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.[42]

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.[43]

[42] 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."

[43] 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 Randomization procedures and check

Interview reports indicate that court randomization practices vary across Circuits. Case assignments fall into two categories: A) Once a case arrives, three randomly chosen judges are assigned to the case. More precisely, a computer program randomly assigns judges to panels two to three weeks before oral arguments. B) The second category is such that, once a year, the calendar is randomly set up in advance determining which judges will sit in which panels on which days in the upcoming year, and when a case comes up it gets assigned to the next panel. Although retired or visiting judges may have a reduced caseload, they are still assigned by a computer algorithm. In some Circuits, senior judges may decline death penalty cases before the random assignment process.

Statistically, we test whether cases are systematically assigned to judges based on their own characteristics. We find that characteristics of cases (the type of case, like criminal, civil rights, first amendment, due process, privacy, labor, economics, or miscellaneous) and characteristics other panelists (proportion of the panel being appointed by Democrats) are not correlated with the judges' own characteristics (absolute ideological distance to the center of the judge pool).

When judges are appointed in the middle of the year, they will be randomly assigned to cases along with the other judges. We construct the center of judge pool on an annual basis, which takes into account the weighted presence of judges during the year. Thus Table A.2 can be interpreted as whether there are significant deviations from random assignment even when judges being appointed mid-year are taken into account.

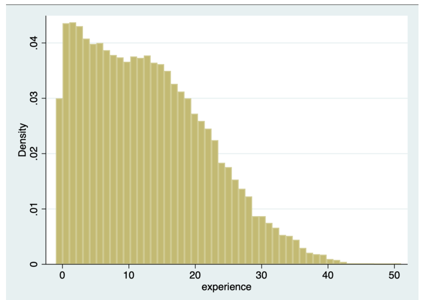
Appendix Table A.2.— Randomization Check

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	% Democrat of Other Panelists	Criminal	Civil Rights	First Amendment	Due Process	Privacy	Labor	Economics	Miscellaneous
Distance to Center of Judge Pool	-0.00492	0.00595	0.00304	-0.000293	-0.00354	-0.000107	-0.000178	-0.00572	-0.00122
	-0.00889	-0.00431	-0.00222	-0.000734	-0.00368	-0.000323	-0.00214	-0.00374	-0.00194
N	659448	659477	659477	659477	659477	659477	659477	659477	659477
R-sq	0.248	0.028	0.014	0.024	0.015	0.008	0.012	0.03	0.014

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). Judicial characteristics come from Federal Judiciary Center/Attributes of U.S. Federal Judges Database. Ideology scores come from the Judicial Common Space database. Column 1 refers to the proportion of other panelists that are Democratic appointees. Columns 2-9 refer to case categories. Circuit-year fixed effects are included.

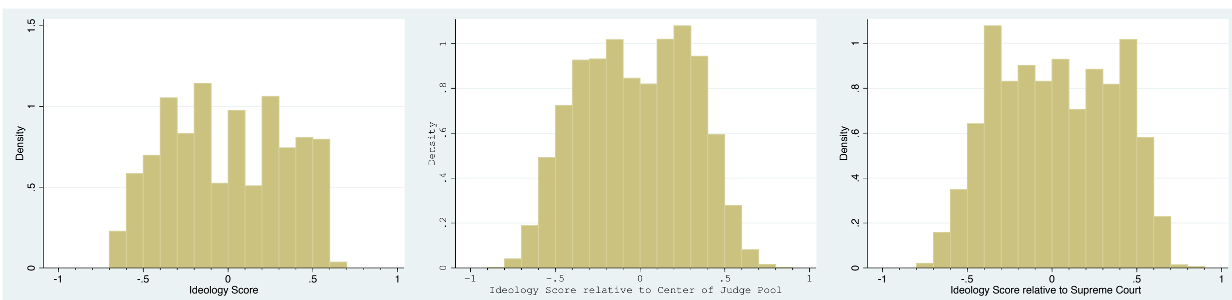
Figure 1 shows that there is a substantial spread of judicial experience over cases. A long right tail can be seen as these are life-tenured positions.

