PREFERENCES FOR PERFECTION: PRIVATE VS. PUBLIC TRUTHS ON JUDICIAL PANELS

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

Abstract: How do individuals perceive the cost of taking actions they disagree with politically or morally? Recent experiments suggest that individuals perceive even small deviations from convictions to be very costly, but once a small deviation has been made, further deviations will entail relatively little additional cost. That implies, individuals tend to give up on their morals if they cannot follow them fully, suggesting a concave cost of deviation. This paper presents a novel and puzzling phenomenon in judicial decisions and show that concave ideological preferences explains this phenomenon along with a number of related empirical facts.

Keywords: Deontological Motivations, Judicial Decision-Making, Concave Costs, Ideology
JEL codes: D7, K0, Z1
Preliminary and Incomplete

*Daniel L. Chen, ETH-Zurich, chendan@ethz.ch; Moti Michaeli, EUI, motimich@gmail.com; Daniel Spiro, Unversity of Oslo, Daniel Spiro daniel.spiro@econ.uio.no.

1 Introduction

Economics tends to gravitate towards the assumption that costs – be they economic, private effort or cognitive – are convex. One rationale for this assumption is that it makes theoretical models analytically tractable. Another rationale is that it seems intuitively plausible. However, such intuition has proved fragile following a number of recent experiments showing that, when it comes to decision making on moral and ethical issues, individuals perceive a concave cost of deviating from what they believe is right. That is, already a small deviation entails a high cost while a larger devitation incurs only a small additional cost. Whether individuals perceive costs to be concave has far reaching implications for the behavior we should expect of them and, as such, it goes to the heart of the empirical predictions theoretical models provide. For instance, individuals with concave moral costs will tend to give up on their morals if they cannot follow them fully. This pattern of behavior has been popularly labeled the "what-the-hell-effect" (Ariely 2012; Baumeister and Heatherton 1996). Theoretically, the concavity of costs associated with deviations from a moral and political bliss point also affects: Polarization in society (Kamada and Kojima 2014); optimal policy when the policy space is multi-dimensional (Eguia 2011); whether regime collapse will be initiated by political insiders or extreme opposition (Michaeli and Spiro 2014); whether effort to achieve high status falls when a person's relative status falls (Clark and Oswald 1998); and whether judicial and social law obedience should be expected by those who are inclined to follow it or by those who have the most inclination to break it. Experimental research has shown that concave preferences are indeed plausible: Where subjects could choose different degrees of cheating, it was shown that most subjects either cheated fully or not at all (Hurkens and Kartik 2009). Likewise, using a dynamic setting, Gino et al. (2010) showed that once individuals are induced to cheat, they succumb to full-blown cheating. Gneezy et al. (2013) show experimentally that the perceived cost of lying is concave in the extent of lying.[1]

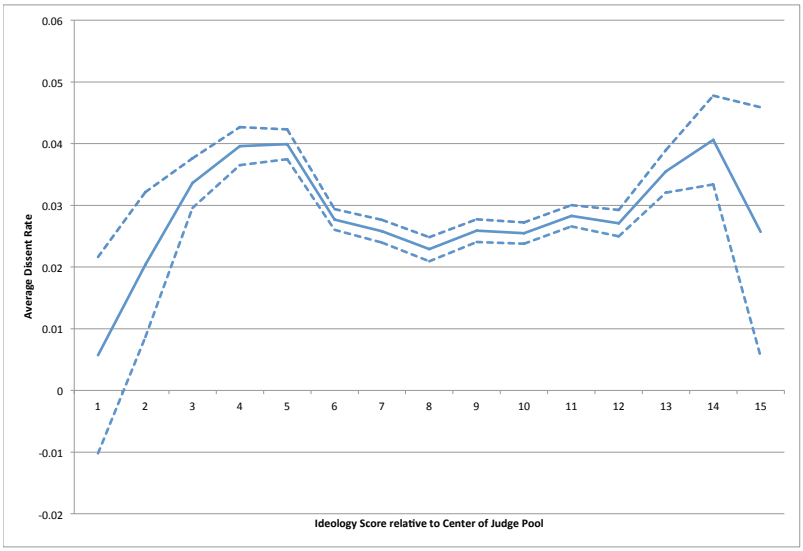
The question remains whether concave preferences have empirically observable implications for important real world decision situations. Showing that they do is the purpose of the current paper. We present a novel and puzzling phenomenon in judicial decisions and show that concave ideological preferences explain this phenomenon along with a number of related empirical facts. In the U.S. Federal Circuit Courts, three judges are randomly assigned to sit in each panel. This means that each judge will, over a period of time, sit in a large number of panels, where the composition of each panel is random. These panels often have ideological and political dimensions; Sunstein et al. (2006) document that a wide range of legal decisions are predicted by panel composition. Apart from the binary choice of verdict (overturn or affirm the lower court decision), a panel also needs to compose an "opinion" (i.e., a text) motivating the verdict. The opinion often serves as precedent for future cases. Furthermore, being a text rather than a binary element, the opinion can reflect the assertiveness of the panel. Finally and importantly, not all judges need to sign this opinion (it suffices to have only one judge signing), but there is a strong collegial pressure for having unanimity: Seventy percent of panels have both Democrats and Republicans, but only 8% have dissents (Berdejó and Chen 2013). Hence, dissenting by not signing comes at a cost, not only by requiring a separate

[1] Kendall et al. (2015) have shown by structural estimation that voters are most likely to have a concave loss function when it comes to deviations from their ideological bliss point. Further, as has been argued by Osborne (1995), in issues of ideology it may be realistic to assume a concave cost function.

opinion, but also in collegiality (Epstein et al. 2011). Dissents also have real policy impacts: About 25% of Circuit Cases are appealed, and a dissent adds 13% points to this appeal rate. The Supreme Court takes 3% of the cases; cases with dissents have 4% points greater likelihood to be heard by the Supreme Court. The Supreme Court reverses 71% of the cases it hears.

How would a judge's ideology affect her choice of when and how often to dissent? Figure 1 shows the rate of dissent on the y-axis as a function of judges' ideological score on the x-axis, where the most liberal judges are on the left and the most conservative on the right (see the next section for more details). Looking at this figure, a striking and empirically robust pattern emerges: Centrist and extremist judges dissent very seldom, while it is the judges with moderate ideologies, on both left and right, who dissent the most.

FIGURE 1.— Dissent Rate and Ideology Score (1950-2007)



Notes: Data on cases comes from OpenJurist. Ideology scores come from the Judicial Common Space database. Ideology scores are demeaned by the center of the pool of judges available to be assigned in a Circuit-year. Bins are evenly spaced from left-to-right in ideology space.

