

Case-by-Case Adjudication and the Path of the Law

ANTHONY NIBLETT*
Draft – October 2010

I. INTRODUCTION

How can a centrist president or governor best influence law through the appointment of judges? Imagine that there are two sitting judges and one of the positions becomes vacant. The other, veteran judge is on the extreme right, from the perspective of the executive, and the executive prefers centrist outcomes. Should the executive appoint a centrist or, instead, appoint a left-wing extremist who might offset the sitting, right-wing judge?

Conventional wisdom holds that judges counteract, or balance, one another; a left-wing appointment carries the best hope offsetting the existing, right-wing judge. As then-Judge Benjamin Cardozo suggested in his 1921 book, *The Nature of the Judicial Process*, “[t]he eccentricities of judges balance one another.”¹ Following this intuition, a moderate appointment would simply cause judicial outcomes to converge to a rule about halfway between the moderate and the right-wing position. There is, however, good reason to think that this conventional wisdom is misguided and that extremists do not necessarily offset one another.

To analyze this problem, an understanding of how legal rules reflect and aggregate the different views of judges is required. Such

* Bigelow Fellow and Lecturer in Law, University of Chicago Law School.

Email: niblett@uchicago.edu

Website: <http://www.law.uchicago.edu/faculty/niblett>

I thank Adam Badawi, Douglas Baird, Scott Baker, Omri Ben-Shahar, Eric Budish, Anthony Casey, Richard Epstein, Lee Fennell, William Hubbard, Louis Kaplow, Ryan Lampe, Saul Levmore, Anup Malani, Ian Martin, Thomas Miles, Eric Posner, Andres Sawicki, Naomi Schoenbaum, Andrei Shleifer, Lior Strahilevitz, and Matthew Tokson, for helpful discussions and comments. I am extremely grateful to Gregory F. Lawler for his help solving the mathematical problem.

¹ Benjamin N. Cardozo, *The Nature of the Judicial Process* (1921), at 177.

legal rules may develop gradually, the culmination of a number of narrow decisions on point. William Landes and Richard Posner contend that the rule created by one decision will be narrow in scope and broader rules usually require a series of decisions.² The decisions in each of the individual cases may be influenced by the preferences of the judges that decided each case.³

Judges often decide only the cases before them. Such behavior may be especially pronounced when judges work with vague standards such as “reasonable,” “excessive,” and “material.” In these contexts, judges will often decide the presented case without venturing to pronounce a bright line rule that would determine the result in all future cases of the same kind. For example, when the Supreme Court took on the task of determining whether punitive damages awards were excessive, it declined to decide the case of first impression with a clear rule, but instead heard several cases over some years and determined from the facts of each case whether the award under review was acceptable or excessive.

It is under this case-by-case approach to judicial decision-making that we would expect legal rules to best take into account different ideologies on the bench.⁴ Unlike broad rules, a narrow decision of one judge has the effect of leaving parts of the law undetermined. This gives judges of future cases greater discretion, affording these judges the opportunity to affect the development of the law. If all judges respect precedent, the gray area of the law will eventually disappear and a clear line in the sand separating compliance from violation will be drawn.

If all judges have the same view, the line of liability will reflect this view. But where will this line in the sand be drawn if judges with heterogeneous preferences resolve disputes? Do the legal rules that emerge from judicial lawmaking reflect the median voter on the bench? How do different distributions of judicial preferences generate different laws? Will the line of liability between compliance and

² William M. Landes and Richard A. Posner (1976), ‘Legal Precedent: A Theoretical and Empirical Analysis,’ 19 *Journal of Law and Economics* 249: “for it is only from a series of decisions, each determining the legal significance of a slightly different set of facts, that a rule applicable to a situation common or general enough to be likely to recur in the future can be inferred.” at 310.

³ The “attitudinal” model of judicial behavior, which claims that judicial decisions are best explained by political preferences has been discussed in academic works too numerous to cite here. I provide some references below, in note 10. An excellent summary of this literature on the attitudinal model can be found in Chapter 2 of Richard A. Posner (2008), *How Judges Think*.

⁴ See e.g., Cass R. Sunstein (2001), *One Case at a Time: Judicial Minimalism on the Supreme Court*.

violation differ under different distributions of judges? Does the law develop differently if the judges issued rules instead of interpreting cases narrowly? These are the questions that this paper seeks to answer.

At least two broad areas of scholarship have explored similar questions. The first branch of scholarship—and the area most closely related to this paper—uses an economic lens to analyze the evolutionary properties of case law. A wave of theoretical literature in the late 1970s and early 1980s sought to explain why the common law may generate more efficient rules than the legislature.⁵ Various explanations were proffered such as judges having preferences for efficiency,⁶ incremental improvements in the informational content of the law,⁷ and litigant behavior.⁸

More recently, researchers have investigated the evolutionary properties of case law when judges have heterogeneous preferences. The assumption of heterogeneous judicial preferences follows the legal realism tradition that suggested personal biases of judges can influence the path of the law.⁹ The assumption also follows the vast empirical

⁵ See generally, Paul Rubin, 'Judge-Made Law,' in *Encyclopedia of Law and Economics*; Lewis A. Kornhauser (1980), 'A Guide to the Perplexed Claims of Efficiency in the Law,' 8 *Hofstra Law Review* 591; Allan C. Hutchinson (2005), *Evolution and the Common Law*; Franciso Parisi and Vincy Fon (2009), *Sources of Law: The Economics of Lawmaking*.

⁶ Richard A. Posner ([1973] 2003), *The Economic Analysis of Law*. See e.g., at 252, where Posner argues that judges "cannot do much. . . to alter the slices of the pie that various groups in society receive, they might as well concentrate on increasing its size." An alternative explanation suggests that precedents alleviate judicial preferences for inefficient outcomes. See, Luca Anderlini, Leonardo Felli, and Alessandro Riboni (2010), 'Why Stare Decisis?' *Mimeo*. The authors argue that judges are affected by a time-inconsistency problem that may lead to excessively lenient decisions. Judges, however, are incentivized toward making tougher, *ex ante* efficient, decisions because they take into account the fact that their decisions will bind courts in future cases.

⁷ See e.g., Robert Cooter, Lewis Kornhauser, and David Lane (1979), 'Liability Rules, Limited Information, and the Role of Precedent,' 10 *Bell Journal of Economics* 366. More recently in this vein, see Scott Baker and Claudio Mezzetti (2010), 'A Theory of Rational Jurisprudence,' *Mimeo*. For an opposing view, suggesting that restricted information and precedent leads to inefficient rules, see Gillian K. Hadfield (1992), 'Bias in the Evolution of Legal Rules,' 80 *Georgetown Law Review* 583.

⁸ See e.g., Paul Rubin (1977), 'Why is the Common Law Efficient?' 6 *Journal of Legal Studies* 51; George Priest (1977), 'The Common Law Process and the Selection of Efficient Rules,' 6 *Journal of Legal Studies* 65; Robert Cooter and Lewis Kornhauser (1980), 'Can Litigation Improve the Law without the Help of Judges?' 9 *Journal of Legal Studies* 139; John C. Goodman (1979), 'An Economic Theory of the Evolution of the Common Law,' 7 *Journal of Legal Studies* 235; R. Peter Terrebonne (1981), 'A Strictly Evolutionary Model of Common Law,' 10 *Journal of Legal Studies* 397.

⁹ See e.g., Oliver Wendell Holmes, Jr. (1897), 'The Path of the Law,' 10 *Harvard Law Review* 457; Jerome Frank (1930), *Law and the Modern Mind*; Joseph C. Hutcheson,

evidence that suggests judicial decisions are influenced by ideology, as well as precedent.¹⁰ Nicola Gennaioli and Andrei Shleifer show that when judges distinguish cases, the legal rules that emerge wash out judicial bias and provide greater precision in the law.¹¹ The authors call their central finding, the “Cardozo Theorem,” in line with the aforementioned proposition that the eccentricities of the judges wash out.¹² Giacomo Ponzetto and Patricio Fernandez present a model where case law is shown to be a continuous and never-ending process of evolution with probabilistic convergence toward efficiency.¹³ Ponzetto and Fernandez also find that the biases of the judges balance one another over time, consistent with the Cardozo Theorem.

