectionism. This is true also for the moderate judge in the figure – the collegial pressure she endures for sticking to her ideology is still lower than the cost of never dissenting. However, since the moderate is less often the median of her panel, her choice to stick to her ideology implies she will have a higher dissent rate than the centrist will. This explains why, as a judge moves from being a centrist to a moderate, the dissent rate goes up (the increasing part of the spider leg). Finally, an extremist judge who would stick to her ideology would face a much larger collegial pressure (see the rightmost point on the dotted line). Hence, she would rather give up on her ideology and just bear the ideological cost associated with never dissenting. Thus, the dissent rate falls sharply as a judge becomes sufficiently extreme. Overall, we get a hill-shaped relationship between a judge's dissent rate

[34] In the corner case where the ideological cost is a step function, the ideological disutility associated with never dissenting is the same for all judges.

and how extreme she is relative to her peers in the pool of judges she interacts with over time. This hill-shape appears only if W is sufficiently large as otherwise the collegial pressure would not be strong enough to make anyone silent. However, if W would be too large, no one would dissent as that would be too costly. Hence, W has to be in an intermediate range.

The above theoretical logic does not hold if the personal cost of bliss-point deviation were instead convex, as stated in the proposition and depicted in Figure 5. The reason is that in this case each judge chooses an inner solution to the optimization problem (see the stars in the figure), where it is very costly for an extremist to dissent as seldom as a moderate or a centrist judge given the increasing marginal cost of ideological deviation.

FIGURE 5.— Why a convex cost cannot generate a spider pattern



6 Conclusions and discussion

We study a high-stakes field setting, the U.S. Courts of Appeals, where decisions have an ideological element and judges are repeatedly randomly assigned into panels of three. In this setting, we present a consistent and robust set of evidence that suggests that judges with non-mainstream world views ("extremists") are less confrontational, despite being less likely to affect the outcome of the case. Our findings further show that the results are not driven by having extreme ideology per se but rather by being extreme relative to the people

one interacts with. Hence, interaction between individuals who disagree ideologically silences those who regularly find themselves far from the mainstream. This may hide undercurrents of dissatisfaction, distort the perception of the distribution of views and create false impressions of consensus where it is absent. To the extent that the results generalize to other settings where ideology is salient, our findings may have important implications, e.g., for the expected behavior of immigrants vis-á-vis native society, for individual differences in confronting social and religious norms and for which factions of society are expected to publicly question the consensus. More narrowly, our results also contribute to the discussion of the balance of powers by raising counter-majoritarian concerns among democracy theorists who advocate the premise that judges should reflect the preferences of the executive branch at the time of their appointment without censoring themselves.

To rationalize our empirical observations, we present a simple model of judicial decision making in panels, where dissenting judges are subject to collegial pressure. This model suggests that the non-confrontational behavior of extreme judges is due to the concavity of ideological preferences (ideological perfectionism), which makes them cave in under peer pressure instead of just compromising a little bit, as would have been the case if the ideological cost were convex. Meanwhile, a concave ideological cost induces moderately ideological judges to stand their ground whenever they disagree with a case's outcome and hence appear more confrontational. A large number of alternative models are further considered (see Section 4 and Appendix D.1). In particular, a competing model about dissent as a way of drawing the attention of the Supreme Court is analyzed and rejected empirically. Our theoretical finding thus suggests that the cost of deviating from one's principles should be concave.[35]

Concave ideological costs have far-reaching implications for the empirical predictions of political-economy models. For instance, as discussed by Osborne (1995) and shown by Kamada and Kojima (2014), concavity drives polarization in political platforms. Furthermore, Eguia (2013) shows that a concave ideological cost affects socially optimal policy in that,

[35] At least for sufficiently large deviations—see earlier footnote on how our results generalize to S-shaped functions and relate to Lim (2013).

when trying to bridge differences of opinion, disagreeing agents should each get to decide on a subset of issues rather than compromise within each debated issue. The concavity of costs associated with deviations from a moral or political bliss point also affects the sustainability of biased norms (Michaeli and Spiro, 2017); the pattern of revolutions (Michaeli and Spiro 2016); and the likelihood of conflict (Gratton and Klose 2017). Given this wide array of applications, more empirical research into the curvature of moral and ideological costs in various field settings is warranted.

References

Altman, N. S. (1992). An introduction to kernel and nearest-neighbor nonparametric regression. *The American Statistician 46*(3), 175–185.

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.

Ariely, D. (2012). *The (Honest) Truth About Dishonesty.* New York: Harper Collins Publishers.

Baumeister, R. F. and T. F. Heatherton (1996). Self-Regulation Failure: An Overview. *Psychological Inquiry 7*(1), 1–15.

Beim, D., T. S. Clark, and B. E. Lauderdale (forthcoming). Republican-majority appellate panels increase execution rates for capital defendants. *The Journal of Politics.*

Beim, D., A. V. Hirsch, and J. P. Kastellec (2014). Whistleblowing and compliance in the judicial hierarchy. *American Journal of Political Science 58*(4), 904–918.

Berdejó, C. and D. L. Chen (2014, December). Electoral Cycles among U.S. Courts of Appeals Judges. Working paper, ETH Zurich.

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

Bernheim, B. D. (1994). A theory of conformity. *The Journal of Political Economy 102*(5), 841–877.

Boyd, C., L. Epstein, and A. D. Martin (2010). Untangling the causal effects of sex on judging. *American Journal of Political Science 54*(2), 389–411.

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.

Carrubba, C., B. Friedman, A. D. Martin, and G. Vanberg (2007). Does the median justice control the content of supreme court opinions? In *2nd Annual Conference on Empirical Legal Studies Paper.*

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

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

Chilton, A. S. and M. K. Levy (2015). Challenging the randomness of panel assignment in the federal courts of appeals. *Cornell L. Rev. 101*, 1.

Clark, T. S. and B. Lauderdale (2010). Locating supreme court opinions in doctrine space. *American Journal of Political Science 54*(4), 871–890.

Clinton, J., S. Jackman, and D. Rivers (2004). The statistical analysis of roll call data. *American Political*

Science Review, 355–370.

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.

Edwards, H. T. and M. A. Livermore (2008). Pitfalls of empirical studies that attempt to understand the factors affecting appellate decisionmaking. *Duke Law Journal 58* (8), 1895–1989.

Eguia, J. X. (2013, January). On the spatial representation of preference profiles. *Economic Theory 52* (1), 103–128.

Epstein, L., W. M. Landes, and R. A. Posner (2011). Why (and when) judges dissent: A theoretical and empirical analysis. *Journal of Legal Analysis 3* (1), 101–137.

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.

Fan, J. and I. Gijbels (1996). *Local polynomial modelling and its applications: monographs on statistics and applied probability 66*, Volume 66. CRC Press.

Fischman, J. B. (2011). Estimating preferences of circuit judges: A model of consensus voting. *Journal of Law and Economics 54* (4), 781–809.

Fischman, J. B. and D. S. Law (2009). What Is Judicial Ideology, and How Should We Measure It? *Washington University Journal of Law and Policy 29*, 133–214.

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.

Gino, F., M. I. Norton, and D. Ariely (2010a). The counterfeit self the deceptive costs of faking it. *Psychological science*.

Gino, F., M. I. Norton, and D. Ariely (2010b, May). The Counterfeit Self: The Deceptive Costs of Faking It. *Psychological Science 21* (5), 712–720.

Gneezy, U., B. Rockenbach, and M. Serra-Garcia (2013, September). Measuring lying aversion. *Journal of Economic Behavior and Organization 93*, 293–300.

Gratton, G. and B. Klose (2017). Moderating conflicts with radical hardliners.

Greene, W. H. (2018). *Econometric Analysis*. Pearson Education.

Hall, M. (2010). Randomness reconsidered: Modeling random judicial assignment in the u.s. courts of appeals. *Journal of Empirical Legal Studies 7* (3), 574–589.

Hammond, T. H., C. W. Bonneau, and R. S. Sheehan (2005). *Strategic behavior and policy choice on the US Supreme Court*. Stanford University Press.

Hangartner, D., B. E. Lauderdale, and J. Spirig (2019). Inferring individual preferences from group decisions: Judicial preference variation and aggregation in asylum appeals.

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.

Hettinger, V. A., S. A. Lindquist, and W. L. Martinek (2007). *Judging on a collegial court: Influences on federal appellate decision making*. University of Virginia Press.

Hix, S., A. Noury, and G. Roland (2006). Dimensions of politics in the european parliament. *American Journal of Political Science 50*(2), 494–520.

Ho, D. E. and K. M. Quinn (2010). How not to lie with judicial votes: Misconceptions, measurement, and models. *California Law Review 98*(3), 813–876.

Hurkens, S. and N. Kartik (2009). Would I lie to you? On social preferences and lying aversion. *Experimental Economics 12*(2), 180–192.

Kajackaite, A. and U. Gneezy (2017). Incentives and cheating. *Games and Economic Behavior 102*, 433–444.

Kamada, Y. and F. Kojima (2014). Voter Preferences, Polarization, and Electoral Policies. *American Economic Journal: Microeconomics 6*(4), 203–236.

Keith T. Poole, H. R. (1997). *Congress: A Political-Economic History of Roll Call Voting*. Oxford University Press.

Kendall, C., T. Nannicini, and F. Trebbi (2015, January). How Do Voters Respond to Information? Evidence from a Randomized Campaign. *American Economic Review 105*(1), 322–53.

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.

Kuran, T. (1997). *Private truths, public lies: The social consequences of preference falsification*. Harvard University Press.

Lauderdale, B. E. and T. S. Clark (2012). The supreme court's many median justices. *American Political Science Review 106*(4), 847–866.

Lauderdale, B. E. and T. S. Clark (2014). Scaling politically meaningful dimensions using texts and votes. *American Journal of Political Science 58*(3), 754–771.

Lauderdale, B. E. and T. S. Clark (2016). Estimating vote-specific preferences from roll-call data using conditional autoregressive priors. *The Journal of Politics 78*(4), 1153–1169.

Lim, C. S. H. (2013, June). Preferences and incentives of appointed and elected public officials: Evidence from state trial court judges. *The American Economic Review 103*(4), 1360–1397.

Martin, A. D. and K. M. Quinn (2002). Dynamic ideal point estimation via markov chain monte carlo for the us supreme court, 1953–1999. *Political Analysis 10*(2), 134–153.

Martin, A. D., K. M. Quinn, and L. Epstein (2004). The median justice on the united states supreme court.

NCL rev. 83, 1275.

Michaeli, M. and D. Spiro (2015). Norm conformity across societies. *Journal of public economics 132*, 51–65.

Michaeli, M. and D. Spiro (2016). The dynamics of revolutions. Technical report.

Michaeli, M. and D. Spiro (2017). From peer pressure to biased norms. *American Economic Journal: Microeconomics 9*(1), 152–216.

NYTimes (2017, September 7th). Appeals court limits trump travel ban and allows more refugees. *New York Times*.

Osborne, M. J. (1995, May). Spatial models of political competition under plurality rule: a survey of some explanations of the number of candidates and the positions they take. *The Canadian Journal of Economics 28*(2), 261–301.

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

Schwartz, E. P. (1992). Policy, precedent, and power: A positive theory of supreme court decision-making. *JL Econ. & Org. 8*, 219.

Spriggs, J. F. and T. G. Hansford (2002). The us supreme court's incorporation and interpretation of precedent. *Law and Society Review*, 139–160.

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.

Sunstein, C. R., D. Schkade, L. M. Ellman, and A. Sawicki (2006). *Are Judges Political?: An Empirical Analysis of the Federal Judiciary*. Brookings Institution Press.

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.

Westerland, C. (2003). Who owns the majority opinion? an examination of policy making on the us supreme court. In *Annual Meeting of the American Political Science Association, Philadelphia*.

For Online Publication

A Empirical Appendix

A.1 Data and Additional Empirical Results

We start with a short discussion of ideological measures and their pros and cons. The ideology score we use—the Judicial Common Space (JCS)—is a unidimensional score. Using a unidimensional score to code judges' views on a variety of ideological issues has been argued to be a valid approximation by, e.g., Keith T. Poole (1997); Martin and Quinn (2002); Clinton et al. (2004); Hix et al. (2006), but was criticized by others (e.g. Fischman and Law 2009; Lauderdale and Clark 2012, 2014, 2016), who claim that the ordering of judges can change substantially between different areas of the law. While we acknowledge the criticism made, the common remedies might be easily applicable to the Supreme Court—where the composition of judges is relatively stable, so that there are enough data points for identifying even 24 different issue dimensions (as in Lauderdale and Clark 2016)—but not to our setting of Appeal Courts.[36]

Among the most commonly used unidimensional scoring systems, two particularly stand out: the JCS scores that we use and the Martin-Quinn scores (henceforth MQ, Martin and Quinn 2002). While MQ scores proved to be very popular when applied to studying the Supreme Court, they are much less suitable for studying Circuit Courts for four main reasons. First, the MQ scores measure ideology by checking the alignment of a judge's votes with that of two exogenously determined judges that form the two ends of the ideology scale. While this makes sense in the Supreme Court, where these two judges sit with all other judges in most cases, it is not applicable to the appeal courts, where judges sit in ad-hoc panels of three that are drawn from the much-larger pool of judges. Second, the calculation of MQ scores requires dropping cases with unanimous decisions, but the strong "norm of consensus" in appeal courts implies that most cases will have to be dropped from the calculation, potentially biasing the measurement. Third, while scholars have cautioned against using the cardinality of ideological distances in the MQ space (see e.g. Ho and Quinn 2010), the absolute difference between common space scores as a proxy for ideological distance between judges has been used previously (see e.g. Hettinger et al. 2007). Finally, and importantly, as stressed in the main text, unlike the MQ scores, the JCS scores are predetermined, hence the direction of causality is clear and there is no concern that our ideological measure is contaminated by panel effects.[37]

[36] Indeed, Fischman and Law (2009, p. 164) state that "[a]lthough methods exist for estimating the dimensionality of ideology on multimember courts ... these methods are not applicable to courts in which judges decide cases alone or on rotating panels."