How can this non-monotonic relationship between ideological extremism and dissent be explained? Using a three-person-bargaining model, we show that if judges have a sufficiently concave ideological cost of signing opinions they do not like, and the cost of dissent is high, the observed pattern emerges. We also test a number of auxiliary model predictions. The intuition for the result is as follows. Consider a judge who is very extreme. Looking across panels, this judge can be expected to be ideologically distanced from the two other judges in most panels. Hence, being in minority, she will very often find herself having to decide whether or not to sign opinions very far from her ideological bliss point. Always dissenting on opinions she does not like would therefore imply a very high collegial pressure as she will be facing such opinions virtually all the time. On the other

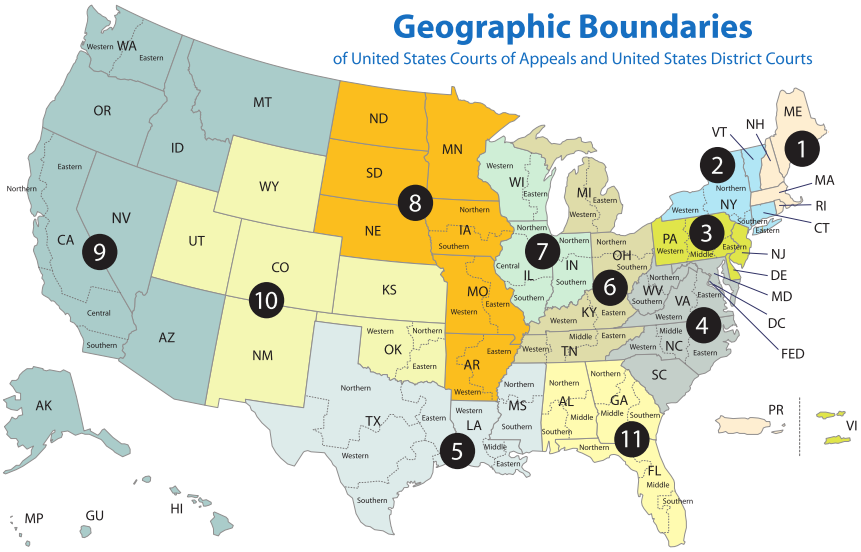
hand, signing some opinions while dissenting against others helps little when the cost incurred by signing few unfavorable opinions instead of many is almost the same. Hence, facing a sufficiently high collegial pressure, an extreme judge will tend to sign virtually all opinions. In comparison, more centrist judges are bound to more often be able to determine the panel's opinion and hence will rarely need to sign unfavorable opinions. In the few occasions where they do need to sign those, they might as well dissent. This means that the less extreme the judge is, the more likely she is to dissent when in minority. However, since centrist judges seldom face such decisions, they will still have a low dissent rate. Vice versa, extremist judges will often face the decision of signing unfavorable opinions but will rarely dissent when in that situation. Finally, judges in between centrists and extremists will often dissent when facing an unfavorable opinion and will face this situation from time to time and hence end up, in aggregate, dissenting the most of all judges. While a concave ideological cost function is not sufficient for this non-monotonicity, we show it is a necessary condition, as this line of reasoning does not work if the ideological cost is linear or convex.

The structure of the paper is as follows. In the next section we outline the judicial process more in detail and describe the data and estimation underlying Figure 1. In Section 3 we describe the model and its results including a number of auxiliary predictions. In Section 4 we test these empirical predictions. Section 5 concludes.

2 Judicial Dissents

2.1 Institutional Background The U.S. Circuit Courts play an important role in the law-making function of common law courts. This making of law occurs because a judge's decisions in current cases become precedent for use in future cases in the same court and in lower courts of the same jurisdiction. There are three layers of federal courts: District, Circuit, and the U.S. Supreme Court. The 94 U.S. District Courts serve as the general trial courts, where a jury is drawn to decide *issues of facts*. If a party appeals the decision, the case goes up to a Circuit Court, which decides *issues of law*; they have no juries. The 12 U.S. Circuit Courts, also known as Courts of Appeals or federal appellate courts, are only to hear cases presenting new legal issues (only 10-20% of District Court opinions are appealed). Cases that reach the Circuit Courts are the more challenging and controversial cases with the greatest likelihood to set new precedent. Figure 2 displays District Court boundaries in dotted lines and Circuit Court boundaries—encompassing between 5 and 13 Districts each—in solid lines.

Figure 2.— Dissent Rate and Ideology Score (1950-2007)



Notes: Data on cases comes from OpenJurist. Ideology scores come from the Judicial Common Space database. Ideology scores are demeaned by the the center of the pool of judges available to be assigned in a Circuit-year. Bins are evenly spaced from left-to-right in ideology space.

In deciding issues of law, Circuit Courts provide new interpretations or distinctions of pre-existing precedents or statutes. These new distinctions expand or contract the space under which an actor is allowed to act (Gennaioli and Shleifer 2007). For example, when the Fifth Circuit in March 2014 upheld a Texas statute requiring doctors in abortion clinics to obtain admitting privileges at local hospitals, one-third of its abortion clinics shut down, leaving only 22. Furthermore, a new Texas statute requires abortion clinics to meet the building standards of ambulatory surgery centers; the statute was allowed by the Fifth Circuit in the Fall of 2014 while it considered an appeal to invalidate the new statute. This statute would reduce the number of centers operating in the state to fewer than 10 (Chen et al. 2011).[2] This example also highlights the power of the U.S. Federal Courts in determining what is the law in both judicial and legislative matters.

Each Circuit Court decides many thousands of cases per year, but only 2% of Circuit cases successfully appeal to the U.S. Supreme Court, so Circuit Courts determine the vast majority of decisions each year that set legal precedent. Each Circuit Court case receives *three randomly assigned judges* out of a pool of judges, presently numbering 17 on average depending on the size of the Circuit. These judges are appointed for life by the U.S. President and their positions and decisions are highly esteemed.[3] Dissents (2-1) decisions are not the only way for judges to express their separate views. They can also concur, which means they agree with the majority opinion but for a different reason. Like dissents, concurrences are also costly in terms of time and collegiality; nor can they be cited as binding precedent. Further, even when they agree with the majority opinion for the same reason, judges can attribute authorship to a particular judge or present a unified front, an unsigned opinion (per curiam). The latter is typically reserved for uncontroversial cases, but what is

[2] http://www.nytimes.com/2014/07/30/us/mississippi-abortion-clinic-Federal-court-blocks-closing.html

[3] Except for retirement, Circuit judges typically leave the bench only for a position in the U.S. Supreme Court.

considered uncontroversial is subject to the judges agreeing the decision to be without controversy.

2.2 Data The data we present come from Openjurist where we collected all cases from 1950 to 2007. We coded whether there was a dissenting opinion, who wrote the opinion, whether there was a concurring opinion, whether the opinion was unsigned (i.e., per curiam), and how the lower court was treated (affirm or reverse, reverse with remand, reverse without remand). We merge each judge with their ideology score, a summary measure using the voting patterns of the appointing president and home state senators, coming from the Judicial Common Space database (Epstein et al. 2007). To buttress assumptions used in our theoretical model, we also employ the U.S. Courts of Appeals Database Project, a random sample of roughly 5% from 1925 to 2002.[4] This database includes additional hand-coded information on decision valence (liberal = 1, conservative = -1, and mixed or unable to code = 0)[5] and is linked by hand to the U.S. Supreme Court, with any subsequent outcomes using the Shepardized Courts of Appeals database. Our sample contains 293,868 decisions for 1950 to 2007 and 18,686 decisions for the period 1925 to 2002. Overall, 7.9% of opinions from 1925 to 2002 have dissents while 6.2% of opinions from 1950 to 2007 have dissents with dissenting opinions.