Appendix Figure 1.— Distribution of the tenure of judges



A.1.3 Distribution of ideological scores

Appendix Figure 2.— Distribution of Ideology Scores



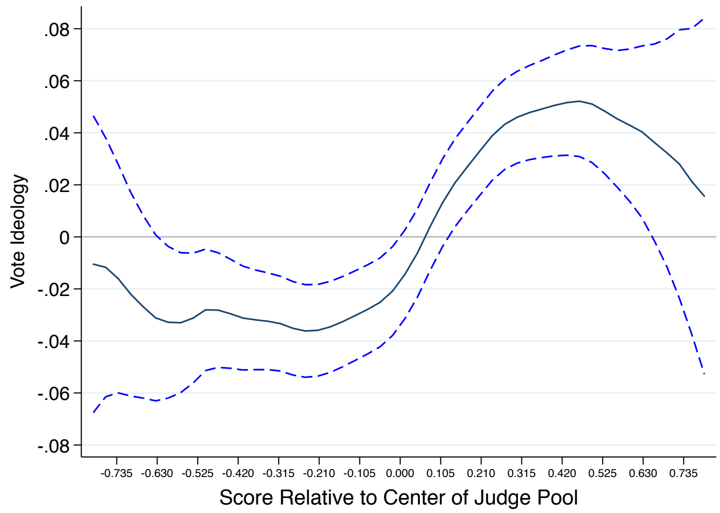
Notes: Ideology scores from the Judicial Common Space database (Epstein et al. 2007). Left panel: raw ideology score. Central panel: ideology score demeaned by the center of the judge pool in a circuit-year. Right panel: ideology score demeaned by the mean ideology score of members of the Supreme Court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

A.1.4 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).[44]

[44] 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.

Appendix Figure 3.— 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).

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 3. 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.[45] The regression results are presented in Appendix Table A.3. 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.

[45] 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.

APPENDIX TABLE A.3

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 average ideology of the pool of judges 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 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.[46]

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

[46] 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.

(see e.g., Schwartz 1992, Hammond et al. 2005 and, most recently, Hangartner et al. 2019).[47] 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.4, we reproduce the finding from Table 6.3 in Cross (2007).[48] 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 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 4 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.

[47] 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.

[48] 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.

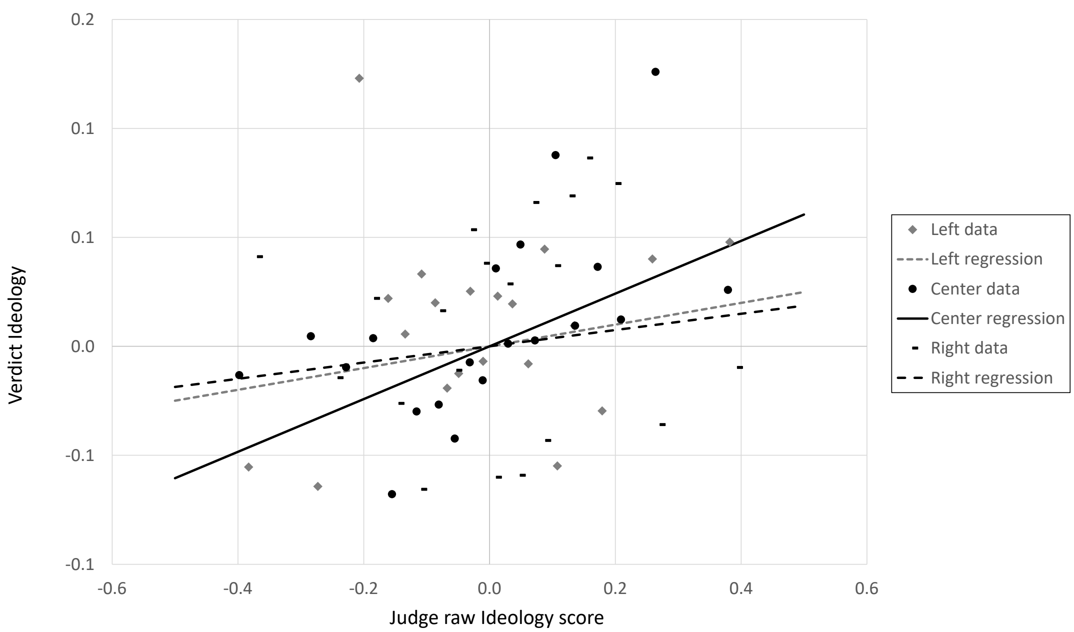
APPENDIX TABLE A.4

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 Ideology Score		0.249*		0.251*
		(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.