[37] Fischman and Law (2009, p. 149) state that: "One challenge that empirical scholars must address, therefore, is the fact that panel composition effects can conceal the true extent of a judge's ideological leanings."

A.1.1 Summary statistics

APPENDIX TABLE A.1

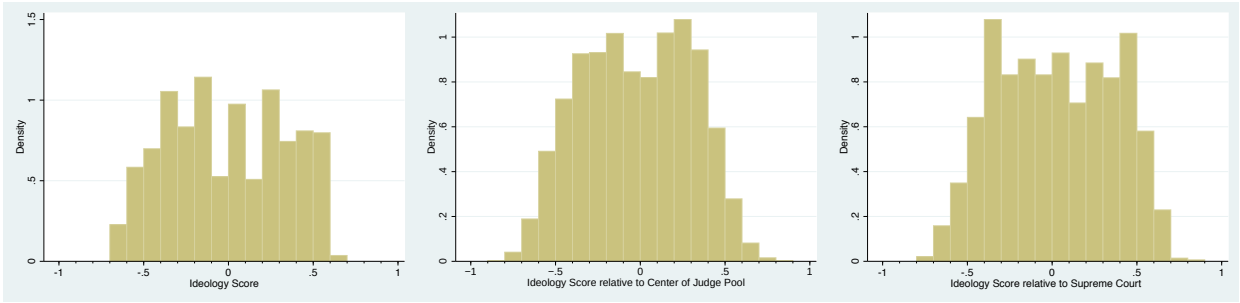
SUMMARY STATISTICS

Vote-Level	Mean	Standard Deviation
Dissent	0.03	0.17
Concur	0.02	0.13
Score Relative to Center of Judge Pool	-0.01	0.30
Distance to Median of Panel	0.18	0.24
Ideology Score	0.01	0.34
N	541182	
Case-Level (Songer-Auburn sample)		
Verdict Ideology (1 = Conservative)	0.19	0.90
Panel Median	0.33	0.47
Score Relative to Center of Judge Pool	-0.01	0.32
Ideology Score	-0.03	0.34
Center of Judge Pool Ideology Score	-0.02	0.16
N	7677	
Circuit-Year Level		
Distance to Center of Judge Pool	0.28	0.17
Distance to Center of Judge Pool based on Party of Appointment	0.46	0.16
Number of Votes Cast Per Judge	65.73	55.88
N	10033	
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 verdict ideology come from the U.S. Courts of Appeals Database Project (1925-2002 5% Songer-Auburn Sample). The sample includes three-judge panels where there are no tied or missing scores. Verdict 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.

A.1.2 Distribution of ideological scores

Appendix Figure 1.— Distribution of Ideology Scores

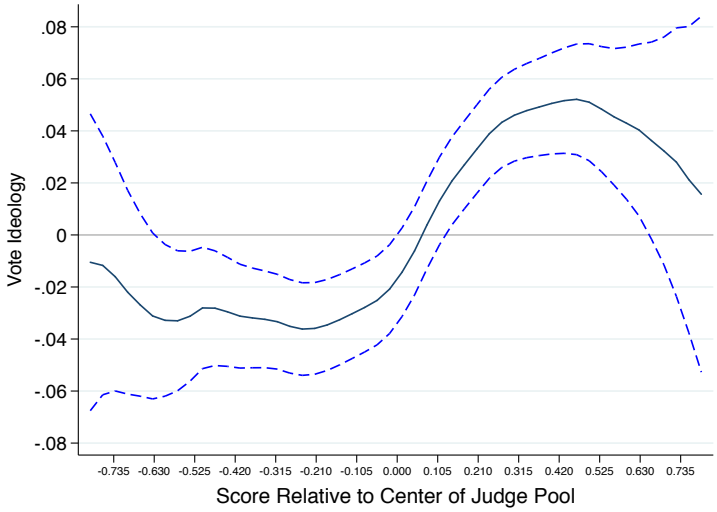


Notes: Ideology scores from the Judicial Common Space database (Epstein et al. 2007). Left panel: raw ideology score. Central panel: ideology score demeaned by the center of the judge pool in a circuit-year. Right panel: ideology score demeaned by the mean ideology score of members of the Supreme Court (The supreme-court ideology comes from Martin and Quinn (2002)).

A.1.3 Which judges have the most ideological signing pattern?

In this section we investigate the relationship between a judge's (relative) ideology score and the ideological flavor of the verdicts she votes for. We call the latter variable Vote Ideology (as defined in Section 2.1).[38]

Appendix Figure 2.— Vote Ideology and Ideology Score of Judge Relative to Center of Judge Pool – local polynomial



Notes: x-axis: Ideology score of a judge demeaned by the center of the *pool* of judges in a Circuit-year. y-axis: Vote ideology, demeaned to be centered at zero. The figure presents a local polynomial regression with an Epanechnikov kernel, where the dependent variable is Vote Ideology, which is coded as 1 for conservative, 0 for mixed or not applicable, and -1 for liberal. The dashed lines depict the 95% confidence interval. Data come from the U.S. Courts of Appeals Database Project (1925-2002 5% Sample).

[38] This is not to be confused with the voting that might takes place when judges deliberate over the verdict, which is unobservable. In the theory section what we call 'voting' is within that deliberation.

We test the relationship between a judge's ideology Score Relative to Center of Judge Pool and her Vote Ideology using a local polynomial regression. The results are presented in Appendix Figure 2. As can be seen, the most ideological voting is obtained for moderately ideological judges and once a judge becomes sufficiently extreme her voting becomes less ideological. To test the statistical significance of this result we run the regression:

$$\begin{aligned} \textbf{Vote Ideology}_{pcit} = \alpha + \gamma_1 \textbf{Score Relative to Center of Judge Pool}_{cit} + \\ \gamma_2 \textbf{Score Relative to Center of Judge Pool}^{\mathbf{2}}{}_{cit} + \\ \gamma_3 \textbf{Score Relative to Center of Judge Pool}^{\mathbf{3}}{}_{cit} + \nu_{pcit} \end{aligned} \tag{7}$$

for judge i on panel p in Circuit c and year t.[39] The regression results are presented in Appendix Table A.2. It confirms (by the negative cubic term) that judges with moderate scores have the most ideological voting pattern while judges with extreme scores have a less ideological voting pattern, implying that being in ideological minority makes a judge behave less ideologically. According to the coefficients in the table, the strongest ideological voting pattern is obtained for judges with scores of -0.36 and +0.47 which are both well within the bounds of our distribution, which goes from around -0.8 to +0.8.

APPENDIX TABLE A.2

VOTE IDEOLOGY AND IDEOLOGY SCORE OF JUDGE RELATIVE TO CENTER OF JUDGE POOL

	(1)
	Vote Ideology
Score Relative to Center of Judge Pool	0.180***
	(0.0308)
Score^2	0.0614
	(0.0659)
Score^3	-0.366***
	(0.125)
N	23031
R-sq	0.002

Notes: Robust standard errors clustered at the circuit-year level in parentheses (* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$). Data on cases comes 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. The dependent variable is Vote Ideology, which is coded as 1 for conservative, 0 for mixed or not applicable, and -1 for liberal. The independent variables are polynomials of the ideology score demeaned by the center of the pool of judges available to be assigned in a Circuit-year.

The result is robust when using polynomials of a higher order and when using the lifetime average for each judge, splitting the sample according to whether the case affirmed the lower court decision, or using the alternative ideology score (the one using the appointing President's party and the share of judges in the

[39] We use polynomial of the third degree in the regression to enable testing for a U-shape on the left and a hill-shape on the right.

pool who are of the other party). These robustness checks are available upon request.

It is worth noting that, like Finding 3, also this result disappears when using raw ideology scores (i.e., not relative to the pool)—see Figure 1. Hence, what drives this result is that a judge is ideologically extreme *relative* to her peers—it is not about extreme ideology per se, but about the interaction between peers who disagree ideologically. It may be noted also that the regression in (7) does not contain fixed effects. Adding circuit and year fixed effects makes the results non significant, possibly due to the much smaller sample used here compared to Findings1-3.[40]

A.2 Robustness of empirical results

A.2.1 Robustness for Finding 1

We start with a brief discussion of related research. It is standard in the literature to assume that the median judge in the panel is pivotal in setting the verdict (Martin et al. 2004 note that the median justice model "figures prominently and crucially in a wide array of research on the Court"). Other competing models, however, do exist, in particular one that claims that the author of the majority opinion is pivotal too (see e.g., Schwartz 1992, Hammond et al. 2005 and, most recently, Hangartner et al. 2019).[41] However, most of the studies analyze a different setting than ours: the Supreme Court. As far as we know, Cross (2007) is the only other study that has examined the median voter theorem in our setting, the Circuit Courts. We here describe and further analyze his results.

In Column 1 of Table A.3, we reproduce the finding from Table 6.3 in Cross (2007).[42] Our data sample is slightly smaller than that of Cross (2007) because we drop panels with judges whose scores are tied (adding ties does not change the results), but we are able to replicate the finding: the sum of the scores of the judges on the panel is correlated in this specification with Verdict Ideology, but the score of the median judge is not. In Column 2, we add the Center of Judge Pool and find that the sum of the scores of the judges is no longer significant. This suggests that the results of Cross (2007)—that the whole panel composition determines the verdict—is in fact driven by a variable he has omitted: the ideology of the circuit. In Column

[40] Recall that in order to construct our measure of Vote Ideology we rely on the Verdict Ideology as coded by the U.S. Courts of Appeals Database Project, which consists of a random sample of only 5% of all cases. Recall also that each of the roughly 10000 observations reported in Table III represents on average about 30 cases.

[41] When it comes to the verbal opinion that accompanies the verdict, the median of the *majority coalition* was also shown to be pivotal (Spriggs and Hansford 2002; Westerland 2003; Carrubba et al. 2007; Clark and Lauderdale 2010). Since in this paper we analyze the ideology of the binary verdict, this option is less relevant, but let us note that most verdicts in the Circuit Courts are unanimous, in which case the median of the majority is also the median of the whole panel.

[42] There are also other studies that construct the relative position of judges. For example, Peresie (2005) sorts the judges on a panel, but presents a regression where judge votes are the unit of observation and the scores of the judge's two colleagues (left and right of the pair) are included as regressors. This means that each judge score enters three times in the dataset, once for the score of the judge whose vote is the dependent variable, and twice more as the left score or right score or both.

3, we separate the sum of the scores into the score of the left judge and score of the right judge. Now we see that the score of the median judge is the main driver of Verdict Ideology, though the score of the right-most judge is also correlated with Verdict Ideology. In Column 4, we show what is arguably the best specification by including all of these measures and, importantly, controlling for the circuit's ideology, and find that the median judge is the only judge affecting the verdict. Appendix Figure 3 visualizes the results. Columns 3 and 4 are close to the specification in Ambrus et al. (2015), a recent experimental paper examining group-decision making. Controlling for Center of Judge Pool instead of subtracting it from the judge's ideology score (like we do in Table I) also shows the relevance of the average ideology score of the pool of judges in each Circuit and each year.

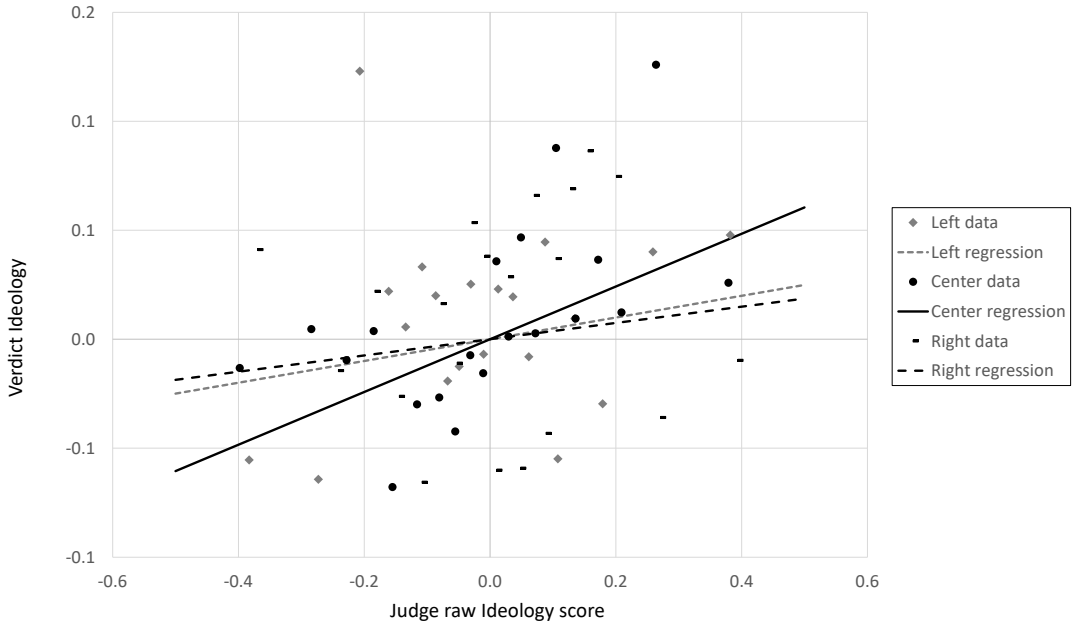
APPENDIX TABLE A.3

VERDICT IDEOLOGY AND IDEOLOGY SCORES OF PANEL MEMBERS – FURTHER INVESTIGATION

	(1)	(2)	(3)	(4)
		Verdict Ideology		
Median of Panel Ideology Score	0.0541	0.0772	0.149***	0.121***
	(0.0607)	(0.0591)	(0.0308)	(0.0354)
Sum of Panel's Ideology Scores	0.0949**	0.0437		
	(0.0347)	(0.0331)		
Center of Judge Pool		0.249*		0.251*
Ideology Score		(0.134)		(0.133)
Left of Panel Ideology Score			0.0869	0.0499
			(0.0717)	(0.0720)
Right of Panel Ideology Score			0.102*	0.0373
			(0.0527)	(0.0448)
N	7677	7677	7677	7677
R-sq	0.007	0.008	0.007	0.008

Notes: Robust standard errors clustered at the circuit-year level in parentheses (* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$). Data 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. The dependent variable is Verdict Ideology, which is coded as 1 for conservative, 0 for mixed or not applicable, and -1 for liberal.