2.3 Specification We begin by calculating the average ideology score of the pool of judges in each Circuit and each year. This average score represents the center of the pool of judges available to be assigned. Next, we calculate the distance from each judge's ideology score to the center of her pool for each Circuit and each year. Intuitively, we might expect that the more distant a judge is from her center in a Circuit-year, the more she will dissent and write separate opinions (including concurrences) and the less she agrees to unsigned opinions (per curiam opinions are presented to the public as a unified voice of the panel without need to attribute authorship, or relatedly, an uncontroversial opinion). We calculate the dissent rate, concurrence rate, and per curiam rate for each judge in each Circuit-year.

We present a non-parametric visualization of the dissent rates by ideological distance to the center of a judge's pool. For each of 15 evenly-spaced bins from the left-most to right-most score in each Circuit-year, we estimate the average of the dissent rates of judges in each bin.[6] The average is a weighted average to account for the number of times the judge actually appeared on cases in that Circuit-year. We repeat the exercise for each judicial behavior.

2.4 Spiders in Dissent and Concurrence Figure 1 presents the average yearly dissent rate of judges according to the distance to the center of their respective pool of judges. We see the most extreme left judges rarely dissent, a marked increase in dissent rates as judges become more centrist (bins 1 to 5), a decrease in dissent rates towards the center of the judicial pool, and then an increase and decrease as judges move farther out to the right. We will refer to this pattern as a spider due to its resemblence of the body and legs of a spider.

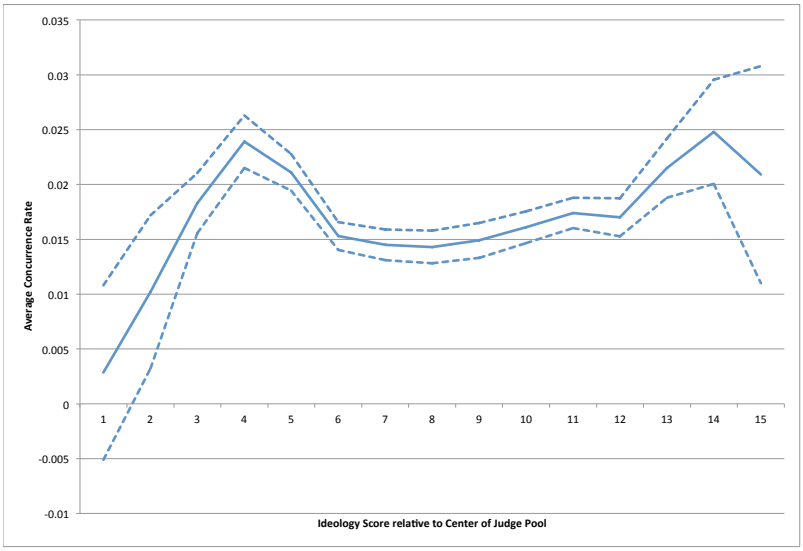
[4]Documentation and data available at http://www.cas.sc.edu/poli/juri/appctdata.htm.

[5]The Appeals Court Database Project states that for most, but not all issue categories, these will correspond to notions of "liberal" and "conservative" that are commonly used in the public law literature. For example, decisions supporting the position of the defendant in a criminal procedure case, the plaintiff who asserts a violation of her First Amendment rights, and the Secretary of Labor who sues a corporation for violation of child labor regulations are all coded as "liberal."

[6]We use Circuit and year since this represents the ideology of the other judges a specific judge will actually be siting with in a panel.

Figure 3 presents the average yearly concurrence rate of judges according to the distance to the center of their respective pool of judges. We analyze concurrence separate from dissents as these are legally distinct. Notably, the pattern is robust: Concurrence rates are also surprisingly uncommon for the most extreme judges. Concurrence rates are higher for judges towards the center, and lower again for the judges at the very center. This finding decreases the likelihood that these patterns are statistical artifacts.

FIGURE 3.— Concurrence Rate and Ideology Score (1950-2007)



Notes: Data on cases comes from OpenJurist. Ideology scores come from the Judicial Common Space database. Ideology scores are demeaned by the center of the pool of judges available to be assigned in a Circuit-year. Bins are evenly spaced from left-to-right in ideology space.

We also 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. We include Circuit and year fixed effects and cluster the standard errors by Circuit-year. Table I indicates that the pattern is robust and Figure 4 shows the pattern obtained using the estimates from the table. Results reported in the table and figure are robust to other polynomials and to using logit.

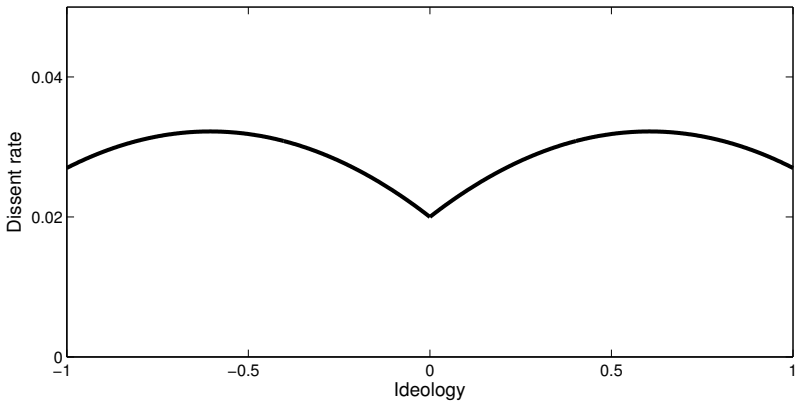
TABLE I

Dissent in Polynomial Distance to Expected Center (1950-2007)

	(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

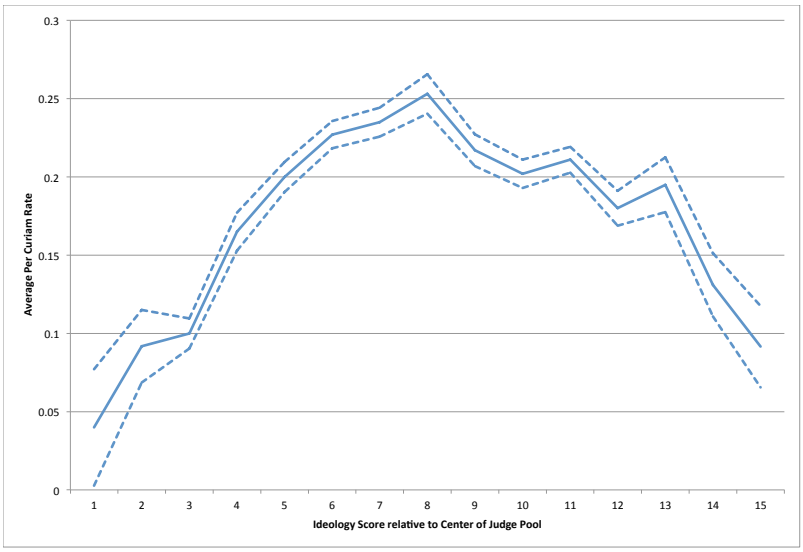
Notes: Robust standard errors clustered at the circuit-year level in parentheses (* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$). Data on cases comes from OpenJurist. Ideology scores come from the Judicial Common Space database. Ideology scores are demeaned by the center of the pool of judges available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

Figure 4.— Predicted Dissent Rate



2.5 Per Curiam This section show that at the extreme ideologies, judges are far less likely to agree to issue a per curiam (jointly issued as an opinion of the court without a designated author) opinion and slightly more likely to author an opinion. We examine whether judges choose *not* to sign an opinion. This can occur when all judges agree there is no controversy whatsoever in the case. Figure 5 shows that the more distant judges are very unlikely to agree to an unsigned opinion, which suggests that extreme judges have the largest number of objections to others' opinions. The figure can also be viewed as robustness as the pattern of behavior in the extreme left and right bins do not appear to be statistical artifacts.