Still within this first branch of scholarship, a series of papers have provided an economic analysis of the process of rule-making.¹⁴ The paper that is perhaps most closely related to the model in this paper is

Jr. (1929), ‘The Judgment Intuitive: The Function of the “Hunch” in Judicial Decision,’ 14 *Cornell Law Quarterly* 274. See generally, Brian Leiter (1995), ‘Legal Realism,’ in *A Companion to Philosophy of Law and Legal Theory* (Dennis Patterson, ed.) and references cited within. Following the legal realist tradition, political theorists began to suggest that private attitudes were becoming public law. See e.g., C. Herman Pritchett (1941), ‘Divisions of Opinions Among Justices of the U.S. Supreme Court, 1939-1941,’ 35 *American Political Science Review* 890.

¹⁰ Amongst many others, see, Tracey George and Lee Epstein (1992), ‘On the Nature of Supreme Court Decision-Making,’ 86 *American Political Science Review* 323; Saul Brenner and Harold Spaeth (1995), *Stare Decisis: The Alteration of Precedent on the Supreme Court, 1946-1992*; Donald Songer and Stefanie Lindquist (1996), ‘Not the Whole Story: The Impact of Justices’ Values on Supreme Court Decision Making,’ 40 *American Political Science Review* 1049; Jeffrey Segal and Harold Spaeth (1996), ‘The Influence of Stare Decisis on the Votes of the United States Supreme Court Justices,’ 40 *American Journal of Political Science* 971; Jeffrey Segal and Harold Spaeth (2002), *The Supreme Court and the Attitudinal Model Revisited*; Richard A. Posner (2005), ‘The Supreme Court 2004 Term: A Political Court,’ 119 *Harvard Law Review* 31; Thomas Hansford and James Spriggs, III (2006), *The Politics of Precedent on the U.S. Supreme Court*.

¹¹ Nicola Gennaioli and Andrei Shleifer (2007), ‘The Evolution of Common Law,’ 115 *Journal of Political Economy* 43.

¹² In a separate paper, Gennaioli and Shleifer find that when judges overrule precedent in a similar model, the evolutionary process is unstable, and case law is not conducive to either convergence or efficiency. Nicola Gennaioli and Andrei Shleifer (2007), ‘Overruling and the Instability of the Law,’ 35 *Journal of Comparative Economics* 309.

¹³ Giacomo A. M. Ponzetto and Patricio A. Fernandez (2008), ‘Case Law versus Statute Law: An Evolutionary Comparison,’ 37 *Journal of Legal Studies* 379.

¹⁴ William M. Landes and Richard A. Posner (1979), ‘Adjudication as a Private Good,’ 8 *Journal of Legal Studies* 235. William M. Landes (1971), ‘An Economic Analysis of the Courts,’ 14 *Journal of Law and Economics* 61; Lewis A. Kornhauser (1989), ‘An Economic Perspective on Stare Decisis,’ 65 *Chicago-Kent Law Review* 63.

Lewis Kornhauser's excellent 1992 article exploring how case law exhibits the property of path dependence.¹⁵ Kornhauser illustrates the decentralized nature of the common law in a system where judges do not issue broad rulings, but rather resolve the disputes before them. I adopt many of the assumptions found in Kornhauser's paper, such as dichotomous decisions and respect for the results of precedent.

The second branch of scholarship examines how groups make decisions and whether those decisions reflect the views of the median voter.¹⁶ Much of the law and economics literature in the first branch assumes that each case is heard by just one judge. Appellate judges, however, often decide cases in panels. When judges decide cases in panels, there are additional considerations of aggregating preferences that must be taken into account.¹⁷ For example, Frank Easterbrook showed that Arrow's Impossibility Theorem¹⁸ applies to judicial panels.¹⁹ A series of empirical papers have shown that the ideological

¹⁵ Lewis A. Kornhauser (1992), 'Modeling Collegial Courts, I: Path-Dependence,' 12 *International Review of Law and Economics* 169. On path dependence, see also: Oona A. Hathaway (2001), 'Path Dependence in the Law: The Course and Pattern of Legal Change in a Common Law System,' 86 *Iowa Law Review* 601; and Alec Stone Sweet (2002), 'Path Dependence, Precedent, and Judicial Power,' in Martin Shapiro and Alec Stone Sweet, *On Law, Politics, and Judicialization*, at 112.

¹⁶ On the median voter theorem, see e.g., the seminal work of Duncan Black (1948), 'On the Rationale of Group Decision-Making,' 56 *Journal of Political Economy* 23. See also, Anthony Downs (1957), *An Economic Theory of Democracy*; and Andrew Caplin and Barry Nalebuff (1991), 'Aggregation and Social Choice: A Mean Voter Theorem,' 59 *Econometrica* 1 (providing a multi-dimensional analog to the median voter result).

¹⁷ Posner, cited above in note 3, refers to this literature as fitting within the "sociological theory of judicial behavior," at 31-35. See also, Edward Schwartz (1992), 'Policy, Precedent, and Power: A Positive Theory of Supreme Court Decision-Making,' 8 *Journal of Law, Economics, & Organization* 219; Lewis A. Kornhauser (1992), 'Modeling Collegial Courts, II: Legal Doctrine,' 8 *Journal of Law, Economics & Organization* 441; Jeffrey Lax and Charles Cameron (2007), 'Bargaining and Opinion Assignment on the U.S. Supreme Court,' 23 *Journal of Law, Economics, & Organization* 276; Gregory Caldeira, John Wright, and Christopher Zorn (1999), 'Sophisticated Voting and Gatekeeping in the Supreme Court,' 15 *Journal of Law, Economics, & Organization* 549.

¹⁸ Kenneth J. Arrow (1950), 'A Difficulty in the Concept of Social Welfare,' 58 *Journal of Political Economy* 328.

¹⁹ See Frank H. Easterbrook (1982), 'Ways of Criticizing the Court,' 95 *Harvard Law Review* 802. See also, Frank H. Easterbrook (1984), 'The Supreme Court, 1983 Term, Foreword: The Court and the Economic System,' 98 *Harvard Law Review* 4. See also, Lewis A. Kornhauser and Lawrence Sager (1986), 'Unpacking the Court,' 96 *Yale Law Journal* 82 (arguing that such cycling may not be such a problem, but multi-judge panels may not satisfy community principles of coherency.)

composition of panels can impact the decision.²⁰ One of the major findings is that there is a moderating effect of panels; a Democrat-appointee may vote differently depending on whether the other two judges on the panel are Republican-appointees or Democrat-appointees, for example.²¹ These papers present a challenge to the traditional median voter model of panel behavior.

I present a model illustrating how law evolves when judges with heterogeneous preferences decide cases narrowly. Consistent with the law and economics literature, I assume that the judges hear cases alone and hear cases sequentially. I further assume that judges respect the precedents set by other judges. Previous attempts to model legal evolution in law and economics have largely focused on situations where judges issue threshold “rules” when deciding cases. In these models, the facts of specific cases that come before the court are largely unimportant to the development of law.²² My model departs from the literature on this point. Judges in my model do not set broad rules. Rather, I assume that judges merely resolve the disputes before the court, leaving parts of the law unsettled and affording judges of future disputes the opportunity to influence the direction of the law.