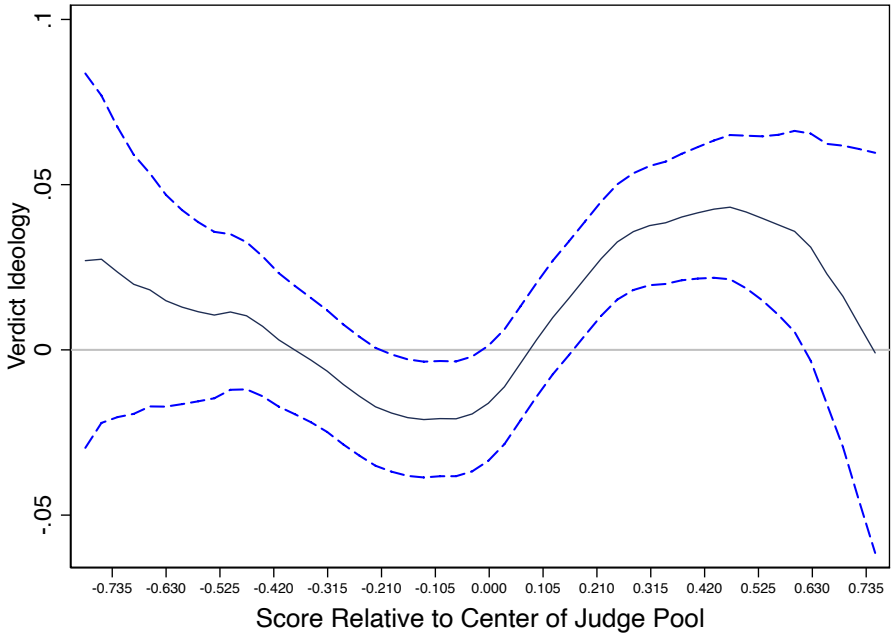
Appendix Figure 3.— Verdict Ideology and Ideology Scores of Panel Members



Notes: x-axis: Ideology score of a judge. y-axis: Verdict 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 verdicts in a bin of judges with similar ideology scores and with the same role within a panel (left, median or right). The lines represent the regression coefficients from column 4 of Appendix Table A.3. As such, each set of dots is plotted after residualizing for the other variables in the regression. The y-axis is demeaned to be centered around zero. Data 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.

Finally, also as a form of robustness, we investigate the relationship between a judge's ideology and the ideological flavor of the verdicts produced by panels she is sitting in. If the median of the panel is the primary driver of the verdict, then extreme judges should rarely influence the outcome of the panel since they are seldom median. Appendix Figure 4 visualizes a local polynomial regression with a judge's ideology as independent variable and the verdict ideology as a dependent variable. As can be seen, the largest proportion of ideological verdicts is obtained when moderates are involved. Put differently, extreme judges are not affecting the verdicts. This result is robust to adding quartic terms and splitting the sample according to whether the lower court decision was affirmed or not and to using the lifetime average for each judge.

Appendix Figure 4.— Verdict Ideology and Ideology Score of Judge Relative to Center of Judge Pool – local polynomial

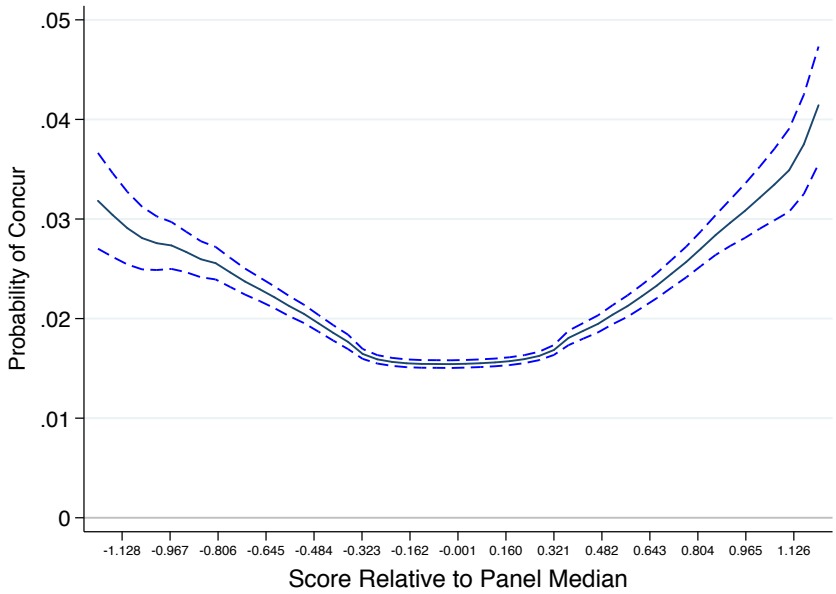


Notes: x-axis: Ideology score of a judge demeaned by the center of the *pool* of judges available to be assigned in a Circuit-year. y-axis: Verdict ideology, demeaned to be centered at zero. The figure presents a local polynomial regression with an Epanechnikov kernel, where the dependent variable is Verdict Ideology, which is coded as 1 for conservative, 0 for mixed or not applicable, and -1 for liberal. The dashed lines depict the 95% confidence interval. Data come from the U.S. Courts of Appeals Database Project (1925-2002 5% Sample).

A.2.2 Robustness for Finding 2

We provide here two robustness checks for Finding 2. First, Appendix Figure 5 (local polynomial) shows that the result that a judge is more likely to dissent when the panel median is ideologically far from her is robust to using concurrence instead of dissent.

Appendix Figure 5.— Concurrence and Ideology Score of Judge Relative to Median of Panel – local polynomial



Notes: x-axis: Ideology score of a judge demeaned by the median of the panel of judges assigned on the case. y-axis: Rate of concur. The figure presents a local polynomial regression with an Epanechnikov kernel, where the dependent variable is rate of concur. The dashed lines depict the 95% confidence interval. Data come from OpenJurist (1950-2007). Ideology scores come from the Judicial Common Space database.

Second, Table A.4 shows that the result—both for dissents and for concurrences—is robust to using a Logit specification.

APPENDIX TABLE A.4

Dissent and Ideological Distance to Median of Panel

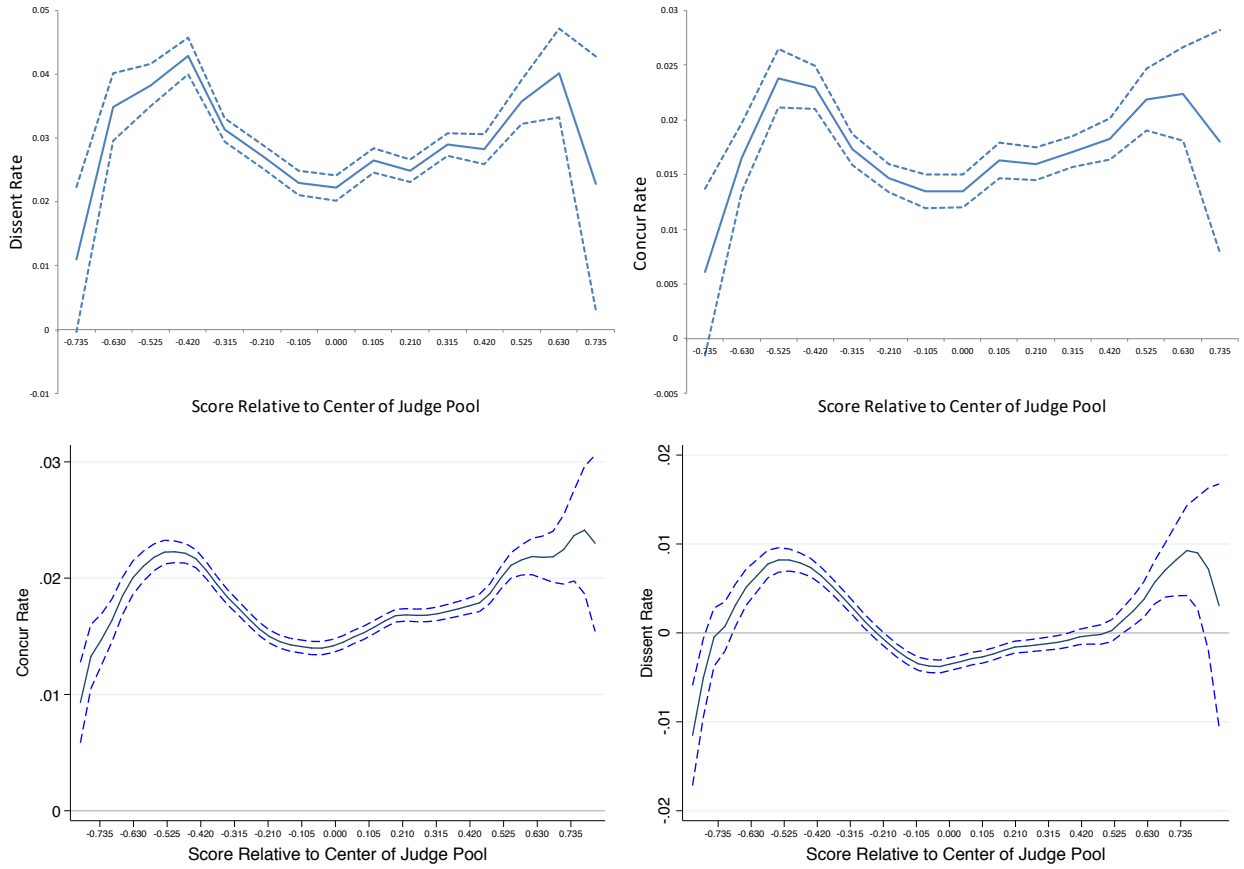
	(1) Dissent	(2) Concur
Distance to Median of Panel	1.192***	0.666***
	(0.119)	(0.142)
Distance^2	-0.252*	0.0437
	(0.140)	(0.171)
Judge Fixed Effects	Y	Y
Circuit Fixed Effects	Y	Y
Year Fixed Effects	Y	Y
N	540385	513597

Notes: Robust standard errors clustered at the judge level in parentheses (* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$). Data on cases comes from OpenJurist (1950-2007). Ideology scores come from the Judicial Common Space database. Ideology scores are demeaned by the actual center of the panel of judges assigned on a case. The dependent variable is a dummy for whether a judge dissented (column 1) or concurred (column 2) in the panel. Fixed effects include year, circuit, and judge.

A.2.3 Robustness for Finding 3

The upper left panel of Appendix Figure 6 shows the raw data of Dissent Rate by Score Relative to Judge Pool when grouping observations with similar relative ideology into separate bins. These scores were divided from left to right into 15 evenly-spaced bins, where for each bin we estimated the average dissent rate in that bin. We also present the 95% confidence interval around the average dissent rate.[43] As can be seen the spider pattern appears clearly here too.

Appendix Figure 6.— Dissent or concur and Ideology Score Relative to Center of Judge Pool



Notes: Data on cases come from OpenJurist (1950-2007). Ideology scores come from the Judicial Common Space database. x-axis: Ideology score of a judge demeaned by the center of the *pool* of judges available to be assigned in a Circuit-year. Upper left panel: Dissent rate (y-axis) when the x-axis is divided into 15 evenly-spaced bins (the number denotes the midpoint of the bin); the y-axis shows the mean concurrence rate for each bin; the means are weighted averages over all judges and all Circuit-years, accounting for the number of times each judge actually appeared on cases in any given Circuit-year.. Upper right panel: Concurrence rate (y-axis) when the x-axis is divided into 15 evenly-spaced bins. Lower left panel: Concurrence rate (y-axis) using a local polynomial regression with an Epanechnikov kernel. Lower right panel: Dissent rate (y-axis) using a local polynomial regression with an Epanechnikov kernel after residualizing by Circuit and Year fixed effects. The dashed lines in all panels depict the 95% confidence interval.