FIGURE 5.— Per Curiam Rate and Ideology Score (1950-2007)



Notes: Data on cases comes from OpenJurist. Ideology scores come from the Judicial Common Space database. Ideology scores are demeaned by the the center of the pool of judges available to be assigned in a Circuit-year. Bins are evenly spaced from left-to-right in ideology space.

3 Theoretical Model

Three judges (t_1, t_2 and t_3) are randomly and independently drawn from a symmetric bell-shaped distribution $F(t)$ to sit together on a panel.[7] For simplicity we set the average of the distribution to be 0. The judges in a panel bargain and eventually produce a verdict $v \in \mathbb{R}$ as the bargaining outcome. The verdict with the most support wins. In case of a tie, the median judge, denoted by t_m, determines the verdict.[8] Judges want the verdict to reflect their bliss points (t). Hence, a judge t gets disutility

$$O(v;t)$$

from verdict v, where $O(v;t)$ is increasing in $|v-t|$. Let $g(v|t)$ be the distribution of verdicts that judge t is facing during her judicial term, with support in $V \subset \mathbb{R}$.[9] For each verdict v, she needs to decide whether to sign it or to dissent.[10] When signing a verdict $v \neq t$, the judge is making a compromise and has to bear the inner discomfort associated with it. This is captured by a separate

[7] Our assumption on the distribution, which is corroborated empirically in the next section, is that the distribution is single-peaked with symmetric downward slopes at each side of the peak.

[8] This is equivalent to having a convention in which first the median judge suggests a verdict and then the other judges accept it, reject it, or suggest an alternative one.

[9] The distribution of verdicts is dependent of t as judge t herself takes part in setting the verdict through the bargaining process.

[10] We treat dissent and concurrence as the same thing – a refusal to sign the chosen verdict.

cost function

$$D = D\left(\int_V |t-v|\, g(v|t)\, s(v)\, dv\right)$$

where $s(v)$ is an indicator function that equals 1 if the judge chooses to sign a verdict and equals 0 if she dissents. For now we assume D is increasing in its argument but the purpose of the later analysis will be to show the effect of a concave D (representing the what-the-hell-effect) as compared to a convex D. In total, a judge of type t has the following loss function.

$$L = \int_V O(v;t)\, g(v|t)\, dv + D\left(\int_V |t-v|\, g(v|t)\, s(v)\, dv\right) + W\int_V (1-s(v))\, g(v|t)\, dv,$$

where W is a fixed cost of dissent.

3.1 Bargaining outcome From the model specification it is clear that for any signing strategy $s(v)$ chosen by a judge t, L is increasing in $|v-t|$. Hence, each judge will strive to minimize $|t-v|$ during the bargaining process. As a consequence, the verdict in each panel will be solely determined by the bliss point of the median judge in that panel.

LEMMA 1 *In all panels $v = t_m$.*

To see this, note that both the other two judges cannot be better off at the same time by signing any $v \neq t_m$. Hence, these two judges cannot form a coalition that will suggest an alternative verdict. Furthermore, if judge t_m suggests $v = t_m$, then either she gets the consent of one or both of the other judges to sign it, in which case t_m is the final verdict, or, if none of the others agrees to sign, then there is a tie, in which case $v = t_m$ is the verdict by convention.

3.2 To sign or not to sign? The deterministic nature of v implies that the individual decision of a judge whether to sign a verdict or not can be analyzed taking v as given. Clearly, for the median judge in a panel, signing is optimal. As for the other judges in that panel, note first that $O(v;t)$ is independent of signing or not. Therefore, the first argument in the loss function, $\int_V O(v;t)\, g(v|t)\, dv$, does not affect any signing decision made by a judge. Each of these judges therefore minimizes

$$l(s(v);t) \equiv L - \int_V O(v;t)\, g(v|t)\, dv = D\left(\int_V |t-v|\, g(v|t)\, s(v)\, dv\right) + W\int_V (1-s(v))\, g(v|t)\, dv.$$

We will now rewrite this expression in a form that is more convenient for analysis. This can be done by noting that, when minimizing her loss, every judge t has a unique cutoff c such that she signs verdict v if and only if $|t-v| < c$.[11] Thus, the choice of a judge boils down to choosing this cutoff c. Rewriting the loss function, the minimization problem for judge t is given by

$$\min_c l = \min_c \left\{ D\left(\int_{t-c}^{t+c} |t-v|\, g(v|t)\, dv\right) + W\left(1 - \int_{t-c}^{t+c} g(v|t)\, dv\right)\right\}.$$

[11] Otherwise, there exist two verdicts v_1 and v_2, such that $|t-v_1| < |t-v_2|$ yet judge t is willing to sign v_2 while refusing to sign v_1, in which case she could lower her loss by inverting this pair of choices.

The first argument expresses the disutility associated with signing verdicts that deviate from the judge's bliss point and the second argument expresses the disutility associated with not signing the other verdicts – i.e., the collegial pressure or cost of writing a dissenting opinion.

We turn now to explicitly express the verdict distribution $g(v|t)$ in terms of the judges' type distribution $F(t)$ and the corresponding density function $f(t)$. For a given judge t, $G(v < t) = \Pr(v < t) = [F(t)]^2$, as this happens if and only if both other judges have bliss points below t. Similarly, $G(v > t) = \Pr(v > t) = [1 - F(t)]^2$. In the remaining cases judge t is the median, in which case $|t - v| = 0$. This means that $g(v|t) = 2F(v) f(v)$ in the range $v < t$ and $g(v|t) = 2[1 - F(t)] f(v)$ in the range $v > t$. It also follows that the probability of dissent for judge t who uses a cutoff strategy c is given by

$$[F(t-c)]^2 + [1 - F(t+c)]^2,$$

and the minimization problem of judge t becomes

$$\min_c \left\{ D\left(\int_{t-c}^{t} (t-v)\, 2F(v) f(v)\, dv + \int_{t}^{t+c} (v-t)\, 2(1-F(v)) f(v)\, dv \right) + W\left[[F(t-c)]^2 + [1-F(t+c)]^2 \right] \right\}.$$

3.3 The probability of dissent for different types of judges The probability that type t dissents is determined by two main factors. The first one is the distribution of other judges' types and the location of t with respect to this distribution. For any given cutoff c, the probability of dissent is determined by the concentration of judges within and outside this cutoff. A judge at the tail of a bell-shaped distribution is therefore bound to encounter more panels where both other judges are outside her cutoff compared to a judge with the same cutoff but whose bliss point is at the center of the distribution. The second decisive factor is the value of the cutoff itself. For any given judge t, larger c naturally implies less dissent. As c is endogenously chosen taken t into account, the function $c(t)$ has implications for the probability of dissent of each type, denoted by $P(t)$.