I illustrate incremental change in law using a very simple example. Drivers may be held to be contributorily negligent in railroad crossing accidents, but this depends on how much of the track the driver could see as he approached the crossing. A case with a random fact pattern comes before the court and a judge is randomly-selected to decide the

²⁰ See generally, Matthew C. Stephenson (2009), ‘Statutory Interpretation by Administrative Agencies,’ Draft chapter for *Research Handbook in Public Law and Public Choice* (Dan Farber and Anne Joseph O’Connell, eds.), at 47; Nancy Maveety (2005), ‘The Study of Judicial Behavior and the Discipline of Political Science,’ in *The Pioneers of Judicial Behavior* (Nancy Maveety, ed.), at 1.

²¹ See e.g., Frank B. Cross and Emerson H. Tiller (1998), ‘Judicial Partisanship and Obedience to Legal Doctrine: Whistleblowing on the Federal Courts of Appeals,’ 107 *Yale Law Journal* 2155; Thomas J. Miles and Cass R. Sunstein (2008), ‘The Real World of Arbitrariness Review,’ 75 *University of Chicago Law Review* 761; Thomas J. Miles and Cass R. Sunstein (2008), ‘The New Legal Realism,’ 75 *University of Chicago Law Review* 831; Cass R. Sunstein, David Schkade, Lisa M. Ellman, and Andres Sawicki (2006), *Are Judges Political? An Empirical Analysis of the Federal Judiciary*; Joshua B. Fischman (2008), ‘Decision-Making under a Norm of Consensus: A Structural Analysis of Three-Judge Panels,’ *Mimeo*.

²² In a recent paper, Baker and Mezzetti, cited above in note 7, model a situation where the facts of the case are important to the development of the law. Glenn Ellison and Richard Holden model a situation where a rule is generated when a principal instructs an agent on how best to respond to idiosyncratic situations. An analogy can be drawn between these idiosyncratic situations and the facts of cases. See Glenn Ellison and Richard Holden (2008), ‘A Theory of Rule Development,’ *Mimeo*.

dispute. If the case is not governed by a previously-decided case, the judge has discretion to decide in line with his or her own views. If a case is governed by precedent, it is assumed that the judge respects that precedent. Over time, precedent covers an increasing proportion of fact patterns and the gray area between precedents narrows. Eventually, the law converges on a bright-line legal rule. But where will that bright-line rule be drawn if judges have different views?

The solution is not intuitively or mathematically obvious. Imagine there are two judges, one who is very pro-railroad and one who is very pro-driver, my model suggests that it is unlikely that the legal rule will represent the compromised point of the view. The bright-line rule is unlikely to be found in the middle ground of the judges' respective ideal points. In fact, under my assumptions, the midpoint of the judges' ideal points is the *least* likely outcome. Judges who decide early cases have a disproportionate influence on the direction of the law, since a precedent-based system of adjudication exhibits strong path dependence. The bottom line is judges' biases do not necessarily counteract or balance one another.

My model illustrates the situation of a standard hardening to a rule. This is consistent with the thinking of Oliver Wendell Holmes, H.L.A. Hart, and Richard Posner, among many others, who have suggested that standards evolve into rules over time if the courts respect precedent.²³ More recently, scholars have suggested a less stable evolution of standards and rules.²⁴ In my model, judges adhere to precedent and, over time, the law becomes less uncertain and the standard hardens to a rule. My model goes a step further and describes the content of this legal rule.

It is important to be very clear that my model does not analyze the effects of differing viewpoints when judges decide a case in a panel. The aggregation of preferences in my model does not take place *within* a

²³ See Oliver Wendell Holmes (1881), *The Common Law*; H.L.A. Hart (1961), *The Concept of Law*, at 129; Posner, cited above in note 6, at 539; and Isaac Ehrlich and Richard Posner (1974), 'An Economic Analysis of Legal Rulemaking,' 3 *Journal of Legal Studies* 257. See also, Louis Kaplow (1992), 'Rules versus Standards: An Economic Analysis,' 42 *Duke Law Journal* 557.

²⁴ Jason Johnston and Adrian Vermuele have separately argued that there may be cycling between standards and rules. Standards may harden into rules if judges follow precedent, but may soften once again into a standard if, for example, judges disagree with the inflexibility that rules offer. See Jason Scott Johnston (1991), 'Uncertainty, Chaos, and the Torts Process: An Economic Analysis of Legal Form,' 76 *Cornell Law Review* 341; Adrian Vermeule (2001), 'The Cycles of Statutory Interpretation,' 68 *University of Chicago Law Review* 149. Frederick Schauer has suggested a different form of convergence, arguing that while standards concretize to rules, rules may become more standard-like over time. See Frederick Schauer [2003], 'The Convergence of Rules and Standards,' *New Zealand Law Review* 303.

particular decision or on a particular panel; but rather over many cases over time. My model does not analyze the effect of having a “Brennan” and a “Scalia” on the same panel. The dynamic effects of intra-panel bargaining are, of course, important influences on judicial outcomes.²⁵ My model, however, focuses on a different dynamic effect. I analyze the effect of having a Brennan deciding one case narrowly on the facts, and a Scalia deciding a subsequent case narrowly on the facts.

The model itself is a stand-alone contribution to the emerging literature on judicial decision-making that seeks to describe the content of law. The key finding of the model, showing that the case law does a poor job of finding the middle ground of judges’ preferences, can be applied to other debates in legal scholarship. In this paper, I use the key finding to explore judicial appointments. I challenge the conventional wisdom that “balancing” an appellate court will wash away the biases of judges. This intuition, that there is a need for balancing a court to generate centrist legal rules, may influence the appointment of judges. For example, Democratic senator Charles E. Schumer contended in 2004 that Democrats should be willing to give more leeway to George W. Bush’s conservative nominees to the Ninth Circuit in order to redress a perceived liberal imbalance on the court: “I am ready and willing to support the appointment of conservatives to this court. While it’s gotten more conservative of late, it’s still the most liberal court in the country, it’s still out of balance, and it still needs some evening out.”²⁶

Since case law converges disproportionately to the respective ideal points of the judges in my model, the conventional wisdom that offsetting the existing biases of the sitting judges to generate centrist laws is shown to be misguided. Rather than strategically offset the biases of existing judges, it is better for the appointer to select a judge who shares his preferences.

This result highlights one key difference between the appointment of judges to the bench and the election of politicians to a legislature. There exists a branch of scholarship analyzing voting behavior of citizens in competitive democratic elections. Political scientists have debated whether citizens vote for parties based on the *platform* of the party or the *outcomes*. Proponents of the proximity model contend that voters prefer parties that are ideologically similar to their views;²⁷

²⁵ See references cited above in note 21.

²⁶ Press Release, Charles E. Schumer, Senator for New York, ‘Remarks on the Nomination of William Myers to the 9th Circuit Court of Appeals,’ (April 1, 2004).

²⁷ See e.g., André Blais, Richard Nadeau, Elisabeth Gidengil, and Neil Nevitte (2001), ‘The Formation of Party Preferences: Testing the Proximity and Directional Models,’ 40 *European Journal of Political Research* 81; and Anders Westholm (1997),

while proponents of the directional model argue that voters prefer parties who have views that are more extreme than their own, to balance out the views of opposing parties.²⁸ Orit Kedar, for example, presents a model that shows moderate voters who are primarily concerned with the policy outcomes will prefer extreme parties.²⁹ The moderate executive in my model—the analog of the moderate voter—is also concerned with outcomes not platforms, but still prefers to appoint a moderate judge. This difference is driven by the sequential nature and narrow effect of judicial opinions that generates only probabilistic legal rules, and, unlike rules created by a legislature, these legal rules are highly unlikely to reflect the expected legal rule.

The remainder of the paper is structured as follows. Section II provides motivating examples illustrating how case law evolves when judges decide cases narrowly. Section III describes how the model operates and outlines the solution to the model. Section IV uses the key finding of the model to analyze judicial appointments. A final section concludes. Mathematical solutions to the model are found in the appendix.