[43] The 95% confidence interval comes from a weighted regression of the dissent rate on a constant for each bin, with weights being the number of votes cast by a judge in a Circuit-year.

The upper right panel (bins) and the lower left panel (local polynomial) of Appendix Figure 6 show the spider pattern appears also when using concurrences instead of dissents, though less markedly on the right. The lower right panel of Appendix Figure 6 (local polynomial) shows the spider pattern appears also when residualizing by circuit and year fixed effects.

We further check if the spider is present under different weighting and scores. Appendix Table A.5 Column 1 repeats the main specification using (the absolute) Distance to Center of Judge Pool. Column 2 includes a cubic term (the quadratic and cubic terms jointly produce a spider). Column 3 weights each judge equally but excludes judges who vote less than 10 times. Column 4 does the same but presents a logit model. Column 5 uses a 2-year binned dissent and concur rates. Column 6 uses the lifetime average rate. Column 7 uses the Distance to the Supreme Court. Column 8 uses both Distance to Center of Judge Pool and Distance to the Supreme Court. As can be seen, the spider pattern is robust in these specifications. Column 9 shows that the spider is robust to including polynomials of the distance to panel median as controls, indicating that the spider pattern is not driven by interactions within particular panels. Finally, Column 10 randomly assigns Distance to Center of Judge Pool to a different judge to mitigate the concern of spurious significance or erroneous clustering level. As can be seen the result then disappears implying the result is not driven by spurious significance or by the chosen level of clustering.

We also check if the spider is robust to dropping one Circuit at a time. Appendix Table A.6 Column 1 repeats the main specification using (the absolute) Distance to Center of Judge Pool. Columns 2 to 13 drop one Circuit at a time. As can be seen the spider pattern is robust. This mitigates the concern that the pattern is driven by outliers.

Appendix Table A.5.— Robustness to Alternative Scores

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
					Dissent or Concur					
Distance to Center of Judge Pool	0.0664***	-0.0140	0.0460***	4.074***				0.0636***	0.0538***	
Scores based on Ideology	(0.0103)	(0.0196)	(0.0121)	(0.649)				(0.0105)	(0.0109)	
Distance^2	-0.0649***	0.243***	-0.0433**	-5.756***				-0.0349**	-0.0891***	
	(0.0156)	(0.0696)	(0.0188)	(1.020)				(0.0157)	(0.0165)	
Distance^3		-0.314***								
		(0.0697)								
Distance to Center of Judge Pool					0.0678***					
2-year bin					(0.0132)					
Distance^2					-0.0670***					
					(0.0197)					
Distance to Center of Judge Pool						0.103***				
Lifetime average						(0.0318)				
Distance^2						-0.114**				
						(0.0516)				
Distance to Supreme Court							0.0622***	0.0416***		
							(0.00956)	(0.00960)		
Distance^2							-0.0897***	-0.0975***		
							(0.0143)	(0.0150)		
Distance to Median of Panel									0.0259***	
									(0.00506)	
Distance^2									0.0215***	
									(0.00670)	
Distance to Center of Judge Pool										0.00325
Resampled										(0.00897)
Distance^2										-0.00482
										(0.0141)
Circuit Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
N	10043	10043	7744	7744	5836	424	10592	10043	509022	10387
R-sq	0.124	0.126	0.111		0.147	0.267	0.114	0.131	0.008	0.111

Notes: Data on cases comes from OpenJurist (1950-2007). Ideology scores come from the Judicial Common Space database. Absolute value of the distance to the center of the judge pool is the main independent variable. The dependent variable is the judge's sum of dissent rate and concurrence rate in this Circuit-year (with the exception of column 5, which is the rate calculated over two years; column 6, which is the lifetime rate; and column 9, which is the decision to dissent or concur on this panel). Fixed effects include circuit and year (year of appointment for column 6). Observations are weighted by the number of votes cast by the judge in the time-unit of observation (with the exception of columns 3 and 4, which do not weight but exclude judges with less than 10 votes in a Circuit-year). Column 4 runs a logit model and all other columns run linear probability models. Distance to Panel Median is not an alternative score, but is presented as a rejection of judge-specific mechanisms. Resampled Distance to Center of Judge Pool is presented as a rejection of spurious significance, where judicial scores have been randomly reassigned. All columns use robust standard errors clustered at the Circuit-year level, except column 5, which clusters at the Circuit-2-year-bin level, and column 6, which clusters at the Circuit level (* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$).

Appendix Table A.6.— Robustness to Dropping One Circuit at a Time

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
							Dissent or Concur						
Distance to Center of Judge Pool	0.0664***	0.0690***	0.0792***	0.0758***	0.0636***	0.0738***	0.0678***	0.0408***	0.0615***	0.0521***	0.0705***	0.0645***	0.0763***
Scores based on Ideology	(0.0103)	(0.0107)	(0.0106)	(0.0107)	(0.0106)	(0.0119)	(0.0105)	(0.0105)	(0.0106)	(0.0111)	(0.0106)	(0.0105)	(0.0104)
Distance^2	-0.0649***	-0.0691***	-0.0836***	-0.0733***	-0.0645***	-0.0730***	-0.0716***	-0.0336**	-0.0596***	-0.0441***	-0.0704***	-0.0621***	-0.0705***
	(0.0156)	(0.0161)	(0.0159)	(0.0162)	(0.0159)	(0.0176)	(0.0157)	(0.0170)	(0.0161)	(0.0168)	(0.0161)	(0.0159)	(0.0163)
Drop Circuit	None	1	2	3	4	5	6	7	8	9	10	11	12
Circuit Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
N	10043	9567	9132	9186	9393	8881	9165	9287	9345	8497	9359	9386	9275
R-sq	0.124	0.113	0.137	0.124	0.123	0.119	0.129	0.133	0.129	0.131	0.119	0.127	0.114

Notes: Data on cases comes from OpenJurist (1950-2007). Ideology scores come from the Judicial Common Space database. Absolute value of the distance to the center of the judge pool is the main independent variable. The dependent variable is the judge's sum of dissent rate and concurrence rate in this Circuit-year. Fixed effects include circuit and year. Observations are weighted by the number of votes cast by the judge in the time-unit of observation. Each one of the 12 columns drops one circuit from the sample. All columns use robust standard errors clustered at the Circuit-year level (* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$).

Appendix Table A.7 (columns 1 and 2) controls for biographical characteristics of the judge. These are controlled using dummy indicators for party of appointment, whether the judge and appointing President were of the same or different political parties, whether government (Congress and President) was unified or divided at the time of appointment, whether the judge was Protestant, Evangelical Protestant, Mainline Protestant, Catholic, Jewish, or non-religious, whether the judge was Black, non-white, or female, whether the judge received a law degree from a public institution, a bachelor's degree from a public institution, a bachelor's degree from within the state of appointment, or obtained further graduate studies in law (LLM or SJD), was born in the 1910s, 1920s, 1930s, 1940s, or 1950s, had previous experience as federal district judge, law professor, U.S. attorney, assistant U.S. attorney, Solicitor-General, mayor, state governor, Attorney-General, Deputy or assistant district/county/city attorney, Bankruptcy judge, U.S. Magistrate, Congressional counsel, District/County/City Attorney, Local/municipal court judge, Sub-cabinet secretary, Cabinet secretary, Special prosecutor, State lower court judge, State high court judge, or Local/municipal court judge, or had experience in City council, Department of Justice, Solicitor-General's office, or served as a member of the State house, State senate, U.S. House of Representatives, or had previous experience in private practice, in government, or in other federal capacity, or received an exceptional rating from the American Bar Association, and, in case the judge was elevated from the district courts, the party of the President who made the district bench appointment. As can be seen, the spider pattern is robust and is not driven by some special characteristics of extreme judges.

APPENDIX TABLE A.7

Dissent and Ideology Score of Judge Relative to Center of Judge Pool

	(1)	(2)	(3)	(4)	(5)	(6)
	Dissent	Concur	Dissents or Concurs Rate		Dissent or Concur Rate	
Distance to Center of Judge Pool	0.0459***	0.0288***	0.0460***	0.0460***	0.0760***	0.0522***
	(0.00784)	(0.00586)	(0.0107)	[0.0184, 0.0708]	(0.0109)	(0.0136)
Distance^2	-0.0403***	-0.0324***	-0.0433**	-0.0433**	-0.0750***	-0.0495**
	(0.0125)	(0.00886)	(0.017)	[-0.0778, 0.0020]	(0.0167)	(0.0215)
Judge Characteristics	Y	Y	N	N	N	N
Circuit Fixed Effects	Y	Y	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y	Y	Y
Sample	All	All	All	All	Affirmed	Not Affirmed
N	8692	8692	7744	7744	9577	9622
R-sq	0.183	0.173	0.111	0.111	0.091	0.072

Notes: Robust standard errors clustered at the circuit-year level in parentheses (* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$). Data on cases comes from OpenJurist (1950-2007). Ideology scores come from the Judicial Common Space database. Judicial characteristics come from Federal Judiciary Center/Attributes of U.S. Federal Judges Database and controlled for as binary indicators. The main independent variable is (the absolute value of) ideology score demeaned by the center of the pool of judges available to be assigned in a Circuit-year. Column 1: dependent variable is dissent rate, controlling for judges' biographical characteristics. Column 2: dependent variable is concurrence rate, controlling for judges' biographical characteristics. Colum 3: dependent variable is dissent+concurrence rate, bootstrapped standard errors clustered at the Circuit-year level. Column 4: dependent variable is dissent+concurrence rate, wild bootstrap. Column 5: dependent variable is dissent+concurrence rate, subsample of decisions that affirmed the lower court verdict. Column 6: dependent variable is dissent+concurrence rate, subsample of decisions that did not affirm the lower court verdict.

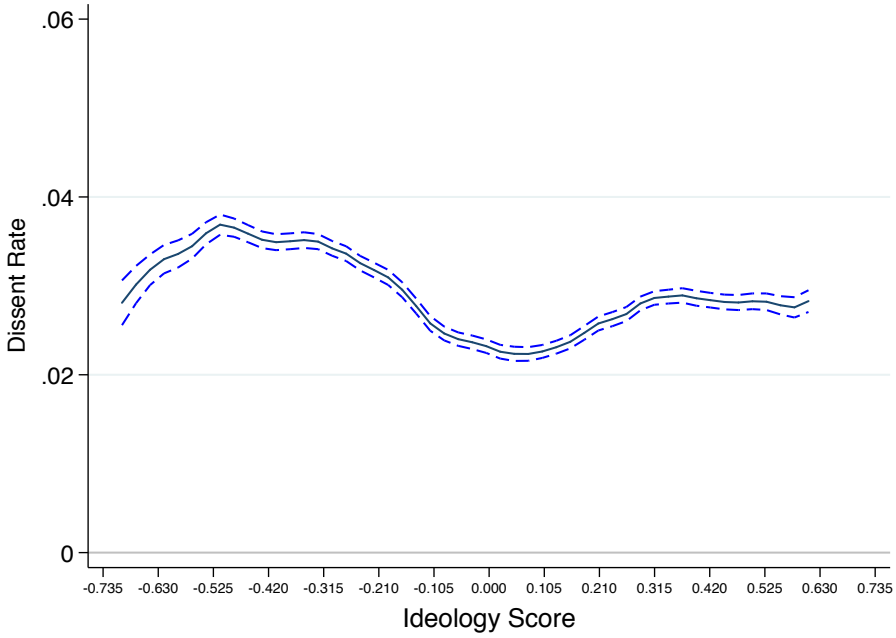
Appendix Table A.7 (columns 3 and 4) presents the main specification but using two forms of bootstrapping. As can be seen, the spider result is robust and the p-values similar as in the main specification.

Appendix Table A.7 (columns 5 and 6) presents the main specification but splits the sample according to whether the decision affirmed the lower court verdict. The sample size differs slightly when there are no affirmances or all affirmances for a judges in a Circuit-year. As can be seen, the spider result is robust.

Finally, Appendix Figure 7 shows the raw data (local polynomial) using the non-demeaned ideology score as the independent variable.[44] As can be seen, the spider pattern is then strongly attenuated (compared to with the equivalent Figure 3 using relative scores), which indicates that the spider result is driven by the interaction of peers who disagree ideologically rather than by extreme ideology per se.

[44] Figure 7 has a slightly more narrow range on the x-axis compared to the equivalent figures using relative ideology scores. This is since the distribution of raw scores has a more narrow support than for the relative score (see Appendix Figure 1).

Appendix Figure 7.— Dissent and (non-relative) Ideology Score of Judge – local polynomial



Notes: x-axis: Non-demeaned Ideology score of a judge. y-axis: Rate of dissent. The figure presents a local polynomial regression with an Epanechnikov kernel, where the dependent variable is rate of dissent. The dashed lines depict the 95% confidence interval. Data come from OpenJurist (1950-2007). Ideology scores come from the Judicial Common Space database.

B Further theoretical predictions

B.1 Predicting the S-shaped signing pattern

Here we show that our main model produces also the empirical observation presented in Section A.1.3—an S-shaped pattern of the ideological flavor of the verdicts she votes for as a function of judges' ideology, reflecting that the moderately ideological judges have the most ideological pattern. To explain this observation analytically we construct a measure of the ideology $J(t;v)$ of a signature of judge t as follows:

$$J(t;v(y)) \equiv \begin{cases} v(y) \;\; if \;\; s\,(v;t) = 1 \\ v_t(y) \;\; if \;\; s\,(v;t) = 0 \end{cases} .$$

In equilibrium, $v(y) = v_{t_m}(y)$, hence, by signing the verdict, a judge in practice aligns with the median's preference. Meanwhile, the verdict voted for by a dissenting judge is assumed to completely align with her own ideology $v_t(y)$ because, in this case, she is free to choose as she likes. This way, $E_y[J(t;v(y))]$ captures the "ideological bias" of judge t's signing pattern.