Lemma 2 *If $c(t)$ is decreasing in $|t|$ then $P(t)$ is increasing in $|t|$.*

The formal proof is in the appendix but to understand the lemma note that $c'(t) < 0$ refers to a case in which judges at the tail of the distribution are less flexible in terms of their chosen cutoff. On top of this they naturally also find themselves more often in panels where the other judges hold views far from them. Put together these two effects imply that judges will tend to dissent more the further they are from the center of the distribution of judges, i.e., $P(t)$ will increase in $|t|$. That $P(t)$ is throughout increasing in $|t|$ is of course inconsistent with the spider pattern and so the lemma says a necessary condition for the spider of dissent to emerge is that $c'(t) > 0$ for large $|t|$.[12] We test this prediction empirically in Section 4. It also has implications for the shape of the function D, which aggregates the inner cost of judges.

[12] Intuitively, an important condition that needs to be met is for extreme judges to use a sufficiently large cutoff that will compensate for their higher tendency to hold views that are far from their panel mates. Alternatively, central judges should be very picky in their agreements, using a small cutoff in their signing strategy.

LEMMA 3 *If $D'' \geq 0$ everywhere, then $P(t)$ is increasing in $|t|$.*

The lemma states that a convex or a linear $D(\cdot)$ cannot give rise to a spider pattern of dissent (see appendix for the formal proof). When D is linear, a judge's location on the type distribution is irrelevant for her choice whether to sign a verdict or not – only the trade-off between the fixed cost W and the distance $|t - v|$ matters. In this case c is the same for all judges, and so judges that are more often far from the verdict, i.e. extreme judges, tend to dissent more. When D is convex, judges incur an inner cost only when deviating quite a lot from their bliss points (in terms of both the number of times and the deviation size). Thus, a judge's location in the type distribution will be crucial for her choice strategy. Judges in the center of the bell-shaped distribution are rarely allocated to panels where their views are far from the verdict (i.e., the median judge), hence will be willing to sign the verdict in these few rare cases. Meanwhile, judges at the tails of the distribution will encounter many such cases, and will have to consider larger deviations. As a convex D implies a steep rising slope of the inner cost when compromising a lot, these judges will not be willing to compromise a lot. In particular, they will not compromise as much as needed in order to produce the spider pattern, in which these judges dissent less than their more moderate colleagues. This shows that a concave D is necessary condition for the spider pattern to arise. So we turn now to show how a concave inner cost D *can* produce a spider of dissent.

3.4 The "what the hell effect" of dissent

An intuitive special case In order to understand how a concave inner cost D can produce a spider pattern of dissent, it is helpful to consider the special case in which D is a step function. This means that signing even one unfavorable verdict ($|t - v| > 0$) comes at a large inner cost, but signing more than one verdict entails no extra cost. It also means that the inner cost is independent of the actual distance between the verdict and the judge's bliss point once they do not equal. In this case the loss function can be written as

$$l = D * A + P(t) * W,$$

where D is a constant and $A = 1$ for a judge who signs at least one unfavorable verdict and $A = 0$ for a judge who never signs an unfavorable verdict. In the former case, judge t incurs an inner cost D, and so minimizing l implies minimizing the probability of dissent $P(t)$, hence she will choose to never dissent ($c \to \infty$, implying $P(t) = 0$ and $l = D$). This behavior reflects the "what the hell effect" of dissent – once the judge realizes she must compromise sometimes, she chooses to compromise all the time. In the latter case, where $A = 0$, the judge dissents whenever $|t - v| > 0$, i.e., whenever she is not the median judge in the panel. In this case her loss l is determined by her location in the judge distribution, and is given by

$$W\left[[F(t)]^2 + [1 - F(t)]^2\right].$$

Hence, judge t will choose to never dissent if and only if

$$D \leq Wq(t),$$

where $q(t) \equiv \left[[F(t)]^2 + [1 - F(t)]^2\right]$ is the probability that judge t is not the median judge in a panel. As the probability of being the median judge in a panel decreases in $|t|$, there will be a cutoff type $|t_c|$, such that types with $|t| \geq |t_c|$ never dissent while types with $|t| < |t_c|$ always dissent against unfavorable verdicts. Looking now at the distribution of dissent as a function of one's type, we get that

$$P(t) = \begin{cases} q(t) \text{ if } |t| < |t_c| \\ 0 \text{ if } |t| \geq |t_c| \end{cases} .$$

That is, $P(t)$ increases in $|t|$ as t grows from 0 to $|t_c|$, reflecting the increasing probability to face a verdict $v \neq t$, but then abruptly falls to zero. This non monotonicity of the probability of dissent captures the pattern of the spider.

Structural estimation We have now shown that a concave D is a necessary condition and that a particular shape, namely the step function, is sufficient in producing the spider. However, deriving more precise necessary and sufficient conditions for the emergence of the spider pattern is rather difficult. But since we are interested in the predictive power of our model, we will choose a representative concave function for D and perform a structural estimation of the model parameters using this function. Then, one may use the elicited parameters to evaluate the goodness of fit and the predictive power of the theory.

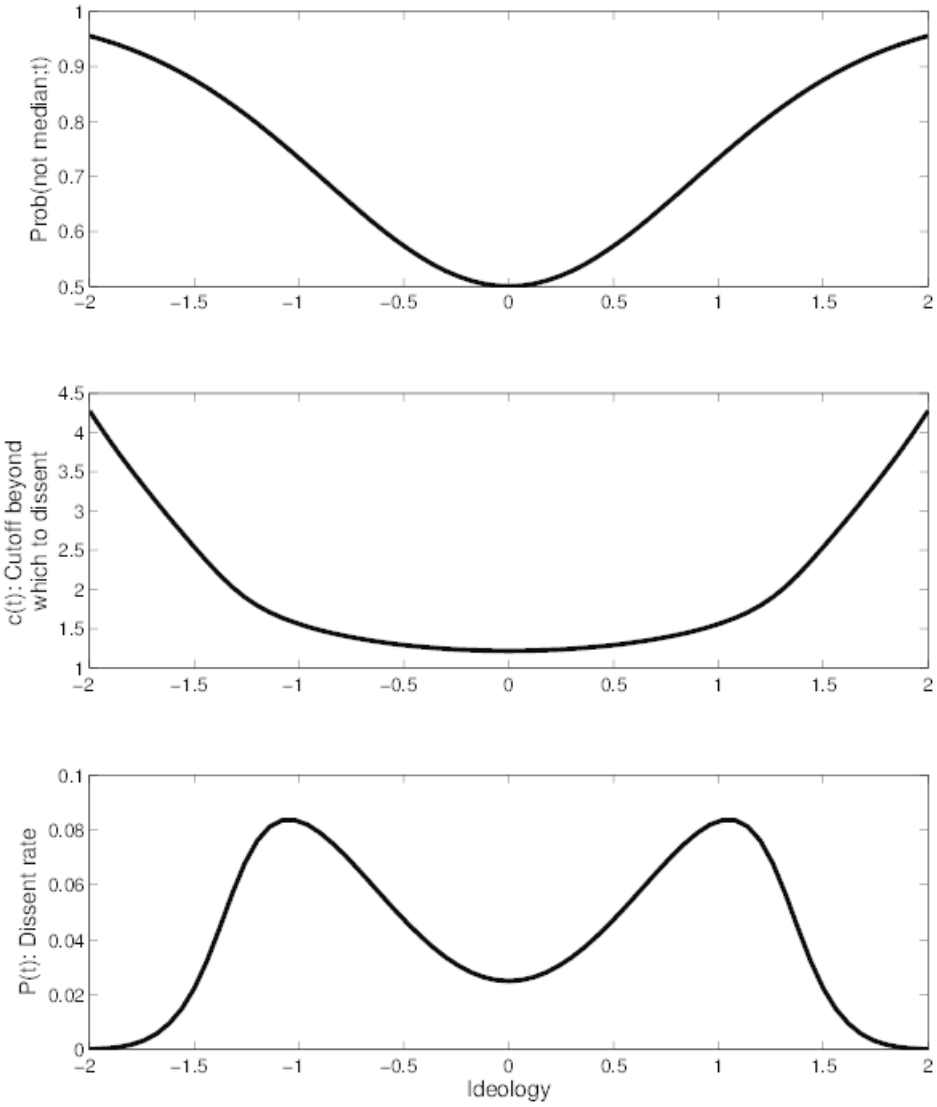
When choosing the D function, one may wish to have a family of functions with one parameter that captures the degree of concavity. Moreover, the inner cost would preferably be bounded from below and from above.[13] One natural choice for a function that satisfies these conditions is the function $K * (1 - e^{-ax})$, where K is a constant cost that can be set to 1 without loss of generality, a captures the degree of concavity and