II. ILLUSTRATIVE EXAMPLES

A. Grossly excessive punitive damages

In this section I present a stylized example illustrating how case law evolves in my model. The example looks at how the Supreme Court has determined whether punitive damages awards by state courts are excessive, violating the due process clause of the Constitution. A state court's award of punitive damages is deemed unconstitutional if it is too high, and not in proportion with the actual damage or potential damage caused by the defendant.³⁰ The Supreme Court has emphasized

'Distance Versus Direction: The Illusory Defeat of the Proximity Theory of Electoral Choice,' 91 *American Political Science Review* 865.

²⁸ See e.g., Stuart Elaine Macdonald, Ola Listhaug, and George Rabinowitz (1991), 'Issues and Party Support in Multiparty Systems,' 85 *American Political Science Review* 1107; and Stuart Elaine Macdonald, George Rabinowitz, and Ola Listhaug (2001), 'Sophistry versus Science: On Further Efforts to Rehabilitate the Proximity Model,' 63 *Journal of Politics* 482.

²⁹ Orit Kedar (2005), 'When Moderate Voters Prefer Extreme Parties: Policy Balancing in Parliamentary Elections,' 99 *American Political Science Review* 185.

³⁰ In *Browning-Ferris Inds of Vermont, Inc. v. Kelco Disposal, Inc.*, 492 U.S. 257 (1989), the Supreme Court held that punitive damages that were over 100 times the

that it would not provide an explicit threshold rule, stating, “[w]e need not, and indeed we cannot, draw a mathematical bright line between the constitutionally acceptable and the constitutionally unacceptable that would fit every case.”³¹ The court instead has decided to provide a standard, resolving disputes case-by-case. Whether or not an award of punitive damages is “grossly excessive” is informed, in part, by looking at the ratio of punitive damages to the harm caused by the defendant. This ratio is not the only factor that the court looks at. The court also takes into account the aggravating behavior of the defendant and other civil and criminal penalties in similar areas of law and similar jurisdictions.³² By simply isolating the ratio, however, we can learn a great deal about how the law evolves when courts interpret cases narrowly. Before the first case on point, it was unclear to state courts what levels of punitive damages would be deemed excessive. What ratio would be acceptable? What ratio would be deemed unacceptable?

In the 1991 case of *Pacific Mutual Life Insurance Co. v. Haslip*,³³ the majority of the court decided that a punitive damages award of \$840,000 against an insurance company was not excessive when the actual compensatory damages were \$200,000. Justice Blackmun wrote the main opinion stating: “We are aware that the punitive damages award in this case is more than 4 times the amount of compensatory damages, is more than 200 times the out-of-pocket expenses of [the plaintiff], and, of course, is much in excess of the fine that could be imposed for insurance fraud [under Alabama legislation.]”³⁴ He concluded: “While the monetary comparisons are wide and indeed, may be close to the line, the award here did not lack objective criteria. We conclude, after careful consideration, that in this case it does not cross the line into the area of impropriety.”³⁵

In the next case in the sequence, *TXO Production Corp. v. Alliance Resource Corp.*,³⁶ the majority found the ratio of punitive damage to the potential damage that could have been caused by the defendant’s behavior may have been less than two-to-one, but could be as high as

compensatory damages did not violate the Eighth Amendment. The appellant did not raise the due process argument before the District Court or the Court of Appeals.

³¹ See e.g., *Pacific Mut. Life Ins. Co. v. Haslip*, 499 U.S. 1 (1991) at 18.

³² *B.M.W., Inc. v. Gore*, 517 U.S. 559 (1996).

³³ *Pacific Mut. Life Ins. Co. v. Haslip*, 499 U.S. 1 (1991).

³⁴ 499 U.S. at 23.

³⁵ 499 U.S. at 23-24.

³⁶ *TXO Prod. Corp. v. Alliance Resource Corp.*, 509 U.S. 443 (1993).

ten-to-one.³⁷ Justice Stevens, writing the lead opinion, wrote that *if* the ratio were as high as ten-to-one, the punitive damages award in this case would not “jar one’s constitutional sensibilities.”³⁸

In 1996, the Supreme Court decided *B.M.W., Inc. v. Gore*.³⁹ In that case, a car dealership was found to have caused \$4,000 worth of damage to a buyer by repainting a car that was damaged pre-delivery and selling the car as new. The jury awarded \$4 million in punitive damages which was later reduced by the state appellate court in Alabama to \$2 million. The Supreme Court distinguished this case from the previous two, holding that the ratio of 500-to-one was grossly excessive. Justice Stevens noted: “When the ratio is a breathtaking 500 to 1. . . the award must surely ‘raise a suspicious judicial eyebrow.’”⁴⁰ The majority held that the defendant did not behave in a manner that warranted such high damages. The *Gore* decision said nothing about the gray area of ratios between ten-to-one (not excessive) and 500-to-one (excessive).

In the 2003 decision of *State Farm Mutual Automobile Ins. Co. v. Campbell*,⁴¹ a punitive damages award of \$145 million was deemed to be grossly excessive when compared to the compensatory damages of \$1 million. The court once again refused to impose a bright line ratio.⁴² Justice Kennedy however noted that few awards that exceed single-digit ratios would probably be acceptable.⁴³ He stated: “Single-digit multipliers are more likely to comport with due process, while achieving the State’s goals of deterrence and retribution, than awards

³⁷ The compensatory damages in this case were only \$ 19,000 while the punitive damages award was \$ 10 million. “While petitioner stresses the shocking disparity between the punitive award and the compensatory award, that shock dissipates when one considers the potential loss to respondents, in terms of reduced or eliminated royalties payments, had petitioner succeeded in its illicit scheme. Thus, even if the actual value of the ‘potential harm’ is not between \$ 5 million and \$ 8.3 million, but is closer to \$ 4 million, or \$ 2 million, or even \$ 1 million, the disparity between the punitive award and the potential harm does not, in our view, ‘jar one’s constitutional sensibilities.” See 509 U.S. at 462 (per Justice Stevens, citing *Haslip*, 499 U.S. at 18.)

³⁸ *Ibid* (my emphasis). Justice Stevens emphasized the reprehensible and malicious behavior of the defendant in this case, suggesting that not all punitive damages awards that are ten times the potential damage caused by the defendant would be deemed constitutionally acceptance.

³⁹ *B.M.W., Inc. v. Gore*, 517 U.S. 559 (1996).

⁴⁰ 517 U.S. at 583 citing Justice O’Connor’s dissent in *TXO*, 509 U.S. at 481.

⁴¹ *State Farm Mutual Automobile Ins. Co. v. Campbell*, 538 U.S. 408 (2003).

⁴² See, e.g., 538 U.S. at 425: “We decline again to impose a bright-line ratio which a punitive damages award cannot exceed” (per Justice Kennedy).

⁴³ *Ibid*.

in the range of 500 to 1, or, in this case, of 145 to 1.”⁴⁴ The ratio of 145-to-one was held to be disproportionate, excessive, and therefore, unconstitutional.⁴⁵

The ratios in the four cases are depicted in **Figure 1**.⁴⁶ The first case holds an award of four times the damage to be constitutional (*Haslip*, 1991). The second case holds that a ratio of ten times would also not violate the fourteenth amendment (*TXO*, 1993). The third case (*Gore*, 1996) shows that a ratio of 500-to-one is grossly excessive, as is ratio of 145-to-one, from the fourth case (*Campbell*, 2003). The gray area has narrowed over time. My simplification of this jurisprudence in this area illustrates that as more cases are decided (and decided narrowly), the “gray area” of the law disappears and we begin to converge toward a rule.