Proposition 4 *Consider the spider equilibrium described for $\alpha < 1$. Then there exists a range of values of W for which $E\,(J\,(t))$ is maximized for an intermediate value of $|t|$ (i.e., $argmax_t E\,(J\,(t)) \notin \{-1, 0, 1\}$).*

Proof: See Appendix C.5.

Q.E.D.

The intuition is straightforward. As explained above, centrists and moderately ideological judges will dissent virtually whenever they are not the median, hence will have $J = v_t\,(y)$ in almost all cases. This implies that among them we will observe an increase in ideological bias as $|t|$ increases. Conversely, extreme judges almost never dissent, hence their voting pattern mostly reflects the verdicts of the panels they sit in, which are determined by the median of these panels, hence their voting pattern tends to be less ideological than the moderates' voting pattern. Hence, a concave ideological cost can rationalize the empirical finding reported in Appendix A.1.3.

B.2 Predicting the effect of collegial pressure

Here we derive two additional predictions that we will use in Section D.1.15 for testing our model against an alternative model (detailed in Section D.2). The additional predictions relate to the effect of collegial pressure on the pattern of dissent. Let $\widetilde{t}$ denote the peak of the hill shape in the range $[0, 1]$ (i.e. the point where the probability of dissent as a function of t drops). The next proposition characterizes how $\widetilde{t}$ changes as a function of the collegial pressure W.

PROPOSITION 5 *Consider the spider equilibrium described for $\alpha < 1$. Then: (i) if $W = 0$, $P^*(t)$ monotonically increases in $|t|$; and (ii) $\widetilde{t}$ decreases in W.*

PROOF: See Appendix C.6.

Q.E.D.

Part (i) says that, should collegial pressure become very small, the spider pattern will disappear and the dissent rate will be a purely increasing function of a judge's extremeness. The intuition for this result is that, without collegial pressure, it becomes possible also for extreme judges to be ideologically picky and hence dissent whenever they disagree with the verdict. Then, given that they are rarely the median in their panels, they will dissent more often than everyone else. Part (ii) of the proposition expresses that, as collegial pressure increases, the range of judges who dissent whenever they are not the median shrinks. This is of course natural since strictly adhering to one's morals becomes more costly under high peer pressure, hence less judges will do so. In Appendix D.1.15 these predictions are corroborated empirically.

C Proofs

C.1 Proof of Proposition 1

For any set of strategies of the other judges, judge t is always strictly better off when $v(y) = 1$ if $y < t$ and $v(y) = 0$ if $y \geq t$, because she can then sign a verdict she agrees with and bear no cost (whereas a verdict she disagrees with always inflicts a cost, either through $D(\cdot)$ or through $P(\cdot)$). Therefore, it is a weakly dominant strategy for judge t to always vote according to her ideological preference. This immediately implies, by the median voting theorem, that there exist an equilibrium in (weakly) dominant strategies in which $v(y) = 1$ whenever $y < t_m$ and $v(y) = 0$ whenever $y > t_m$, in which case $v(y) = v_m(y)$ in all panels.[45] Finally, the sequential voting ensures that this holds in any equilibrium of the game. This is so because if the voting in a given case would have resulted in $v(y) \neq v_m(y)$, then backward induction implies that the first judge who had voted for this verdict yet preferred $v(y) = v_m(y)$ over this verdict could have changed the verdict (and strictly increase her payoff) by deviating and voting according to her preference.

C.2 Proof of Proposition 2

Proposition 1 establishes that the median's preference determines the verdict. Consider now a judge who's signing strategy includes dissenting on a case with characteristic $y \in [-1, 1]$ iff she disagrees with the verdict. Compare the probability of dissent of judge t when the median is t_{m_1} and when the median is t_{m_2}, where $t_{m_2} < t_{m_1} < t$. The verdicts of t_{m_1} and t_{m_2} differ only when $t_{m_2} < y < t_{m_1}$, in which case the

[45] Note that the signing strategy of a judge has no effect on the payoff of other judges hence the equilibrium whose existence is guaranteed consists of the voting strategies just described and the signing strategies that maximize the payoff of each judge given these voting strategies.

decision of t_{m_1} coincides with the preferences of t while the decision of t_{m_2} contradicts these preferences. Hence follows that the judge is weakly more likely to dissent under t_{m_2}. Since this holds for any y follows that the judge is weakly more likely to dissent under t_{m_2} for any distribution of y that the judge may have as her set of case characteristics for which she dissents upon disagreeing with the verdict.[46] Given that the distribution of cases is independent of the identity of the panel, it follows that the probability that judge t dissents is weakly higher when t_{m_2} is the panel's median compared to when t_{m_1} is the median. The same holds if $t < t_{m_1} < t_{m_2}$.

C.3 Proof of Lemma 1

Given t, the event of an unfavorable verdict, $v(y) \neq v_t(y)$, requires by Proposition 1 that either $t_m < y < t$, or $t < y < t_m$. Since the probability of having a case $y < t$ is $F(t)$, and this is also the probability of randomly drawing a judge with ideology to the left of t, we get that the probability that the other two judges in the panel and the case characteristic y are all to the left of t is $[F(t)]^3$, and the probability that of these three independent draws (the draw of y and of the other two judges' ideologies) it is y that is the closest to t (as necessary for generating $t_m < y < t$) is $\frac{1}{3}[F(t)]^3$. Similarly, the probability of $t < y < t_m$ is given by $\frac{1}{3}[1 - F(t)]^3$. Suppose without loss of generality that $t > 0$, so that $F(t) = 1/2 + z$ for some $z > 0$. Then $Prob(v(y) \neq v_t(y)) = \frac{1}{3}(1/2 + z)^3 + \frac{1}{3}(1/2 - z)^3$, whose derivative with respect to z, $(1/2 + z)^2 - (1/2 - z)^2 = 2z$ is positive. This implies that, for $t > 0$, $Prob(v(y) \neq v_t(y))$ increases in $F(t)$ hence also increases in t. Due to symmetry, an equivalent result holds for $t < 0$, implying that $Prob(v(y) \neq v_t(y))$ increases in $|t|$.

C.4 Proof of Proposition 3

Define $\tau \equiv Prob(v(y) \neq v_t(y)) = \int_{-1}^{1} I(v(y) \neq v_t(y)) f(y)\, dy$ for a judge t. From (5) and Proposition 1 it follows that $x = \tau - P$, and the objective function (6) simplifies to $min_P L(P;\tau) = min_P \{(\tau - P)^\alpha + WP\}$.

Consider now the case of $\alpha < 1$. By differentiating L twice w.r.t. P, we get that the second-order condition does not hold ($L'' < 0$) hence all judges have a corner solution: either $P = \tau$ or $P = 0$. Define $\nabla L \equiv L(P = 0) - L(P = \tau) = (\tau)^\alpha - W\tau$. If W is sufficiently small, ∇L will be positive for any feasible value of τ, in which case we get that the probability of dissent is $P = \tau = Prob(v(y) \neq v_t(y))$, which by Lemma 1 is increasing in $|t|$, and this does not produce the spider pattern. However, if W is sufficiently large, we get that ∇L is positive for small τ (α<1 implies that for any W there exists a small enough $\bar{\tau}$ s.t. ∇L> 0 for any $\tau < \bar{\tau}$) and negative for sufficiently large τ. Hence, for small τ, we have $P^*(\tau) = \tau$, so that in this range $P^*(\tau)$ is increasing in τ, while for large τ, $P^*(\tau) = 0$. Thus, if W is sufficiently large, $P^*(\tau)$ is first increasing and then drops to zero and by Lemma 1 the same pattern holds as a function of t. The

[46] Recall that we restrict the equilibria to those where dissent is based solely on case characteristics and disagreement.

value of t at which $P^*(t)$ drops to zero is denoted by $\widetilde{t}$ (see Section B.2). From Proposition 5 we know that $\widetilde{t}$ decreases in W, hence $\widetilde{t} < 1$ for sufficiently large W (note that $lim_{W\to\infty}\bar{\tau} = 0$). However, we know from the proof of Lemma 1 that τ is bounded from below by $\tau(t=0) = \frac{1}{3}[F(t)]^3 + \frac{1}{3}[1-F(t)]^3 = \frac{2}{3}(1/2)^3 = 1/12$, hence, for $\widetilde{t}$ to be in the range $(0,1)$, W cannot be too large. This proves the proposition's statement on the existence of a spider pattern for $\alpha < 1$ and an intermediate range of W.

If $\alpha = 1$, then L increases in P iff $W > 1$ independently of τ. Hence either all judges choose $P(\tau) = 0$ or all judges choose $P(\tau) = \tau$, which precludes a spider pattern. Finally, consider $\alpha > 1$. Then $L'' > 0$ implying that for any τ there is an inner solution given by the first-order condition $L' = 0 \iff \alpha(\tau - P)^\alpha = W \iff P^* = \tau - (\frac{W}{\alpha})^{1/\alpha}$. P^* is thus increasing in τ which, since τ is increasing in $|t|$, precludes a spider pattern.

C.5 Proof of Proposition 4

LEMMA 2 *The probability that judge t is the median is denoted by $P_m(t)$ and given by*

$$P_m(t) = 2F(t)[1-F(t)]. \tag{8}$$

PROOF: Judge t is the median when one other judge is to her left, which happens with probability $F(t)$, and the other is to her right, which happens with probability $1 - F(t)$. As judge's types are i.i.d. and the order of the assignment of judges is irrelevant, we get (8).

Q.E.D.

C.5.1 Proof of the proposition

We prove the proposition for $t \geq 0$. Equivalent statements can be made for $t \leq 0$. When the spider pattern holds, we know from the proof of Proposition 3 that judges with $t \leq \widetilde{t}$ sign always according to their preferences (dissent whenever they disagree with the verdict), hence for them the average signature is simply $Prob(y < t) = F(t)$. Consider now a judge t with $t > \widetilde{t}$. This judge always signs the verdict. With probability $P_m(t)$ this judge is herself the median, and then she signs a $v = 1$ with probability $Prob(y < t) = F(t)$. Otherwise she signs $v(y) = 1$ if the median has $v_{t_m}(y) = 1$, i.e. if $y < t_m$. Noting that the probability that the ideology of the median equals some value $t_m < t$ is computed by differentiating the probability that this ideology is below t_m, i.e. it is given by $\left([F(t_m)]^2\right)'$, and similarly the probability that the ideology of the median equals some value $t_m > t$ is given by $\left([1-F(v)]^2\right)'$, we have

$$\begin{aligned} E[J(t)] &= P_m(t)F(t) + \left[\int_{-1}^{t}\left([F(v)]^2\right)'F(v)\,dv + \int_{t}^{1}\left([1-F(v)]^2\right)'F(v)\,dv\right] \\ &= 2[F(t)]^2[1-F(t)] + \left[\int_{-1}^{t}\left([F(v)]^2\right)'F(v)\,dv + \int_{t}^{1}\left([1-F(v)]^2\right)'F(v)\,dv\right]. \end{aligned}$$

Differentiating $E[J(t)]$ by t we get

$$\begin{aligned}\frac{dE[J(t)]}{dt} &= 4[1-F(t)]F(t)f(t) - 2[F(t)]^2 f(tt) + \left([F(t)]^2\right)' F(t) - \left([1-F(t)]^2\right)' F(t) \\ &= 2F(t)f(t)(2[1-F(t)] - F(t)) + [2F(t)f(t) + 2[1-F(t)]f(t)]F(t) \\ &= 2F(t)f(t)(2[1-F(t)] - F(t) + 1) = 6F(t)f(t)[1-F(t)] > 0.\end{aligned}$$

Thus, $E[J(t)]$ gets its maximum value in the rage $t > \widetilde{t}$ when $t = 1$, where it equals

$$\begin{aligned}E[J(t=1)] &= \int_{-1}^{1} \left([F(v)]^2\right)' F(v)\,dv \\ &= \left[[F(v)]^3\right]_{-1}^{1} dv - \int_{-1}^{1} [F(v)]^2 f(v)\,dv = \frac{2}{3}.\end{aligned}$$

It thus follows that $argmax_t \mathrm{E}[J(t)] \notin \{0,1\}$ if $F\left(\widetilde{t}\right) > \frac{2}{3}$, or, put differently, if $\widetilde{t} > F^{-1}\left(\frac{2}{3}\right)$. Proposition 5 guarantees that this holds for an intermediate range of values of W (the range for which $\widetilde{t} \in \left[F^{-1}\left(\frac{2}{3}\right), 1\right]$).

C.6 Proof of Proposition 5

We prove the proposition for $t \geq 0$. Equivalent statements can be made for $t \leq 0$.

Proof of part (i): When $W = 0$ the objective function for all types becomes $min\, D$ which clearly is achieved by dissenting whenever not being median. By Lemma 1 this probability increases in t.