$$x \equiv \int_{t-c}^{t} (t - v)\, 2F(v)\, f(v)\, dv + \int_{t}^{t+c} (v - t)\, 2\,(1 - F(v))\, f(v)\, dv.$$

Simulating our model with this function and a standard normal distribution to represent judges' bliss points (which fits the data quite well), we get that for a large enough a, i.e. high enough degree of concavity, there exists a range of dissent cost W for which the spider pattern is produced (TBD). One such example is depicted in Figure 6 ($a = 1, W = 1$).

[13] This is to avoid infinite disutitilities and utilities which seem unrealistic. Note for example that this doesn't hold for the log function, which is unbounded at both edges ($\log(0)$ and $\lim_{x\to\infty} \log(x)$) and for power functions x^α with $\alpha < 1$, which are unbounded from above.

Figure 6.— Potential Dissent Patterns



4 Additional Empirical Results

This section will test the empirical predictions obtained in the previous section and the assumptions underlying the model.

4.1 Does the median judge determine the verdict? We seek to investigate whether the result of a median voter holds in the data, i.e., whether Lemma 1. To do this, we examine the correlation between the liberal valence of the decision with the ideology scores of each panel member. Then panel members who are at the center of their panel, at the closer edge of the 3-judge panel, and at the farthest edge of the 3-judge panels are identified. Finally, we regress the liberal verdict (+1/0/-1) on each of the three judges' scores. Table II reveals that only the median judge's ideology score is correlated with the verdict.

TABLE II

Decision Valence and Ideology Scores of Panel Members (1925-2002 5% Sample)

	(1)
	Liberal Verdict
Most Distant Judge Ideology Score	0.0684
	(0.0938)
Center Judge Ideology Score	0.466***
	(0.172)
Closer Edge Judge Ideology Score	-0.142
	(0.153)
N	2986
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 comes from U.S. Courts of Appeals Database Project. The outcome variable is Liberal verdict, which is coded as 1 for liberal, 0 for mixed or not applicable, and -1 for conservative. Fixed effects include year and circuit.

4.2 Is there a cutoff verdict distance beyond which a judge dissents? In the model section we established that it must be optimal for a judge to establish a cutoff distance such that she dissents against any verdict beyond this cutoff and signs any verdict closer to her than the cutoff. Having established theoretically and empirically that the median judge more or less single handedly determines the verdict, testing this boils down to testing whether the probabiliity of dissent is increasning with the distance to the median judge.

$$Prob(dissent\, or\, concur) = a + b * abs(t - t_m)$$

The results are displayed in Table III. The distance to the median judge on this panel is, on average, 0.18, and the standard deviation is 0.22. Table III implies that a standard deviation increase in distance to the median judge corresponds to a 0.9% point increase in the probability of dissenting (the baseline is 3%).

TABLE III

Dissents and Concurrences vs. Distance to Center of Judge Panel (1950-2007)

	(1)
	Dissents or Concurs
Distance to Center of Panel	0.0399***
	(0.00580)
Circuit Fixed Effects	Y
Year Fixed Effects	Y
N	541182
R-sq	0.008

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. Ideology scores come from the Judicial Common Space database. The outcome variable is whether a judge dissents on this panel. Fixed effects include year and circuit.

4.3 Are more extreme judges signing verdicts which are more unfavorable? In Lemma 2 we showed that a necessary condition for the spider pattern to emerge is that the cutoff beyond

which dissent is increasing the extremeness of her ideology. We test this by extending the previous regression.

$$Prob(dissent\, or\, concur) = a + b_1 * abs(t - t_m) + b_2 abs(t - t_m) f(t) + b_3 f(t)$$

We are interested in how the interaction of the extremeness of the judge and the distance to the verdict (i.e., the median judge) affects the probability of dissent. The model predicts the interaction should be negative to capture that the cutoff is falling with extremeness (along with the first term being positive). The results are displayed in Table IV and Figure 7. Table IV reports that for judges who are at any distance to the center of the judge pool, dissents increase with the distance to the center of the panel. But for judges who are quite far from the center of the judge pool, dissents are no longer related to the distance to the center of the panel. These patterns are also reflected in Figure 7 which shows the relationship between dissent and distance to the center of the panel for three groups of judges. Those who are closest to the center of the pool, those who are at moderate distances away from the center, and those who are furthest from the center. Note that judges closest to the center of the judge pool still are, on average, 0.1 distance away in ideology score. The regression coefficients are such that even for the closest judges, a positive relationship between dissent and distance to the panel center should arise.

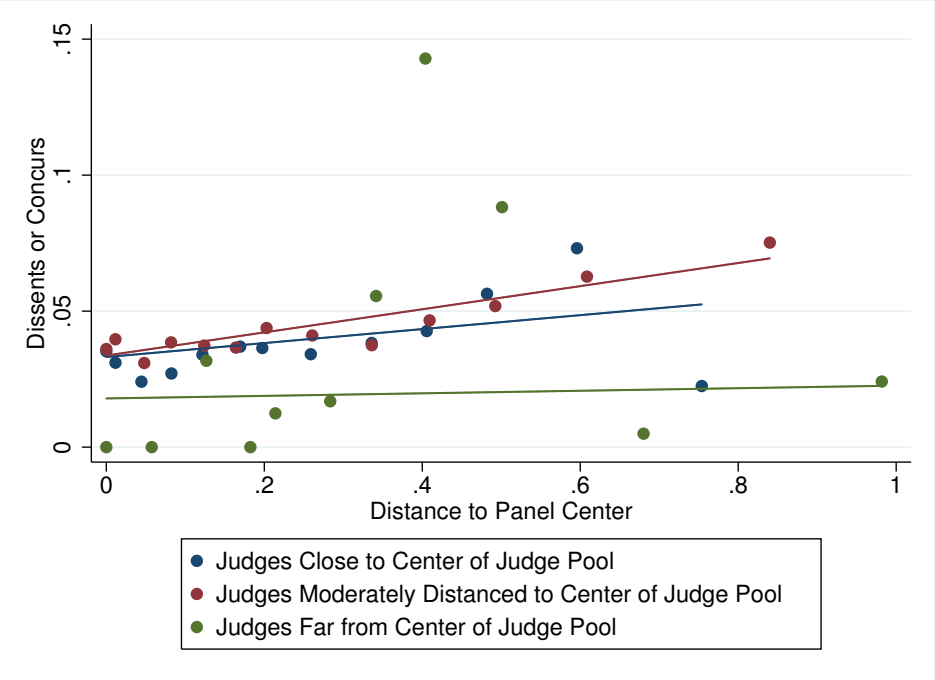
TABLE IV

Dissents and Concurrences vs. Distance to Center of Judge Panel (1950-2007)