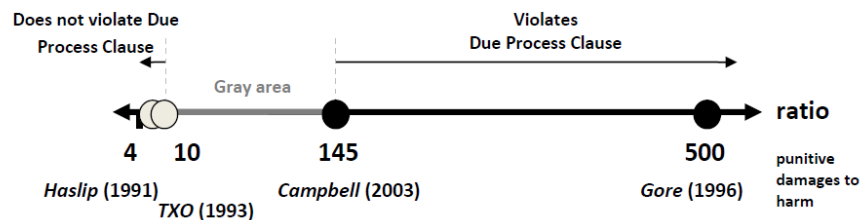


Figure 1: The four cases are shown on a spectrum that indicates the ratio of punitive damages to harm. Ratios lower than 10:1 have been found to be constitutionally acceptable. Ratios higher than 145:1 have been found to be unacceptable. In between ratios of 10:1 and 145:1 is a gray area.

⁴⁴ *Ibid* (references to *Gore* have been removed).

⁴⁵ Other factors such as the wealth of the defendant and the fact that the defendants had been similarly punished elsewhere (Texas) were not held to be relevant. See 538 U.S. at 426-428.

⁴⁶ *Philip Morris v. Williams*, 549 U.S. 346 (2007) and *Exxon Shipping v. Baker*, 128 S. Ct. 2605 (2008) are not included in this analysis. In *Philip Morris*, the compensatory damages award from the Oregon court was around \$ 821,000 and the punitive damages award was about \$ 79.5 million. The Supreme Court did not decide whether the ratio of 97:1 was grossly excessive or not. The decision of the majority instead focused on whether the punitive damages award was improperly calculated because it took into account harm to other smokers who were not party to the case. The *Exxon* decision is not included because it focuses on admiralty law, rather than state-court punitive damages awards. In this case, the Supreme Court ruled that a punitive damages award of \$ 2.5 billion should be reduced to around \$ 500 million, when the compensatory damages were around \$ 507 million. The court held that under federal maritime jurisdiction the fair upper limit was a compensatory-to-punitive damages ratio of 1:1. This admiralty case does not affect state-court punitive damages awards.

B. Other examples

The example above is merely used to illustrate the incremental nature of legal evolution. The example uses Supreme Court cases; however, my model is not *per se* a model of Supreme Court judicial behavior. It does not analyze how a panel of (the same) nine justices decide cases. My model rather illustrates how the law evolves when the bench consists of *different* judges with heterogeneous preferences who resolve disputes individually and sequentially. In the context of the Supreme Court, the ideology of the median judge may change over time,⁴⁷ but the model is a better fit for all other courts where judges are selected randomly to hear cases. There are myriad other examples that one could point to involving standards that evolve toward rules in this manner. I provide a very brief overview of such four examples.

i. Non-compete covenants in employment contracts

Most state courts decide whether a non-compete covenant in an employment contract is reasonable or not in a case-by case manner.⁴⁸ Covenants against competition are tolerated only if they are strictly limited in time and territorial effect, and these limits are reasonable.⁴⁹ The Supreme Court of Georgia provides a nice example of how courts resolve these disputes narrowly and adhere to precedent.⁵⁰ In 1968, the court decided that a restrictive covenant that lasted for just one year and restricted the former employee from working within a 10

⁴⁷ See e.g., Andrew D. Martin, Kevin M. Quinn, and Lee Epstein (2005), 'The Median Justice on the United States Supreme Court,' 83 *North Carolina Law Review* 1275; Keith Krehbiel (2007), 'Supreme Court Appointments as a Move-the-Median Game,' 51 *American Journal of Political Science* 231.

⁴⁸ Not all state courts analyze non-compete covenants this way. Indeed, in some states, such as California, the use of non-compete covenants in employment contracts is prohibited. Forty-four states have not prohibited such clauses in employment contracts, however.

⁴⁹ See e.g., *Shirk v. Loftis Bros. & Co.*, 148 Ga. 500 (1918).

⁵⁰ In this example, I include only cases where non-compete clauses (not solicitation or non-disclosure clauses) were included only in employment contracts. Non-compete clauses are also found in partnership, sales of business, and franchise contracts; but what is reasonable in these cases may differ to what is reasonable in employment contracts. See e.g., *Rash v. Toccoa Clinic Medical Assoc.*, 253 Ga. 322 (1984) (partnership); *Dalrymple v. Hagood*, 246 Ga. 235 (1980) (sale of business); *Johnson v. Lee*, 243 Ga. 864 (1979) (sale of business); *T. E. McCutcheon Enterprises, Inc. v. Snelling & Snelling, Inc.*, 232 Ga. 609 (franchise); and *Farmer v. Airco, Inc.*, 231 Ga. 847 (1973) (sale of business).

mile radius of the former employer was “reasonable both as to time and territory.”⁵¹ Two years later, the court decided that a covenant of two years covering an area of 31 states was an unreasonable restriction, because it was “oppressive to the party restrained and opposed to the interests of the public.”⁵² The 1970s and 1980s saw a series of decisions that filled the gap between these two cases.⁵³ The gap between covenants that were deemed unreasonable and those that have been deemed reasonable narrowed, and the gray area diminished. When cases fell within the ambit of a precedent, the court issued decisions that were consistent with precedent.⁵⁴

ii. Ministerial exception to Title VII

Convergence and line drawing in the law, obviously, does not always involve a spectrum with numbers. Courts can still rank and compare fact patterns when these facts are not easily described by numbers. As an example, take the “ministerial exception” to Title VII. Federal courts have held that they do not have jurisdiction to apply Title VII to employment relationships involving a church and a minister because the state would be encroaching upon the forbidden area of

⁵¹ *Baxley v. Black*, 224 Ga. 456 (1968) at 457.

⁵² *Moore v. Dwoskin, Inc.*, 226 Ga. 835 (1970) at 837 (citing *Rakeshaw v. Lanier*, 104 Ga. 188 (1898) at 202.)

⁵³ A covenant that covered a radius of approximately twelve miles for eighteen months was deemed reasonable. See *Mike Bajalia, Inc. v. Pike*, 226 Ga. 131 (1970). The territorial limit here was Lowndes Co., Georgia. This county is 511 sq. miles; approximately twelve miles radius. Covenants that covered a radius of 25 miles but lasted only a year were also held to be reasonable. See e.g., *Preferred Risk Mutual Ins. Co. v. Jones*, 233 Ga. 423 (1974); and *Landmark Financial Serv., Inc. v. Tarpley*, 236 Ga. 568 (1976). In 1976, “reasonable” was extended to include a covenant that covered a radius of 50 miles and lasted for one year. *Edwards v. Howe Richardson Scale Co.*, 237 Ga. 818 (1976). Two years earlier, however, a covenant that covered an area of 50 miles and lasted one year was not enforced by the court. See *Worley & Assoc., Inc. v. Bull*, 233 Ga. 276 (1974). In that case, the defendant former employee was working in a county where the plaintiff former employer had no customers. The court held that a non-compete clause was overly broad if it prohibits employment in areas where the employer does not do business. The injunctive relief sought by the former employer was not granted.

⁵⁴ For example, a covenant that covered a radius of 35 miles but lasted only six months was deemed reasonable and not unduly broad. See *Beckman v. Cox Broadcasting Corp.*, 250 Ga. 127 (1982). Similarly, covenants that covered radii of five miles and ten miles for two years were deemed reasonable. See, respectively, *Watson v. Waffle House, Inc.*, 253 Ga. 671 (1985) and *Nunn v. Orkin Exterminating Co., Inc.*, 256 Ga. 558 (1986).

religious freedom.⁵⁵ The landmark decision *McClure v. The Salvation Army*⁵⁶ found that an exception applies to ordained ministers. This was later extended to include associates in pastoral care⁵⁷ and probationary ministers.⁵⁸

A series of cases followed, with litigants testing where the courts draw the line. A typist-receptionist at a church was held not to fall within the exception;⁵⁹ neither did an editorial secretary at a religious printing press⁶⁰ or a director of plant operations at a religious hospital.⁶¹ On the other hand, a communications manager who acted a liaison between the church and parts of the congregation was held to fall under the ministerial exception.⁶² Generally, teachers at religious institutions have not been found to fall within the ministerial exception;⁶³ but if the teacher's primary duty is spreading faith, religious participation, then the exception applies.⁶⁴ Over time, the gray area between what is covered and what is not covered by the exception has narrowed.

iii. MAC clauses in merger agreements

⁵⁵ See e.g., *McClure v. The Salvation Army*, 460 F.2d 553 (5th Cir. 1972), *Bryce v. Episcopal Church in the Diocese of Colo.*, 289 F.3d 648 (10th Cir. 2002), *Minker v. Baltimore Annual Conference of the United Methodist Church*, 894 F.2d 1354 (D.C. Cir. 1990).