Proof of part (ii): From the proof of Proposition 3 we know that $\widetilde{t}$ solves $\nabla L = (\tau)^{\alpha} - W\tau = 0$. Solving for τ yields $\tau|_{t=\widetilde{t}} = W^{\frac{1}{\alpha-1}}$. Since, by Lemma 1, $\tau = Prob(v(y) \neq v_t(y))$ is monotonic in t, and given that $\alpha < 1 \Rightarrow \frac{1}{\alpha-1} < 0$, we get that $\widetilde{t}$ decreases in W.

D Appendix: Alternative Explanations

D.1 Refuting of alternative theories

This subsection discusses in brief a number of alternative theories and why they cannot be the explanation for the patterns we find empirically. Among these alternative explanations only one can explain our empirical patterns. This is the instrumental model mentioned in the introduction. We treat this model (judges use dissent as an instrument in trying to reverse the verdict using the Supreme Court) in great detail at the end of the list and in the next section by developing a proper theoretical model (Section D.2), deriving predictions by which it differs from our main model (of peer pressure with a concave cost of bliss-point deviations) and testing these predictions against each other (Section D.1.15). We find that the instrumental model is refuted by the data.

The other stories can be divided into four categories (though many stories can fit under more than one category). The first group (see D.1.1-D.1.2) consists of mechanisms that seem reasonable but ultimately

fail to explain the main Finding 3, part of this finding or some of the other findings. The second category (D.1.3-D.1.6) consists of explanations that are inconsistent with some other aspects of the data. The third group (D.1.7-D.1.9) consists of theories that can explain the empirical patterns only if judges have a concave cost of bliss point deviations, hence are basically variants of our main model. The final category (D.1.10-D.1.14) consists of potential threats to our empirical variables or results. Many of these are rebutted by the robustness checks in the empirical section (Section 3) but here we discuss further grounds for why some empirical concerns are not likely to be the reason for our empirical patterns. While this, of course, is not a final proof that our main model is correct, the long list of alternative explanations is meant to show that the "usual-suspect" mechanisms can be refuted as explanations for our observations (which of course does not mean that these mechanisms are not valid in general).

D.1.1 Signaling for conformity.

Mechanism: A model where dissent is a signal of being an extremist (which is supposedly a bad thing) and where peer pressure is applied to judges based on the type they are perceived to be (in expectation in equilibrium) would supposedly imply extreme judges dissent less. **Refutation**: Bernheim (1994) shows this kind of model will produce dissent (non-conformity in his model) that is *increasing* in the extremeness of the type, hence cannot explain why extreme judges dissent less.

D.1.2 Signaling for extremeness.

Mechanism: Judges aim to please voters or party members or those who selected them. Moderately ideological judges need to signal their ideological belonging while everyone knows already that extremists are extreme hence they do not need to use dissent to signal ideology. **Refutation**: i) This does not explain why centrists dissent less than moderately ideological judges. Alternatively, it assumes that centrists, unlike moderates, do not want to signal they are extreme. ii) As explained under D.1.10 our results are driven by ideology relative to Center of Judge Pool and not by ideology per se and they hold also when using the alternative score based on party of appointing President (where a judge cannot be more "extreme", only more or less in minority).

D.1.3 Bargaining.

Mechanism: Through the bargaining process, extreme judges are more successful than moderately ideological judges in pulling the court's decision in their direction hence have less reasons to dissent. **Refutation**: This description is inconsistent with Appendix Figure 4 which, on the contrary, shows that at large distances from the center of the pool of judges there is a negative correlation between judges' ideology and the ideology of the verdict. It is also very implausible when considering the alternative ideological score (party of appointing President), because it would suggest that a Republican-nominated judge would have more bargaining power the larger is the share of Democrat-nominated judges in the pool.

D.1.4 Log-rolling.

Mechanism: An extreme judge joins an unfavorable decision in return for getting the other judges to agree on the extremist's view in the next case. **Refutation**: This description is inconsistent with Appendix Figure 4 which, on the contrary, shows that at large distances from the center of the pool of judges there is a negative correlation between judges' ideology and the ideology of the verdict. If extreme judges refrain from dissenting in return for getting to determine future cases, then this influence should take away that negative correlation.

D.1.5 Tie the hands of the Supreme Court.

Mechanism: Extreme judges sign the verdict in order to be able to add facts to it and thereby tie the hands of the Supreme Court. **Refutation**: i) If extremists do this, then moderates may want to do this too. ii) The facts added should increase the propensity of verdicts to be ideological, but this is inconsistent with Appendix Figure 4 which, on the contrary, shows that at large distances from the center of the pool of judges there is a negative correlation between a judge's ideology and the ideology of the verdict.

D.1.6 Moderates as official party spokesmen.

Mechanism: Moderate judges are the modal judges in their corresponding (liberal or conservative) group and are very close to their party leadership, while centrists and extremists are located at the tails of their party of affiliation. The moderates thus tend to signal the official party line and act accordingly by being vocal (i.e. dissent more and signal the party position on the issue). **Refutation**: i) Figure 1 shows a clear monotonic relationship between ideology of voting and a judge's ideology score that is not demeaned by the Center of Judge Pool. This is an indication that, abstracting from the censoring effect of the environment, extreme judges adopt a more ideological line than moderate judges. ii) The judge fixed effects exercise in Section 3.3 shows that extreme judges are dissenting more when the pool changes so they become "moderates". So it cannot be that one type of judge is designated to be the main defender of the party line since (by the judge fixed effects exercise) this person changes behavior when the environment changes. iii) Our results are robust to using party of appointing President as ideology score. Here, all judges in the pool who are from the same party get the same score, and "moderate" judges are simply (all) the judges who happen to sit in a well-balanced pool.

D.1.7 Peer pressure increases in extremeness of the dissenting opinion.

Mechanism: Extreme minority opinions (dissents or concurrences) are sanctioned by other judges in the pool more heavily than moderate minority opinions are. Hence, an extremist judge will find it harder to express her true view by dissenting. **Refutation**: Extreme judges can always imitate moderate ideologists, thus dissent at least as much as (instead of less than) the moderates do. To make the extremists prefer to dissent strictly less than the moderates, judges must have a concave personal cost of bliss-point deviations

also in this mechanism. This way, moderate judges, who are under small peer pressure, would choose to dissent and avoid the personal cost, while extreme judges, who are under severe social pressure, would pay the full personal cost and completely refrain from dissenting. It is thus clear that this mechanism is only a variant of our main model in the body of the paper. We prefer our own model because it does not require the additional assumption of increasing social pressure and because we believe the personal cost is likely to apply to the *number* of dissents at least as much as it applies to the size of dissent.

D.1.8 Risk of impeachment.

Mechanism: The risk of impeachment, given dissent, could be increasing in extremeness of the dissenting opinion. **Refutation**: Impeachment is essentially a form of pressure, hence the refutation in mechanism D.1.7 applies.

D.1.9 Extreme dissents get fewer citations

Mechanism: Minority opinions that are very extreme get fewer citations hence are not worth the effort of writing. **Refutation**: We can see two reasons why a judge may want citations. The first is the value of simply being cited many times. Such a motive is devoid of ideology hence is silent on Finding 2. The second motive is ideological, where the purpose of the dissent is to improve the total ideological image of the particular case by moving the attention (citations) from a (subjectively) less favorable majority opinion to a more favorable minority opinion. To see why this cannot explain the observations, note first that while it may be true that extreme opinions are cited less, an extremist can always write less extreme dissents thereby increasing her citations (essentially emulating the moderate). Suppose now that a moderate chooses to write a moderate dissent while the extremist does not (a necessary, but not sufficient, condition for the moderate's dissent rate to be higher than the extremist's dissent rate). Then it must be that the moderate thinks the ideological utility increases substantially (at least sufficiently to be worth the writing effort) when moving attention from, say, a centrist opinion to a moderate one, but that the extremist does not think there is a large utility difference when doing precisely the same. Following the same intuition of our main model, this means that the judge's perceived cost of an unfavorable opinion getting cited has to be concave in the ideological distance of the opinion from the judge's bliss point, which makes it clear that this mechanism is essentially a variant of our main model where the dissent pattern is driven by concave ideological costs.

D.1.10 Unobserved heterogeneity among judges.

Mechanism: Extreme judges have some personal characteristics that are different from the others, and these characteristics make them dissent less. **Refutation**: i) As mentioned in the main text our results are driven by ideology relative to Center of Judge Pool and not by ideology per se (as evident by using judge fixed effects in Section 3.3; the attenuation of Finding 3 and disappearance of the S-shaped voting pattern presented in A.1.3 when using raw ideology score; and the alternative score based on party of appointing

President where a judge's ideology score only depends on her own party and the number of judges of the other party). That is, relative position is not a personal attribute, and when considering extremism per se the empirical patterns are weak to non-existent. ii) Our results are robust to using controls for judge personal characteristics (Table A.7). iii) The judge characteristics driving the result have to be positively correlated with ideology for extreme judges but negatively correlated for less extreme judges (since Finding 3 shows non-monotonicity).

D.1.11 Score bias.

Mechanism: Our ideology score is constructed by the voting behavior of the appointing President and home state senators (see Section 2). This score may be flawed if extreme Presidents appoint non-extreme judges to show they are non–biased. **Refutation**: i) Figure 1 shows an almost linear relationship between ideology of voting and judge's ideology score that is not demeaned by the Center of Judge Pool. This is an indication that the scoring system we use is indeed a good proxy of judges' ideology. ii) The results are robust to the alternative ideology score based on party of appointing President, i.e., a score that is not dependent on how extreme the President is. iii) This mechanism cannot explain non-monotonicity unless one assumes that moderately ideological Presidents for some reason do not at all need to signal that they are not biased.

D.1.12 Results driven by outliers.

Mechanism: The result of low dissent rate for extreme judges could be driven by outliers **Refutation**: i) The pattern is robust to using concurrences rather than dissents (see Appendix Figure 6). ii) There is a lower bound of dissent at zero, hence single outliers cannot pull down the average very much. iii) We test robustness for outlier Circuits: Appendix Table A.6 shows that the results are robust to dropping one Circuit at a time.

D.1.13 Random opinion writing.

Mechanism: The opinion is written by a random judge who gets to decide its content, hence extreme judges are not under-represented among the judges who set the court's opinion. **Refutation**: i) This cannot explain the spider. ii) It is false since we show that the median is pivotal in determining the verdict (Finding 1).

D.1.14 Outlier circuits.

Mechanism: The results are driven by a few circuits with many judges since only large circuits would have large enough variation in judges to include extremists **Refutation**: i) This cannot explain the spider. ii) Appendix Table A.6 shows that the results are robust to dropping one Circuit at a time.

D.1.15 Empirical testing of our model against an alternative model that can explain the spider pattern

This section describes an alternative model that may explain the main stylized finding (Finding 3), derives two predictions from this model that differ from the two predictions outlined in Section B.2, and tests the different predictions against each other empirically.

The alternative theory is one where a judge dissents, at a cost of collegial pressure, in the hope that the U.S. Supreme Court (SCOTUS) will use this as a signal to review the case and overturn the (binary) verdict.[47] Majority voting within panels implies also here that the median judge decides the verdict for the panel. In Appendix D.2 we develop a simple model capturing this mechanism (the model is inspired by the model in Beim et al. 2014). The intuition for why this alternative SCOTUS model can produce a spider-shaped dissent pattern is as follows. A judge compares, case by case, the cost of dissent (W) with how wrongful she thinks a certain verdict is. This means that two prerequisites need to be in place for a judge to dissent: i) she needs to think that the verdict is sufficiently bad to warrant the cost of dissent and ii) she needs to have the Supreme Court on her side as otherwise the verdict will not be overturned anyway. Here, centrists on the one hand usually have the Supreme Court on their side but on the other hand, often being the median, seldom encounter verdicts that are too far from what they think is right. Hence they rarely dissent. Conversely, extremists often dislike the verdict sufficiently to dissent but rarely have the Supreme Court on their side, hence dissent seldom too. Finally, moderately ideological judges may have a larger set of cases where they both sufficiently oppose the verdict and have the Supreme Court on their side. In the appendix we show that this may create a spider pattern of dissent. Two additional predictions (the equivalent of Proposition 5) can be derived from the SCOTUS model.

PROPOSITION 6 *Consider the SCOTUS model. Then: (i) if $W = 0$, $P(t)$ is monotonically decreasing in $|t|$; and (ii) $argmax_{|t|}P(|t|)$ is increasing in W.*

PROOF: See Appendix D.2.

Q.E.D.

Prediction (i) of the SCOTUS model says that, as the collegial cost of dissent (W) approaches zero, the dissent rate becomes a decreasing function of judge's extremeness. This is intuitive since, when the collegial pressure is low, the only factor that determines whether a judge dissents is whether she has the Supreme Court on her side (because there is no collegial pressure against dissenting). This means that centrist judges will dissent very often. Furthermore, the more extreme a judge is, the less likely it is that her preferences will be aligned with the Supreme Court, which means that the dissent rate falls. This way, the prediction of the SCOTUS model is opposite to the prediction of our main model (with a concave ideological

[47] 2% of the Circuit Court decisions are appealed to the Supreme Court, of which 30% are affirmed.

cost) where, as the collegial pressure goes to 0, the dissent rate increases with how extreme the judge is (Proposition 5 part i).