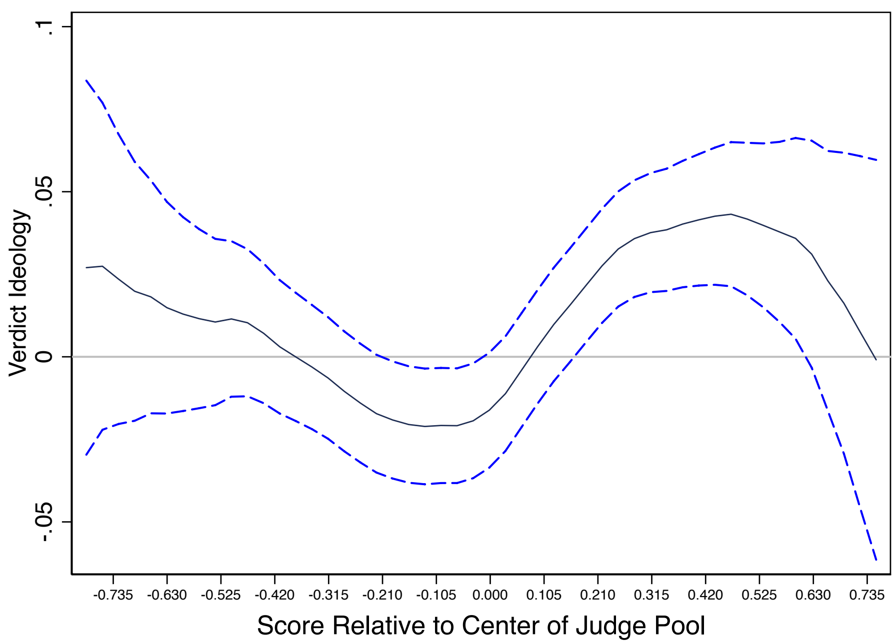
APPENDIX FIGURE 4.— Verdict Ideology and Ideology Scores of Panel Members



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.4. 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 5 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 5.— Verdict Ideology and Ideology Score of Judge Relative to Center of Judge Pool – local polynomial

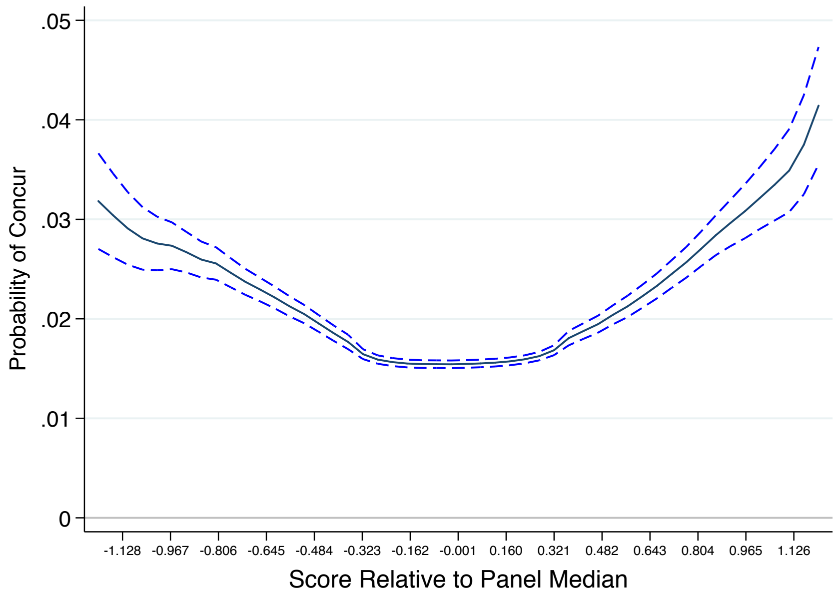


Notes: x-axis: Ideology score of a judge demeaned by the average ideology of the *pool* of judges in a Circuit-year. y-axis: Verdict ideology, demeaned to be centered at zero. The figure presents a local polynomial regression with an Epanechnikov kernel, where the dependent variable is Verdict Ideology, 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 6 (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 6.— 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.5 shows that the result—both for dissents and for concurrences—is robust to using a Logit specification.