	(1)
	Dissents or Concurs
Distance to Center of Panel	-0.00335
	(0.00892)
Distance to Center of Judge Pool	0.0180
	(0.0225)
Distance to Center of Panel *	0.244***
Distance to Center of Judge Pool	(0.0572)
Distance to Center of Judge Pool^2	-0.0403
	(0.0389)
Distance to Center of Panel *	-0.287**
Distance to Center of Judge Pool^2	(0.103)
Circuit Fixed Effects	Y
Year Fixed Effects	Y
N	509022
R-sq	0.008

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. Ideology scores come from the Judicial Common Space database. The outcome variable is whether a judge dissents on this panel. Fixed effects include year and circuit.

FIGURE 7.— Dissents and Concurrences vs. Distance to Center of Judge Panel and Distance to Center of Judge Pool (1950-2007)



Notes: Data on cases comes from OpenJurist. Ideology scores come from the Judicial Common Space database. Ideology scores are demeaned by the the center of the pool of judges available to be assigned in a Circuit-year. Absolute value of the distance to the center of the judge pool is on the x-axis. This plot shows the raw data, no fixed effects are included.

5 Conclusion

This paper has presented a novel pattern in judicial making. Namely, that there is a non-monotonic relationship between a judge's ideological extremeness and her proneness to dissent against verdicts, implying that extreme judges are less likely to dissent than moderate judges. We present a simply model of judicial bargaining over verdicts and a judge's decision to dissent and show that the observed pattern is consistent only if she has perceives a concave cost of agreeing to verdicts she finds dissatifying. A number of other model predictions are corroborated empirically. This shows that a behavioral pattern found in recent lab experiments – that individuals give up on their morals if they cannot follow them fully – has implications for important real world decision making.

References

Ariely, D. (2012): *The (Honest) Truth About Dishonesty : How We Lie to Everyone–Especially Ourselves*, New York: Harper Collins Publishers.

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

Berdejó, C. and D. L. Chen (2013): "Priming Ideology? Electoral Cycles without Electoral Incentives among Elite U.S. Judges," Tech. rep., ETH Zurich, Mimeo.

Chen, D. L., V. Levonyan, and S. Yeh (2011): "Do Policies Affect Preferences? Evidence from Random Variation in Abortion Jurisprudence," Manuscript.

Clark, A. E. and A. J. Oswald (1998): "Comparison-concave utility and following behaviour in social and economic settings," *Journal of Public Economics*, 70, 133–155.

Eguia, J. X. (2011): "On the spatial representation of preference profiles," *Economic Theory*, 52, 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, 101–137.

Epstein, L., A. D. Martin, J. A. Segal, and C. Westerland (2007): "The Judicial Common Space," *Journal of Law, Economics, and Organization*, 23, 303–325.

Gennaioli, N. and A. Shleifer (2007): "The Evolution of Common Law," *The Journal of Political Economy*, 115, 43–68.

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

Gneezy, U., B. Rockenbach, and M. Serra-Garcia (2013): "Measuring Lying Aversion," *Journal of Economic Behavior and Organization*, 93, 293–300.

Hurkens, S. and N. Kartik (2009): "Would I lie to you? On Social Preferences and Lying Aversion," *Experimental Economics*, 12, 180–192.

Kamada, Y. and F. Kojima (2014): "Voter Preferences, Polarization, and Electoral Policies," *American Economic Journal: Microeconomics*, 6, 203–36.

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

Michaeli, M. and D. Spiro (2014): "Skewed Norms under Peer Pressure: Formation and Collapse," Memorandum 15/2014, Department of Economics, University of Oslo.

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

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.

A Proofs

A.1 Proof of Lemma 3 Noting that $P(t) = \left[\left[F\left(t-c\right)\right]^2 + \left[1-F\left(t+c\right)\right]^2\right]$ and denoting

$$z \equiv \int_{t-c}^{t} (t-v)2F(v)f(v)dv + \int_{t}^{t+c} (v-t)2\left[1-F(v)\right]f(v)dv,$$

we can write $L = D\left(z\right) + WP(t)$. We need to show that If $D'' \geq 0$ everywhere, then $P(t)$ is increasing in $|t|$. We provide the proof for $t > 0$ due to symmetry.

(1)

$$\begin{aligned} \frac{dP(t)}{dc} &= \frac{d}{dt}\left[\left[F\left(t-c\right)\right]^2 + \left[1-F\left(t+c\right)\right]^2\right] \\ &= 2F(t-c)f(t-c) + 2\left[1-F(t+c)\right]f(t+c). \end{aligned}$$

Let

$$M \equiv F(t-c)f(t-c) + \left[1-F(t+c)\right]f(t+c) > 0.$$

Then $\frac{dP(t)}{dc} = 2M$ and

$$\frac{dz}{dc} = c2\left\{F(t-c)f(t-c) + \left[1-F(t+c)\right]f(t+c)\right\} = 2cM. \tag{1}$$

We then get

$$\begin{aligned} \frac{dL}{dc} &= D'()2cM - 2WM = 2M\left[cD'() - W\right] \\ L'|_{c=0} &= 0D'(0) - 2MW < 0 \\ L'|_{c\to\infty} &\to \infty D'(\infty) > 0 \end{aligned}$$

So every type has (at least one) inner solution.

(2) In an inner solution we have

$$\frac{dL}{dc} = D'()2cM - 2WM = 2M\left[cD'() - W\right] = 0.$$

As $M > 0$, W is fixed, and

$$\frac{d\left(cD'()\right)}{dc} = D'() + cD''()\frac{dz}{dc} = D'() + 2Mc^2D''() > 0,$$

it follows that there is a unique value of c for which $cD'() = W$ and $\frac{dL}{dc} = 0$, hence the inner solution is unique.