Prediction (ii) of Proposition 6 refers to the consequences of an increase in the cost of dissent. In the SCOTUS model, the judge at the peak of the spider pattern is the one for whom the threshold cutoff for dissent, as determined by the cost of dissent, exactly equals her ideological distance to the Supreme Court, implying that she dissents against any verdict that both she and the Supreme Court view as biased to the "wrong" side. Judges who are more extreme dissent in these cases as well, but in total they are predicted to dissent less because, compared to the judge at the peak, they have less objection to verdicts that manifest extreme ideology on their side of the ideological spectrum, yet do not have the Supreme Court on their side for overturning verdicts they consider to be too moderate. If the cost of dissent increases, judges have to censor themselves more, hence the cutoff for dissent is larger, implying that a judge has to be more ideologically extreme to be at the peak of the spider pattern. Therefore, the prediction of the SCOTUS model is that, as the cost of dissent increases, the spider peak would move outwards. In the main model we had the opposite prediction: an increase in the cost of dissent would have pushed the peak of the spider inwards (Proposition 5 part ii).

Testing predictions for W$\rightarrow$ 0

When testing these opposite predictions (part i of propositions 5 and 6), it is important to note that in the main model it is the **Distance to Center of Judge Pool** that measures the extremeness, while in the SCOTUS model it is the **Distance to SCOTUS** that measures a judge's extremeness.[48] To get a measure of the Distance to SCOTUS, we use the Martin Quinn (MQ) Supreme Court scores (Martin and Quinn 2002). In particular, since Giles et al. (2001), when constructing the ideological scores for Circuit Court judges, put these scores and the MQ scores on the same scale, we calculate the distance to SCOTUS by merging in the annual MQ scores from Giles et al. (2001).[49]

To test the predictions, we use retired status of a judge as a proxy for a very low cost of dissent. The motivation for this is as follows. Firstly, judges who have retired take a reduced caseload, hence have more time to write dissents. Secondly, they arguably have lower collegial pressure from colleagues or are less sensitive to such pressure. We first verify that judges who have retired have a discontinuous drop in caseload from about 100 per year to 30 per year (Figure 8) and that caseload continues to decline gradually thereafter. Next, we show that retired judges dissent more, discontinuously at the year of retirement (Figure 9), and verify that the increase in dissents is not due to age. In fact, older judges are less likely to dissent,

[48] Since Distance to SCOTUS correlates with Distance to Center of Judge Pool, we present also a regression where we control for the latter.

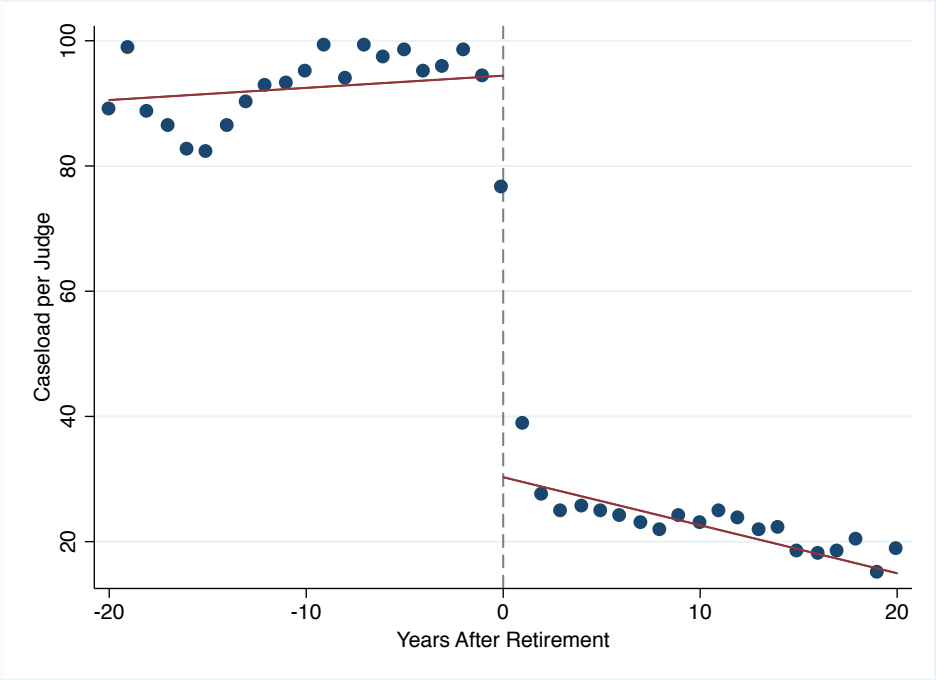
[49] A histogram of **Score Relative to Supreme Court** is presented in the right panel of Appendix Figure 1.

which also explains the decline in dissent before retirement. Figure 9 visualizes the following regression.[50]

$$Dissent\ Rate_{it} = a + b * \mathbf{1}(\textbf{Years after Retirement} \geq \mathbf{0})_{it} +$$
$$c * \textbf{Years after Retirement}_{it} +$$
$$d * \textbf{Years after Appointment}_{it} + \nu_{it}$$

for judge i and year t.

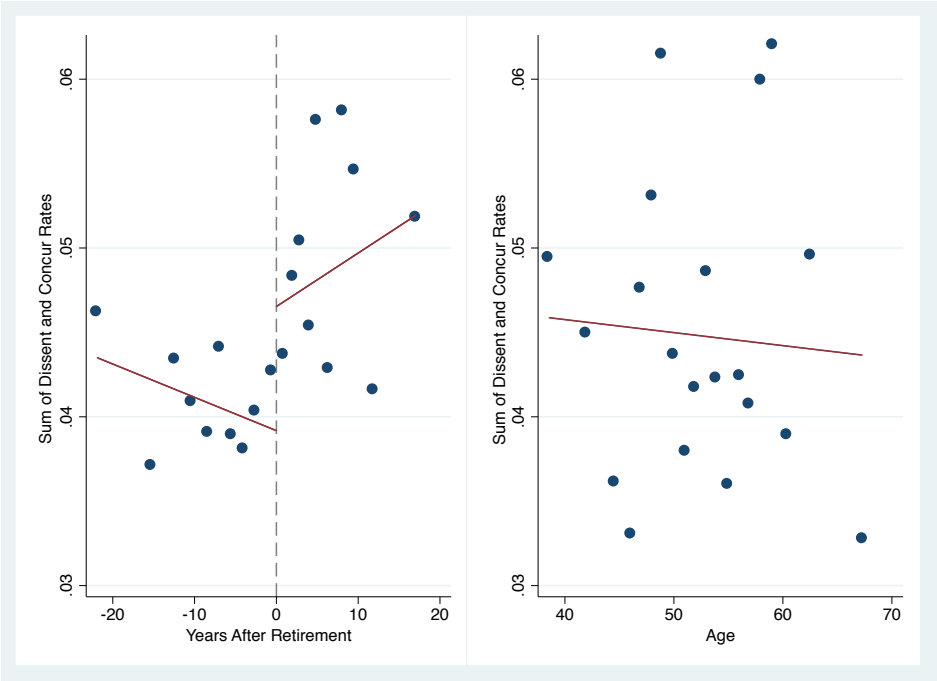
Appendix Figure 8.— Caseload and Years from Retirement



Notes: Each dot represents the average caseload of judges with the same number of years relative to retirement. Data on cases comes from OpenJurist (1950-2007).

[50] We have verified that age and experience vary smoothly around the retirement decision.

Appendix Figure 9.— Dissent or Concurrence and Years from Retirement vs. Age



Notes: Each dot represents the average sum of dissent rate and concurrence rate for judges with the same number of years relative to retirement (left panel) or the same age (right panel) in a Circuit-year. The average is a weighted average to account for the number of times the judge actually appeared on cases in that Circuit-year. Data on cases comes from OpenJurist (1950-2007).

To test the opposing prediction (i) of Propositions 5 and 6 we run the regression in equation (3), limiting the sample to retired judges only. As can be seen in Table A.8, for retired judges, the rate at which judges dissent or concur is positively correlated with **Distance to Center of Judge Pool** (columns 1 and 5), which supports our main model. The rate at which judges dissent or concur is also positively correlated with **Distance to SCOTUS** (columns 3 and 7), which goes against the prediction of the SCOTUS model. Table A.9 (columns 1 and 3) further shows that these two positive correlations hold also when both variables—**Distance to Center of Judge Pool** and **Distance to SCOTUS**—are jointly included as regressors (where the latter is not statistically significant, but is clearly not negative as predicted by the SCOTUS model). Note also that the spider pattern disappears (the even-numbered columns of Tables A.8 and A.9), as predicted in Proposition 5. In total, these results provide support for the main model and go against the prediction of the SCOTUS model.

APPENDIX TABLE A.8

Dissent or Concurrence and Ideology Score among Retired Judges

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Dissent or Concur				
Distance to Center of Judge Pool	0.0323***	0.000569			0.0337***	0.0115		
	(0.00810)	(0.0244)			(0.00814)	(0.0250)		
Distance to Center of Judge Pool^2		0.0524				0.0365		
		(0.0418)				(0.0420)		
Distance to Supreme Court			0.0226***	0.0488*			0.0253***	0.0427
			(0.00849)	(0.0266)			(0.00848)	(0.0264)
Distance to Supreme Court^2				-0.0444				-0.0296
				(0.0466)				(0.0463)
Control for Age and Experience	N	N	N	N	Y	Y	Y	Y
Circuit Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Sample				Senior Judges				
N	3353	3353	3673	3673	3353	3353	3673	3673
R-sq	0.090	0.091	0.081	0.082	0.094	0.094	0.087	0.088

Notes: Robust standard errors clustered at the circuit-year level in parentheses (* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$). Data on cases comes from OpenJurist (1950-2007). Ideology scores come from the Judicial Common Space database. Absolute values of the distance to the center of the judge pool or the Supreme Court are the main independent variables. The dependent variable is the judge's sum of dissent rate and concurrence rate in a Circuit-year. Fixed effects include year and circuit. Observations are weighted by the number of votes cast by the judge in the Circuit-year. Columns 1-4 use the same set of controls as in Table III.

APPENDIX TABLE A.9

Dissent or Concurrence and Ideology Score among Retired Judges: further results

	(1)	(2)	(3)	(4)
		Dissent or Concur		
Distance to Center of Judge Pool	0.0319***	-0.0120	0.0324***	0.000529
	(0.00932)	(0.0250)	(0.00937)	(0.0260)
Distance to Center of Judge Pool^2		0.0749*		0.0541
		(0.0405)		(0.0412)
Distance to Supreme Court	0.000531	0.0393	0.00180	0.0308
	(0.0104)	(0.0285)	(0.0104)	(0.0287)
Distance to Supreme Court^2		-0.0689		-0.0516
		(0.0498)		(0.0502)
Control for Age and Experience	N	N	Y	Y
Circuit Fixed Effects	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y
Sample		Senior Judges		
N	3353	3353	3353	3353
R-sq	0.090	0.092	0.094	0.095

Notes: Robust standard errors clustered at the circuit-year level in parentheses (* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$). Data on cases comes from OpenJurist (1950-2007). Ideology scores come from the Judicial Common Space database. Absolute values of the distance to the center of the judge pool or the Supreme Court are the main independent variables. The dependent variable is the judge's sum of dissent rate and concurrence rate in a Circuit-year. Fixed effects include year and circuit. Observations are weighted by the number of votes cast by the judge in the Circuit-year. Columns 1-2 use the same set of controls as Table III.

Testing predictions about the most dissenting judge

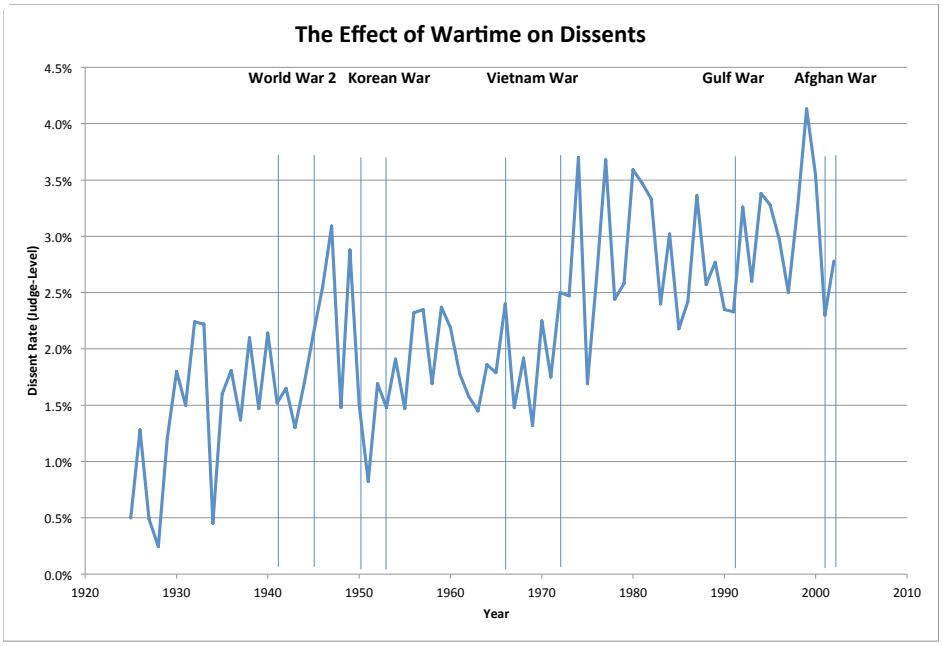
Now we move to testing the effect of an increase in the cost of dissent on who the most dissenting judge is (part ii of propositions 5 and 6). To test this we run two different tests. In the first one, we follow Berdejó and Chen (2014), who find that dissents decrease during wartime (Figure 10). For the purpose of

our test, our assumption is that this decrease is due to wars tending to increase social cohesion (increase W). We test the main (concave ideological cost) model's prediction by examining how the dissent rate is affected by the interaction between a judge's Distance to Center of Judge Pool and wartime.

$$\begin{aligned} Dissent\,Rate_{cit} = a + b * \mathbf{Distance}_{cit} + c * \mathbf{Distance^2}_{cit} + d * \mathbf{Distance}_{cit} * wartime_t + \\ e * \mathbf{Distance^2_{cit}} * wartime_t + f * wartime_t + \nu_{cit} \end{aligned} \tag{9}$$

for judge i in Circuit c and year t.