APPENDIX TABLE A.5

Dissent and Ideological Distance to Median of Panel

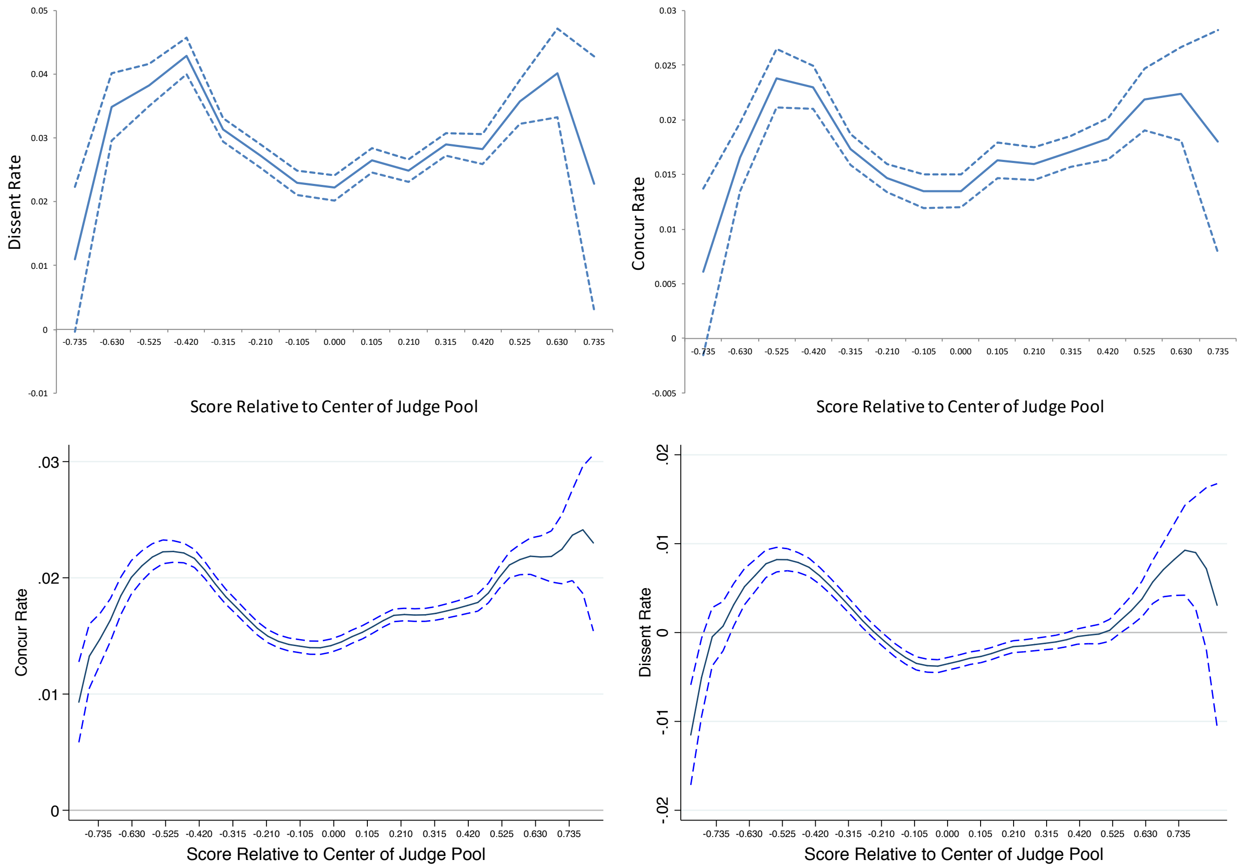
	(1)	(2)
	Dissent	Concur
Distance to Median of Panel	1.192***	0.666***
	(0.119)	(0.142)
Distance^2	-0.252*	0.0437
	(0.140)	(0.171)
Judge Fixed Effects	Y	Y
Circuit Fixed Effects	Y	Y
Year Fixed Effects	Y	Y
N	540385	513597

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 7 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.[49] As can be seen the spider pattern appears clearly here too.

Appendix Figure 7.— 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 average ideology of the *pool* of judges 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.

[49] 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 7 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 7 (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.6 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.7 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.6.— 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.7.— 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.8 (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.8

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 average ideology of the pool of judges 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. Column 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.