⁵⁶ *McClure v. The Salvation Army*, 460 F.2d 553 (5th Cir. 1972).

⁵⁷ *Rayburn v. Gen. Conf. of Seventh-Day Adventists*, 772 F.2d 1164 (4th Cir. 1985).

⁵⁸ *Young v. The Northern Illinois Conference of United Methodist Church*, 21 F.3d 184 (7th Cir. 1994).

⁵⁹ *Whitney v. Greater N.Y. Corp. of Seventh-Day Adventists*, 401 F.Supp. 1363 (S.D.N.Y., 1975)

⁶⁰ *EEOC v. Pacific Press Pub. Ass'n*, 676 F.2d 1272 (9th Cir. 1982).

⁶¹ *Lukaszewski v. Nazareth Hospital*, 764 F.Supp. 57 (E.D.Pa., 1991).

⁶² *Alicea-Hernandez v. Catholic Bishop of Chicago*, 320 F.3d 698, 703 (7th Cir. 2003).

⁶³ See *EEOC v. Fremont Christian School*, 781 F.2d 1362 (9th Cir. 1986), *Geary v. Visitation of the Blessed Virgin Mary Parish School*, 7 F.3d 324 (3d Cir. 1993), *EEOC v. Mississippi College*, 626 F.2d 477 (5th Cir. 1980).

⁶⁴ *EEOC v. Catholic Univ.*, 83 F.3d 455 (D.C. Cir. 1996) and *EEOC v. Southwestern Baptist Theological Seminary*, 651 F.2d 277 (5th Cir. 1981). Similarly, the exception applies to music directors at churches and religious schools if spreading religious faith. See e.g., *Starkman v. Evans*, 198 F.3d 173 (5th Cir. 1999) and *EEOC v. Roman Catholic Diocese*, 213 F.3d 795 (4th Cir. 2000).

The line of liability may be drawn at one end of the spectrum very early in the evolution of a law, even if the disputes are decided narrowly. Take, for example, the way the Delaware Chancery Court has interpreted “material adverse change” clauses (“MAC” clauses) in merger contracts. A MAC clause in a corporate merger agreement allows the acquirer to walk away from the merger if the target suffers a material adverse change (or effect) between the time of signing and the time of closing; however, the Court of Chancery has never recognized a change or event as being materially adverse. In deciding cases, Delaware courts have placed “strong emphasis on the particular facts.”⁶⁵

The first case, *In re IBP, Inc. S’holders Litig.*,⁶⁶ represents an example of an early case in the evolution of law being decided narrowly, and implicitly resolved most potential disputes. Tyson Foods discontinued its purchase of IBP after signing the merger agreement, claiming that an unexpected forty per cent drop in the target’s sales constituted a material adverse change, and gave rise to a right to terminate.⁶⁷ Vice Chancellor Strine found the case to be a close one,⁶⁸ but decided that the drop in earnings was not a material adverse change.⁶⁹ V.C. Strine wrote that this finding reflected his own personal perspective that a “reasonable acquirer” should be viewing takeovers from a “longer-term perspective”;⁷⁰ explicitly noting that his own view was “seller friendly.”⁷¹

⁶⁵ R. Samuel Snider, Rahul Patel, and Will Smoak (2009), ‘Accounting for the Unknowable: Risk Allocation & Current Insights on Material Adverse Change Clauses,’ 2009 *Emerging Issues* 3600 (Lexis).

⁶⁶ *In re IBP, Inc. S’holders Litig.*, 789 A.2d 14 (2001).

⁶⁷ Poor weather conditions led to an unexpected “sharp drop” in earnings from \$2.38 per share to \$1.44 a share. 789 A.2d at 69. IBP could only point to “two weeks of truly healthy results in 2001 before the contract termination date.” 789 A.2d at 71. The discovery that an IBP subsidiary had been a victim of fraud was also claimed. The fraud was expected to have a small effect upon the entire company however. The subsidiary was found to have “a tiny fraction of IBP’s overall business and that a total shutdown of [the subsidiary] would likely have little effect on the future results of a combined Tyson/IBP.” See 789 A.2d at 70.

⁶⁸ V.C. Strine “admit[ted] to reaching this conclusion with less than the optimal amount of confidence.” 789 A.2d at 71. Further, V.C. Strine noted: “I am confessedly torn about the correct outcome.”

⁶⁹ Although the target did not perform as well as Tyson had hoped, V.C. Strine found that IBP was still “in sound enough shape to deliver results of operations in line with the company’s recent historical performance.” *Ibid.*

⁷⁰ 789 A.2d at 68.

⁷¹ The Vice Chancellor explains that his conclusion is heavily influenced by his own temporal perspective that may not be shared by other judges. See 789 a.2d at 71,

The *IBP* decision was seen by commentators as surprising.⁷² Commentators suggested that the interpretation of the materiality standard is “so demanding that—absent a cataclysm of biblical proportions—it cannot be met.”⁷³ The benchmark was set “impossibly high,” ensuring that the MAC clause will almost certainly not be invoked.⁷⁴ The line was drawn heavily in favor of sellers. It became very difficult for buyers to walk away under the auspices of a material change. This example illustrates how the decisions of early cases can lead to the law settling at one end of the spectrum relatively early in the evolution of the law. A standard that settles quickly is less likely to be litigated. Indeed, it took another seven years before the Delaware Chancery Court heard another case where a buyer sought to enforce a MAC clause.⁷⁵

footnote 170. He also wrote: “Tyson has evinced more confidence in stock market analysts than I personally harbor.” 789 A.2d at 71.

⁷² See e.g., Andrew A. Schwartz (2010), ‘A “Standard Clause Analysis” of the Frustration Doctrine and the Material Adverse Change Clause’, 57 *U.C.L.A. Law Review* (forthcoming). There were precedents from New York suggesting a materially adverse change can be found if a target suffers a short-term decline in profitability. See e.g., *In Pan Am Corp. v. Delta Airlines*, 175 B.R. 438 (S.D.N.Y. 1994) (Pan Am Airlines suffered a sharp decline in bookings over a three-month period), and *Katz v. NVF Co.*, 100 A.D.2d 470 (N.Y. App. Div. 1984) (two merger partners agreed that one partner suffered a materially adverse change after making a loss over the course of one year of \$6.3 million, compared to a profit the year before). There was, however, a more recent case, *Bear Stearns Co. v. Jardine Strategic Holdings*, No. 31371187, slip. Op. (N.Y. Supr. June 17, 1988), where Bear Stearns’s losses on Black Monday, October 19, 1987 (up to \$100 million) did not represent a materially adverse change, and the contract to purchase 20% of Bear Stearns could not be avoided.

⁷³ See Schwartz, cited above in note 72. See also, Jeffrey Thomas Cicarella (2007), Note, ‘Wake of Death: How the Current MAC Standard Circumvents the Purpose of the MAC Clause’, 57 *Case Western Reserve Law Review* 423, at 450 (MAC case law uses “a test for materiality that can almost never be met”).

⁷⁴ See Schwartz, cited above in note 72, at 46.