Appendix Figure 10.— The Effect of Wartime on Dissents



Notes: Each dot represents the proportion of dissents over many votes on cases with the same publication year. Figure reproduced from Berdejó and Chen (2014).

Our second test assumes that judges who are new in a circuit are under increased collegial pressure (higher W). We therefore test the main (concave ideological cost) model's prediction by examining how the dissent rate is affected by the interaction between a judge's Distance to Center of Judge Pool and being new in the circuit's pool of judges.[51] Since this is a variation at the judge level (rather than at the circuit level like wartime), the specification we test is the following.

$$\begin{aligned} Dissent\,Rate_{cit} = a + b * \mathbf{Distance}_{cit} + c * \mathbf{Distance^2}_{cit} + d * \mathbf{Distance}_{cit} * new\_judge_{cit} + \\ e * \mathbf{Distance^2_{cit}} * new\_judge_{cit} + f * new\_judge_{cit} + \nu_{cit} \end{aligned} \tag{10}$$

[51] Being new is defined here as belonging to the pool for a full calendar year for the first time. Transition years (the year in which a judge joins a new circuit) are excluded from the sample for this test.

for judge i in Circuit c and year t.

We test the SCOTUS model's predictions by running the same regressions (9 and 10) but using Distance to SCOTUS as a measure of extremeness. The peak of the spider is determined by the first-order condition of the regression equation. Therefore, to test for a shift in the peak of the dissent rate, we test for a significant difference between $\frac{-b}{2c}$ and $\frac{-(b+d)}{2(c+e)}$. If an increase in W (whether due to wartime or due to the judge being new in the circuit) shifts the peak inwards ($\frac{-b}{2c} > \frac{-(b+d)}{2(c+e)}$), this would corroborate the main model and weaken the SCOTUS model and vice versa. The regressions, the ratios, and the test statistics for the equality of the coefficient ratios are reported in Tables A.10 and A.11 respectively for the two tests. Using **Distance to Center of Judge Pool**, Column 2 of Table A.10 reports that during war there is a statistically significant inward shift of the peak of the spider, which is consistent with the main model. Using **Distance to Supreme Court**, Column 4 of the same table reports an outward shift, as predicted by the SCOTUS model, but it is very far from being statistically significant. Similarly, Column 2 of Table A.11 reports that for new judges there is a statistically significant (at the 10% level) inward shift of the peak of the spider, which is consistent with the main model, and Column 4 reports an outward shift, as predicted by the SCOTUS model, but again it is very far from being statistically significant (the P-value approaches 1).

Judged together, the three tests presented in this section seem to refute the SCOTUS model while supporting the main model presented in this paper.

APPENDIX TABLE A.10

Dissent or Concurrence and Ideology Score among Judges: Tests for changes in who dissents the most due to increase in dissent costs during wartime

	(1)	(2)	(3)	(4)
	Dissent or Concur			
Distance	0.0806***		0.0962***	
	(0.0120)		(0.0103)	
Distance^2	-0.0819***		-0.137***	
	(0.0182)		(0.0160)	
Distance * Wartime	-0.0861***		-0.127***	
	(0.0255)		(0.0255)	
Distance^2 * Wartime	0.127***		0.172***	
	(0.0450)		(0.0379)	
Test for Difference in		-0.431**		0.0910
Ratio of Coefficients		(0.207)		(0.142)
Distance to:	Center of Judge Pool		Supreme Court	
Circuit Fixed Effects	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y
N	8760	8760	8760	8760
R-sq	0.128		0.125	

Notes: Results of regression 9. Robust standard errors clustered at the circuit-year level in parentheses (* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$). Data on cases comes from OpenJurist (1950-2007). Absolute value of the distance to the center of the judge pool (columns 1 and 2) and absolute value of the distance to Supreme Court (columns 3 and 4) are the main independent variables of interest. The dependent variable is the sum of a judge's rates of dissent and concurrence in a Circuit-year. The test statistic reported in columns (2) and (4) is the difference between $\frac{-b}{2c}$ and $\frac{-(b+d)}{2(c+e)}$ in equation (9) for the distance to the center of the judge pool and the distance to Supreme Court respectively. Observations are weighted by the number of votes cast by the judge in the Circuit-year.

APPENDIX TABLE A.11

DISSENT OR CONCURRENCE AND IDEOLOGY SCORE AMONG JUDGES: TESTS FOR CHANGES IN WHO DISSENTS THE MOST DUE TO AN INCREASED COST OF DISSENT FOR NEW JUDGES

	(1)	(2)	(3)	(4)
	Dissent or Concur			
Distance	0.0714***		0.0702***	
	(0.0103)		(0.00981)	
Distance^2	-0.0727***		-0.102***	
	(0.0156)		(0.0149)	
Distance * New Judge	-0.0660***		-0.0631***	
	(0.0156)		(0.0167)	
Distance^2 * New Judge	0.110***		0.101***	
	(0.0360)		(0.0350)	
Test for Difference in		-0.564*		6.190
Ratio of Coefficients		(0.300)		(406.2)
Distance to:	Center of Judge Pool		Supreme Court	
Circuit Fixed Effects	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y
N	10038	10038	10038	10038
R-sq	0.126		0.121	

Notes: Results of regression 9. Robust standard errors clustered at the circuit-year level in parentheses (* p $<$ 0.10; ** p $<$ 0.05; *** p $<$ 0.01). Data on cases comes from OpenJurist (1950-2007). Absolute value of the distance to the center of the judge pool (columns 1 and 2) and absolute value of the distance to Supreme Court (columns 3 and 4) are the main independent variables of interest. The dependent variable is the sum of a judge's rates of dissent and concurrence in a Circuit-year. The test statistic reported in columns (2) and (4) is the difference between $\frac{-b}{2c}$ and $\frac{-(b+d)}{2(c+e)}$ in equation (10) for the distance to the center of the judge pool and the distance to Supreme Court respectively. Observations are weighted by the number of votes cast by the judge in the Circuit-year.

D.2 The SCOTUS model

In this section we present a hierarchical model that is able to produce the spider-shaped pattern of dissent rate as a function of ideological distance to SCOTUS and we derive the predictions that are presented in Proposition 6 and are tested empirically in Section D.1.15. The hierarchical model presented here is developed along the lines of the model of Beim et al. (2014).

Every period, three judges are randomly and independently drawn from a uniform distribution of types $t \sim U\ (-1, 1)$ to sit together on a panel. The panel produces a *binary* verdict ("conservative" or "liberal") for a case with characteristics x, where $x \in [-1, 1]$. A judge of type t prefers a conservative verdict over a liberal one iff $x < t$. The panel determines the verdict by a majority voting, implying that the verdict is conservative if and only if the median judge, denoted t_m, is such that $x < t_m$.

A panel member may also dissent. Upon noticing a dissent, the Supreme Court may decide to review the case. The bliss point of the Supreme Court is normalized to 0. Thus, the Supreme Court rules conservatively on a reviewed case iff $x < 0$. The cost of dissenting is denoted W (and represents writing costs of the minority opinion or collegial pressure). When a judge t is able to reverse the binary verdict in case x her utility gain is $|t - x|$. Hence, a judge will never dissent if $|t - x| \leq W$. We can calculate the

type-dependent probability of dissent $P(t)$ while considering only a judge with $t > 0$ (by symmetry the same applies to judges $t < 0$).

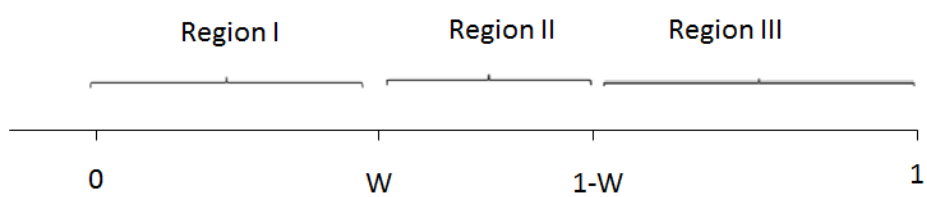
Under this framework, judge $t > 0$ may dissent in two scenarios:

1. $t_m > 0$ and $x \in [t, t_m]$, so that both t and the Supreme Court prefer a liberal verdict while the panel produces a conservative verdict.
2. $t_m < 0$ and $x \in [t_m, 0]$, so that both t and the Supreme Court prefer a conservative verdict while the panel produces a liberal verdict.

Under scenario 1, the judge indeed dissents if $x - t > W$, i.e., if $x \in [t + W, t_m]$. Under scenario 2, the judge indeed dissents if $t - x > W$, i.e., if $x \in [t_m, \min\{0, t - W\}]$. We will now show that the model produces a spider-shaped pattern of dissent rate for $W < 1/2$. In this case, we have $W < 1 - W$.

Figure 11 is helpful in distinguishing between three regions of t.

Appendix Figure 11.— Regions in SCOTUS model



A judge t in region I dissents if either $x \in [t + W, t_m]$ or $x \in [t_m, t - W]$.

A judge t in region II dissents if either $x \in [t + W, t_m]$ or $x \in [t_m, 0]$.

A judge t in region III dissents if $x \in [t_m, 0]$.

Calculating the type-dependent probability of dissent $P(t)$, we get (each line represents one region in the graph)

$$P(t) = \begin{cases} \frac{1}{3}\left[\frac{1}{2}\left[1 - (t + W)\right]\right]^3 + \frac{1}{3}\left[\frac{1}{2}\left[(t - W) - (-1)\right]\right]^3 & \text{if } t \in [0, W] \\ \frac{1}{3}\left[\frac{1}{2}\left[1 - (t + W)\right]\right]^3 + \frac{1}{3}\left(\frac{1}{2}\right)^3 & \text{if } t \in [W, 1 - W] \\ \frac{1}{3}\left(\frac{1}{2}\right)^3 & \text{if } t \in [1 - W, 1] \end{cases}$$

To understand the calculations of the expression of $P(t)$, note first that the event $x \in [t_m, 0]$ is independent of t and it occurs iff $\min t < t_m < x < 0$. As $\min t, t_m$ and x are all drawn from a uniform distribution over $[-1, 1]$, the probability that all three of them are negative is $\left(\frac{1}{2}\right)^3$, and the probability that x is the largest among the three is $1/3$, yielding the expression $\frac{1}{3}\left(\frac{1}{2}\right)^3$. Next, the event $x \in [t_m, t - W]$ is an adjustment of this calculation for the event $x \in [t_m, t - W]$. In particular, we now need $\min t < t_m < x < t - W$, so x needs to be the largest of the three uniformly-distributed variables, which all need to be in the region $[-1, t - W]$, and this event corresponds to probability $\frac{1}{3}\left[\frac{1}{2}\left[(t - W) - (-1)\right]\right]^3$ (i.e., the probability of being

negative, 1/2, is replaced with the probability of being smaller than $t - W$, which is $\frac{1}{2}\left[(t-W)-(-1)\right])$. Finally, the event $x \in [t+W, t_m]$ occurs iff $t + W < x < t_m < \max t$. So x needs to be the smallest of three uniformly-distributed variables, which all need to be in the region $[t+W, 1]$, and this event has probability $\frac{1}{3}\left[\frac{1}{2}\left[1-(t+W)\right]\right]^3$.

Differentiating with respect to t yields

$$\frac{dP(t)}{dt} = \begin{cases} -\frac{1}{8}\left[1-(t+W)\right]^2 + \frac{1}{8}\left[(t-W)+1\right]^2 \text{ if } t \in [0, W] \\ -\frac{1}{8}\left[1-(t+W)\right]^2 \text{ if } t \in [W, 1-W] \\ 0 \text{ if } t \in [1-W, 1] \end{cases}$$

It is immediate to see that $dP(t)/dt$ is negative in region II. To get the sign of $dP(t)/dt$ in region I, note that $t > 0$ implies that $t + W$ is closer to the right edge of the type distribution (1) than $t - W$ is to the left edge of the type distribution (-1), implying that

$$\left[1-(t+W)\right]^2 < \left[(t-W)+1\right]^2,$$

implying that $dP(t)/dt > 0$ in region I. Overall, we get that $P(t)$ increases in region I and then decreases in region II and stays flat in region III, implying a spider-shaped pattern of dissent rate.

D.2.1 Proof of Proposition 6

(i) When $W = 0$, regions I and III disappear and we are left only with region II where $dP(t)/dt$ is negative. Symmetry implies that for any type t, $P(t)$ is decreasing in $|t|$.

(ii) The value of t for which $P(t)$ is maximal is the border between regions I and II, i.e. $t = W$. It is thus immediate that $\arg\max_t P(t)$ increases in W.