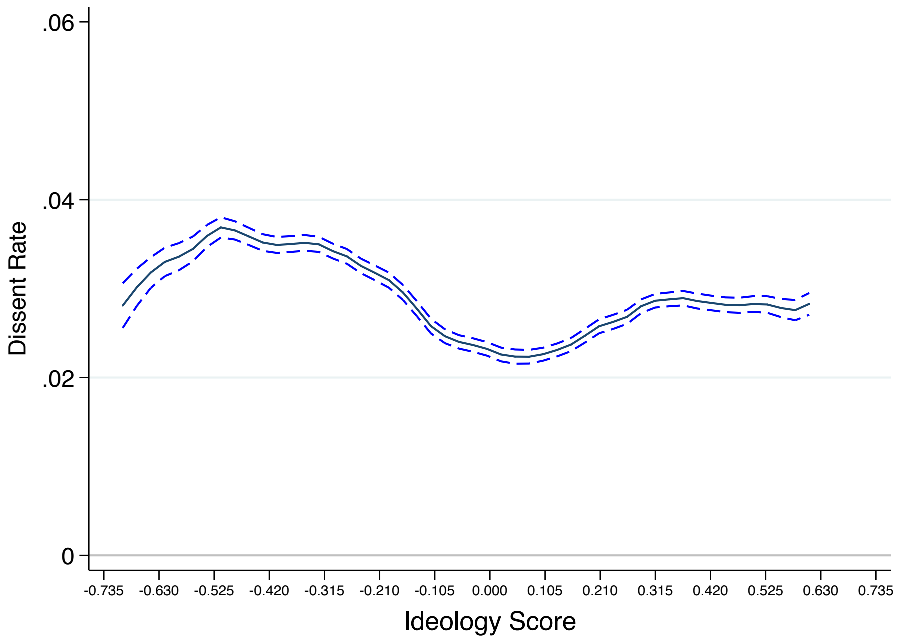
Appendix Table A.8 (columns 3 and 4) presents the main specification but using two forms of bootstrapping. As can be seen, the spider result is robust and the p-values similar as in the main specification.

Appendix Table A.8 (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.

Appendix Figure 8 shows the raw data (local polynomial) using the non-demeaned ideology score as the independent variable.[50] 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.

[50] Figure 8 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 2).

Appendix Figure 8.— 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.

Finally, Appendix Table A.9 repeats the main specification in (3) but using Circuit-year Fixed Effects. The only coefficient whose statistical significance becomes weaker is the quadratic coefficient in column (1).

APPENDIX TABLE A.9

Dissent and Ideological Distance to Center of Judge Pool with Circuit-year FE

	(1)	(2)	(3)
	Dissent Rate	Concur Rate	Dissent or Concur Rate
Distance to Center of Judge Pool	0.0334***	0.0295***	0.0588***
	(0.00773)	(0.00609)	(0.0107)
Distance^2	-0.0209*	-0.0344***	-0.0527***
	(0.0123)	(0.00899)	(0.0163)
Circuit-Year Fixed Effects	Y	Y	Y
N	10040	10040	10040
R-sq	0.222	0.171	0.243

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 average ideology of the pool of judges 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 Circuit-year. Observations are weighted by the number of votes cast by the judge in the Circuit-year.

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.1.4—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.1.4.

B.2 Predicting the effect of collegial pressure

Here we derive two additional predictions that we will use in Section D.1 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 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.[51] 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

[51] 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.[52] 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 ($\alpha<1$ implies that for any W there exists a small enough $\bar{\tau}$ s.t. $\nabla L > 0$ 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

[52] 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 $lim_{W\to\infty}\overline{\tau} = 0$). 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 $L' = 0 \Longleftrightarrow \alpha(\tau - P)^\alpha = W \Longleftrightarrow P^* = \tau - (\frac{W}{\alpha})^{1/\alpha}$. 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 $E\left[J\left(t\right)\right]$ by t we get

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

$$= 2F(t)f(t)\left(2\left[1-F(t)\right]-F(t)+1\right)=6F(t)f(t)\left[1-F(t)\right]>0.$$

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

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

It thus follows that $argmax_{t}\mathrm{E}[J\left(t\right)]\notin\{0,1\}$ 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, $\tau=Prob(v\left(y\right)\neq v_{t}\left(y\right))$ 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: An alternative Explanation

D.1 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.[53] 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

[53] 2% of the Circuit Court decisions are appealed to the Supreme Court, of which 30% are affirmed.