(3) Let $c(t)$ denote the value of c for which type t has an inner solution to the minimization problem. To investigate the properties of $c(t)$, let $G \equiv D'() - \frac{W}{c}$, so that in the inner solution $G = 0$.

$$c' = \frac{dc}{dt} = -\frac{\frac{dG}{dt}}{\frac{dG}{dc}} = -\frac{D''()\frac{dz}{dt}}{D''()\frac{dz}{dc} + \frac{W}{c^2}}.$$

where $\frac{dz}{dc}$ is given by equation (1) and

$$\begin{aligned} \frac{dz}{dt} &= -2cF(t-c)f(t-c) + 2c\left[1-F(t+c)\right]f(t+c) + \int_{t-c}^{t} 2F(v)f(v)dv - \int_{t}^{t+c} 2\left[1-F(v)\right]f(v)dv \\ &= -2cm + Q, \end{aligned}$$

where

$$m \equiv F(t-c)f(t-c) - \left[1-F(t+c)\right]f(t+c) \geq 0$$

and

$$Q \equiv \left[F(t)\right]^2 - \left[F(t-c)\right]^2 + \left[1-F(t+c)\right]^2 - \left[1-F(t)\right]^2 \geq 0.$$

(4) Differentiating now $P(t)$ with respect to t we get

$$
\begin{aligned}
(2) \qquad \frac{1}{2}P'(t) &= F(t-c)f(t-c)\left(1-c'\right) - \left[1-F(t+c)\right]f(t+c)\left(1+c'\right) \\
&= m - c'M \\
&= m + M\frac{D''()\frac{dz}{dt}}{D''()\frac{dz}{dc} + \frac{W}{c^2}} \\
&= m + M\frac{D''()\left(Q-2cm\right)}{2cMD''() + \frac{W}{c^2}} \\
&= m + \frac{D''()\left(Q-2cm\right)}{2cD''() + \frac{W}{c^2M}} \\
&= \frac{2cmD''() + \frac{Wm}{c^2M} + D''()\left(Q-2cm\right)}{2cD''() + \frac{W}{c^2M}} \\
&= \frac{\frac{Wm}{c^2M} + QD''()}{2cD''() + \frac{W}{c^2M}},
\end{aligned}
$$

and so both the numerator and the denominator are positive when $D''() \geq 0$, implying that $P(t)$ increases in the distance from center.

Q.E.D.

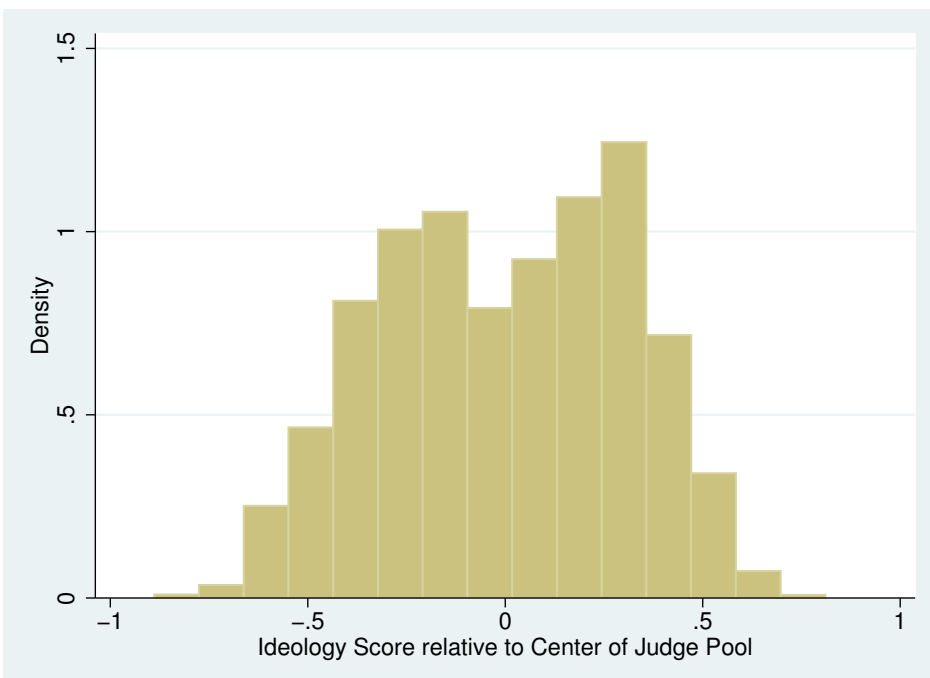
A.2 Proof of Lemma 2 Revisiting equation (2) reveals that if $c'(t) \leq 0$ then $P'(t) = m - c'M > 0$, hence $P(t)$ is increasing in $|t|$.

B Additional Empirical Evidence

This section presents some empirical results supporting the assumptions and intermediate results from the model.

B.1 Histogram of Ideology Scores An assumption in the model is that ideology scores of judges are roughly normally distributed. Figure 8 shows that they are for the judges from each party.

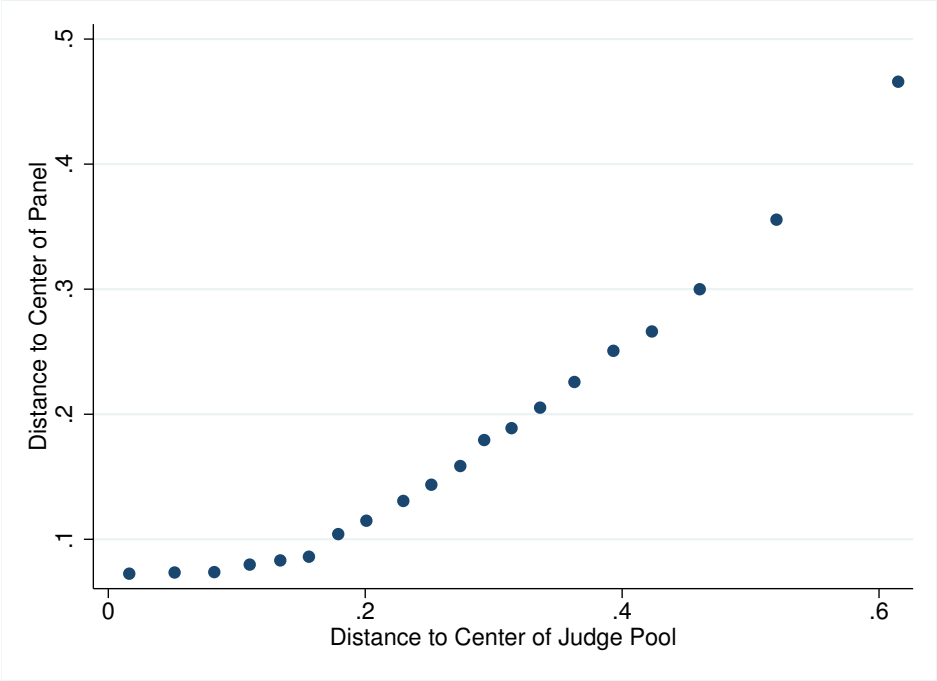
FIGURE 8.— Distribution of Ideology Scores (1950-2007)



Notes: Ideology scores come from the Judicial Common Space database (Epstein et al. 2007), which provides a summary measure using the voting patterns of the appointing president and home state senators.

B.2 Distance to the Panel Center We check if the most extreme judges are indeed, on average, distant from the center of their panels. Figure 9 shows that this is the case. We plot the relationship between the distance to the expected center of the judge pool and the average distance to the center of their panels.

FIGURE 9.— Distance to Panel Center and Distance to Center of Judge Pool (1950-2007)



Notes: Data on cases comes from OpenJurist. Ideology scores come from the Judicial Common Space database. Ideology scores are demeaned by the the center of the pool of judges available to be assigned in a Circuit-year. Absolute value of the distance to the center of the judge pool is on the x-axis.