⁷⁵ The facts of *Hexion Speciality Chems, Inc. v. Huntsman Corp.*, 965 A.2d 715 (2008) suggest that the case falls within the ambit of the *IBP* precedent. The buyer, Hexion, contended that the target, Huntsman, had suffered a materially adverse change in conditions. In the *Hexion* case, Vice Chancellor Lamb noted the seller friendly philosophy of Vice Chancellor Strine in the *IBP* precedent, contending that a “buyer faces a heavy burden when it attempts to invoke a material adverse effect (‘MAE’) clause in order to avoid its obligation to close.” See 965 A.2d at 738. Hexion suggested that there had been an MAE because of Huntsman’s disappointing earnings performance in the year following the agreement. There was some dispute over which metric to use to measure Huntsman’s performance. Hexion claimed that projections for earnings had fallen somewhere between 19 and 32 per cent from the year before. See 965 A.2d at 740. The judge found that a more accurate metric indicated a drop in earnings of around four per cent. See 965 A.2d at 743. Vice Chancellor Lamb held that such results “do not add up to an MAE”. See 965 A.2d at 743.

iv. Contributory negligence of drivers

Finally, judges may not decide cases narrowly all the time, but any dicta handed down by judges in early cases may be nullified in subsequent cases. In this way, judges are constrained in their attempts to set rules. Take, for example, the cases determining the contributory negligence of plaintiff automobile drivers injured in railroad crossing accidents from the late 1920s and early 1930s. Railroad companies could be held negligent if drivers were injured or killed on railroad crossings without bells and whistles to indicate an on-coming train. Drivers were sometimes held to be contributorily negligent. Courts took into account factors such as how much of the track the driver could see when he approached the danger zone. In the 1927 case, *Baltimore & Ohio Railroad Co. v. Goodman*,⁷⁶ the Supreme Court found that the driver was contributorily negligent in spite of the fact that he could only see 243 feet of track. Justice Holmes, writing for the unanimous court, added dicta stipulating that the driver always has a responsibility to stop, look, and listen—and even get out of his car—if he is not certain whether a train is coming.

The Supreme Court revisited the issue seven years later in *Pokora v. Wabash Rwy Co.*⁷⁷ Contrary to the dicta in *Goodman*—but not inconsistent with the result—the court found that a driver who could see only 130 feet of the track was *not* contributorily negligent. The court held that the plaintiff could not have been protected by having such a short glimpse of the track. Justice Cardozo, writing for the unanimous court,⁷⁸ wrote that there “is no doubt that the opinion in [*Goodman*] is correct in its result,”⁷⁹ but noted that he was not bound by Justice Holmes’s dicta.⁸⁰ I use this final example in section III to flesh out the model.

III. THE MODEL

A. The set-up

⁷⁶ *Baltimore & Ohio Railroad Co. v. Goodman*, 275 U.S. 66 (1927).

⁷⁷ *Pokora v. Wabash Rwy Co.*, 292 U.S. 98 (1934).

⁷⁸ Justices Van Devanter, McReynolds, Brandeis, Sutherland, Butler, and Stone were on both the *Goodman* and *Pokora* panels.

⁷⁹ 292 U.S. at 102.

⁸⁰ 292 U.S. at 104.

In this section I provide a sketch of my model. The model illustrates how case law evolves when judges resolve disputes narrowly. I do not address *why* judges resolve cases narrowly; it is rather assumed that they do.⁸¹ I show how a rule emerges when judges sequentially resolve disputes and interpret vague standards such as “reasonable,” “excessive,” and “material.” It is *ex ante* unclear what the dividing line will be between compliance and violation. The “rule,” however, emerges from judges resolving disputes and acting consistently with precedents.

I use a very simple example. Suppose there is a jurisdiction with just one level of court.⁸² When a case comes before a court, a single judge—chosen at random—decides the case. The court has been asked to decide whether an automobile driver was contributorily negligent when injured or killed in a railroad crossing accident. I assume that the only key fact that all judges take into account is the field of vision the driver had of the railroad track as he approached the crossing.⁸³ If the driver could not see any length of the track, he will not be held to be contributorily negligent. If the driver could see a very long length of the track, he will be held to be contributorily negligent. In line with Kornhauser’s 1992 model, the decisions of judges are dichotomous.⁸⁴ A

⁸¹ An alternative explanation for this assumption is that if judges do add dicta, the dicta can be subsequently ignored, rendering all decisions narrow in effect—as in the *Goodman* and *Pokora* examples above in section II.B.iv. Two recent law and economics papers have explored the issue of breadth of judicial decisions. In these two papers, the breadth of the decision is chosen by the judge. See, Baker and Mezzetti, cited above at n 7; and Anderlini, Felli, and Riboni, cited above at n 6.

⁸² In assuming only one level of court, my model abstracts from the important effects of hierarchy, control, and aspects of the appeals process. See e.g., Richard A. Posner (1990), *The Problems of Jurisprudence*, at 224; Donald R. Songer, Jeffrey A. Segal, and Charles M. Cameron (1994), ‘The Hierarchy of Justice: Testing a Principal-Agent Model of Supreme Court-Circuit Court Interactions,’ 38 *American Journal of Political Science* 673; Lewis A. Kornhauser (1995), ‘Adjudication by a Resource-Constrained Team: Hierarchy and Precedent in a Judicial System,’ 68 *Southern California Law Review* 1605; McNollgast (1995), ‘Politics and the Courts: A Positive Theory of Judicial Doctrine and the Rule of Law,’ 68 *Southern California Law Review* 1631; Steven Shavell (1995), ‘The Appeals Process as a Means of Error Correction,’ 35 *Journal of Legal Studies* 1; Ethan Bueno de Mequita and Matthew C. Stephenson (2002), ‘Informative Precedent and Intrajudicial Communication,’ 96 *American Political Science Review* 755; David E. Klein (2002), *Making Law in the United States Court of Appeals*; Tonja Jacobi and Emerson H. Tiller (2007), ‘Legal Doctrine and Political Control,’ 23 *Journal of Law, Economics, & Organization* 326; Pauline T. Kim (2007), ‘Lower Court Discretion,’ 83 *New York University Law Review* 383.

⁸³ See *Pokora*, 292 U.S. at 100, where Justice Cardozo makes reference to the “zone of danger.”

judge simply decides if the driver was contributorily negligent (i.e., had a sufficient field of vision and should have stopped) or not. Different judges, however, have different views about where the line should be drawn.

i. Facts

There exists a spectrum of facts along one dimension.⁸⁵ The facts of any case are described by a point on this line. The point is essentially a “bundle” of different key facts that will be relevant to liability. Bundles of facts are ranked in order of egregiousness. These facts are assumed to be objective and verifiable.⁸⁶ In my example, I make a very simple assumption about the nature of “facts.” I assume that prior to any judicial intervention the length of track that the drivers can see when they cross a railroad track is uniformly distributed between 0 feet and 400 feet. That is, there are some crossings where the driver is afforded no visibility, there are some crossings where the driver is afforded a great deal of visibility, and there are many different intermediate types of crossing. The assumption of a uniform distribution merely suggests that there is some variation in the facts of cases, but the facts are not necessarily skewed toward the extremes, nor skewed toward the middle. The facts of cases that come before the court are drawn at random. The assumption of randomness abstracts away problems of selection effects.⁸⁷

⁸⁴ Kornhauser, cited above in note 15, at 171. Kornhauser explains that this assumption is representative of legal decisions. “Legal questions almost always have yes or no answers. While the plaintiff seeks a remedy from a potentially large number of possible outcomes, the question of law generally is simply whether she has a right to recover or not.”

⁸⁵ A model that places factual scenarios on just one dimension may appear to have limited application. I discuss the effects and sources of multiple dimensions below in section III.C.ii.