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 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→ 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.[54] 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).[55]

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 9) and that caseload continues to decline gradually thereafter. Next, we show that retired judges dissent more, discontinuously at the year of retirement (Figure 10), and verify that the increase in dissents is not due to age. In fact, older judges are less likely to dissent, which also explains the decline in dissent before retirement. Figure 10 visualizes the following regression.[56]

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

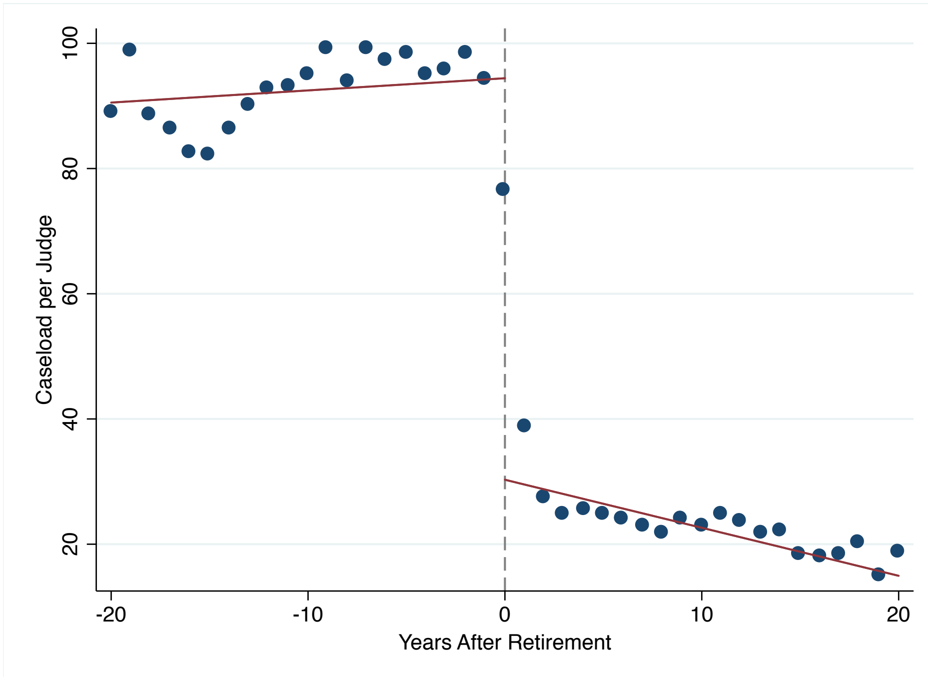
for judge i and year t.

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

[55]A histogram of **Score Relative to Supreme Court** is presented in the right panel of Appendix Figure 2.

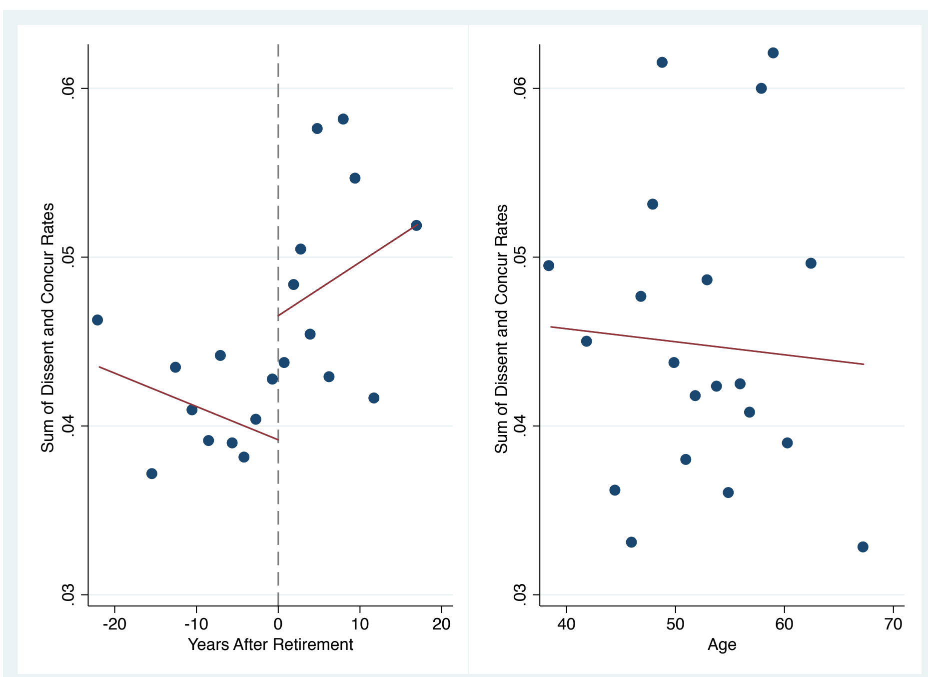
[56]We have verified that age and experience vary smoothly around the retirement decision.

Appendix Figure 9.— 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).

Appendix Figure 10.— 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.10, 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.11 (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.10 and A.11), 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.10

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.11

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.

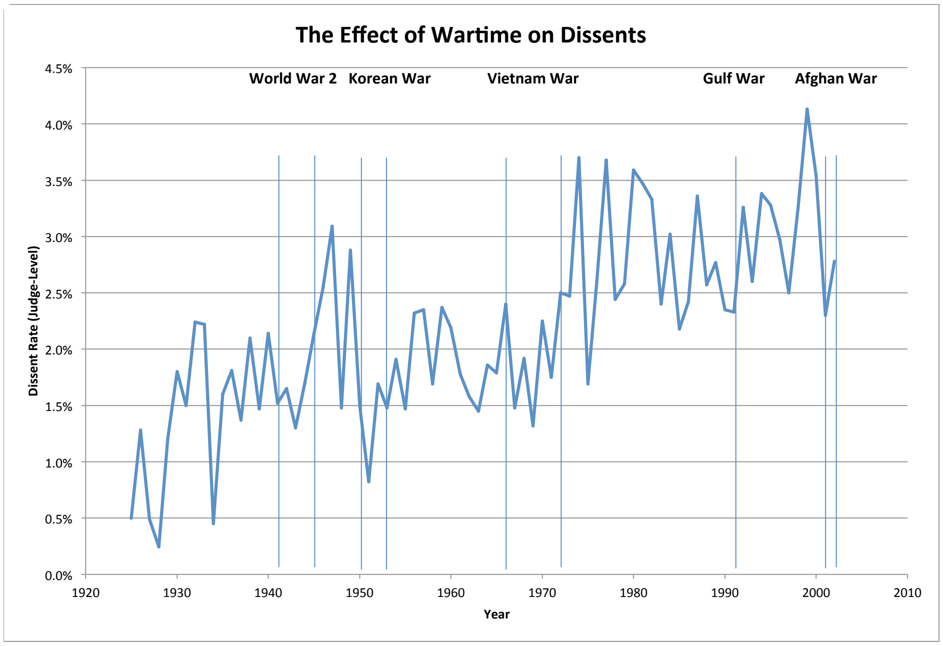
Testing predictions about the most dissenting judge

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 11). 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.

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

for judge i in Circuit c and year t.

APPENDIX FIGURE 11.— 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.[57] Since this is a variation at the judge level (rather than at the circuit level like wartime), the specification we test is the following.

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

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.12 and A.13 respectively for the two tests.

[57]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.

Using **Distance to Center of Judge Pool**, Column 2 of Table A.12 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.13 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.12

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 Ratio of Coefficients		-0.431**		0.0910
		(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.13

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. 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:

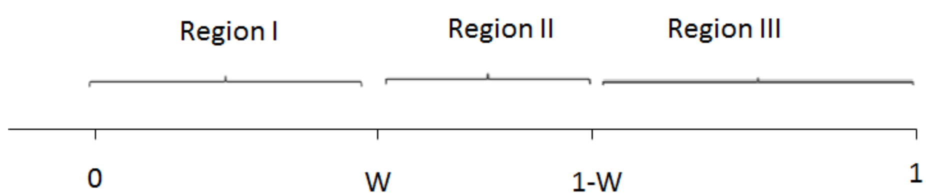
1. $t_m > 0$ and $x \in [t, t_m]$, so that both t and the Supreme Court prefer a liberal verdict while the panel produces a conservative verdict.

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 12 is helpful in distinguishing between three regions of t.

Appendix Figure 12.— 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.