⁸⁶ This abstracts away from any possibility of judicial fact discretion. In real cases, judges may have the ability to selectively choose which facts to emphasize in a decision or perhaps the ability to nudge the bundle of facts to give the impression that the facts of the case are more or less egregious than they are. This is a feature of legal realist thinking. See e.g., Jerome N. Frank (1931), ‘Are Judges Human? Parts I and II,’ 80 *University of Pennsylvania Law Review* 17 and 233; Jerome N. Frank (1949), *Courts on Trial: Myth and Reality in American Justice*, esp. Chapter III; Nicola Gennaioli and Andrei Shleifer (2008), ‘Judicial Fact Discretion,’ 37 *Journal of Legal Studies* 1 (exploring the economic effects of such discretion).

⁸⁷ The assumptions of the model can be tweaked to permit situations where facts are not drawn randomly. For example, if litigants only bring, or judges only allow, extreme cases early on in the development of the law, the law will converge much faster. There is a substantial literature in law and economics discussing the sources and

ii. Judges and ideal points

Different judges have different views of where the line of liability should be drawn. To align the model with the empirical literature in this area, I shall refer to each individual's threshold line as their "ideal point."⁸⁸ Suppose that there are just two judges with different views. One judge, Lewis, is very pro-railroad company. He has an ideal point of 0 feet. He believes that any driver who gets hit by a train is at fault. If the driver is unsure whether a train is coming as he drives toward a railroad crossing, he should get out of his car and look down the track. The other judge, Rita, is very pro-automobile driver. Her ideal point lies at the other end of the factual spectrum, at 400 feet. She believes that any driver who gets hit by a train should recover, unless the driver could see further than 400 feet when they crossed the track. The midpoint of the two judges' ideal points is 200 feet. After a random case comes to court, a judge to hear the case is randomly selected.⁸⁹ There is a fifty per cent chance of drawing Lewis and a fifty per cent chance of drawing Rita for each case. I have chosen these two "extreme" ideal points for the purposes of illustration only. I could have, for example, assumed that Lewis has an ideal point of 100 feet and Rita has an ideal

effects of case selection by litigants, including Marilyn J. Simon (1981), 'Imperfect Information, Costly Litigation, and Product Quality,' 12 *Bell Journal of Economics* 171; Steven Shavell (1982), 'The Social versus the Private Incentive to Bring Suit in a Costly Legal System,' 11 *Journal of Legal Studies* 333; George L. Priest and Benjamin Klein (1984), 'The Selection of Disputes for Litigation,' 13 *Journal of Legal Studies* 1.

The selection of cases for trial is also affected by a court's discretion to choose the cases that come before it. In my model, if case selection is not random the distribution of facts will not be uniform. See e.g., Vanessa Baird (2007), *Answering the Call of the Court: How Justices and Litigants Set the Supreme Court Agenda*; Tonja Jacobi (2008), 'The Judicial Signaling Game: How Judges Shape Their Dockets,' 16 *Supreme Court Economic Review* 1; Jonathan P. Kastellec and Jeffrey R. Lax (2008), 'Case Selection and the Study of Judicial Politics,' 5 *Journal of Empirical Legal Studies* 407; Steven Shavell (2010), 'On the Design of the Appeals Process: The Optimal Use of Discretionary Review versus Direct Appeal,' 39 *Journal of Legal Studies* 63.

⁸⁸ I assume that judges receive utility from the outcomes of particular cases. In a world unconstrained by precedent, each judge has preferences for the desired outcome in all factual situations, $f \in [0, 400]$. We will call the dichotomous outcomes, L and NL , for liability and non-liability respectively. The desired outcome in a case with facts f is $r^*(f)$, the outcome that generates greater utility. The ideal point for judge i is the fact pattern, f_i^* , where the judge is indifferent between liability and non-liability: $U_i(L|f_i^*) = U_i(NL|f_i^*)$.

⁸⁹ For a defense of the randomization of decision-makers in the adjudicative system, see, Adam M. Samaha (2009), 'Randomization in Adjudication,' 51 *William & Mary Law Review* 1.

point of 300 feet. Below, I point out how such a change in assumptions would affect the results.

The assumption of just two judges is clearly not entirely realistic; however, one can think of two equally-sized blocs of judges. Judges within each bloc share similar preferences and ideal points of the law. The assumption of two blocs of ideologically opposed judges is a simplification or abstraction of the empirical evidence relating to judicial ideal points. Empirical calculations of the ideal points of federal judges have been estimated by Andrew Martin and Kevin Quinn, along with various co-authors.⁹⁰ For example, plotting these ideal points for United States Court of Appeals over the period 1953 to 2007 revealed a bi-modal distribution of ideal points.⁹¹ There is a bloc of judges to the left of center; there is a bloc of judges to the right. This is represented in **Figure 2**.

⁹⁰ See e.g., Andrew D. Martin and Kevin M. Quinn (2001), 'Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953-1999,' 10 *Political Analysis* 134; Martin, Quinn, and Epstein, cited above in note 47; Lee Epstein, Andrew D. Martin, Jeffrey A. Segal, and Chad Westerland (2007), 'The Judicial Common Space,' 23 *Journal of Law, Economics, & Organization* 303.

⁹¹ Epstein, Martin, Segal, and Westerland, cited above in note 90, use the NOMINATE Common Space scores of home state Senators and appointing Presidents to measure the ideology of federal judges. The measure uses the scores developed in Keith T. Poole (1998), 'Estimating a Basic Space From a Set of Issue Scales,' 42 *American Journal of Political Science* 954 (legislative and executive) and Micheal W. Giles, Virginia A. Hettinger, and Todd C. Peppers (2002), 'Picking Federal Judges: A Note on Policy and Partisan Selection Agendas,' 54 *Political Research Quarterly* 623.

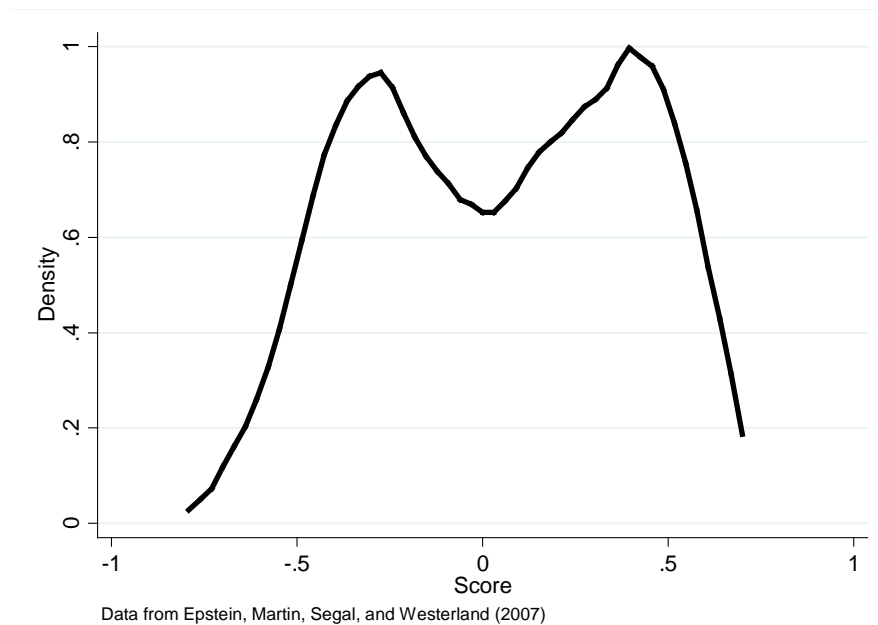


Figure 2: The density of the ideal points of all U.S. Court of Appeals judges from 1953-2007. The data indicate that the ideal points have a bimodal distribution.⁹²

iii. Decisions: stare decisis and discretion

It is commonly stated that judicial behavior is not adequately explained by either extreme realism or extreme legalism.⁹³ In some cases, a judge may have discretion to decide a case and may deliver a decision that reflects his or her own view of what the optimal law is. In other cases, however, the judge will feel bound by precedent. Even if the judge believes the decision is incorrect, over-ruling precedent can be costly. My model reflects this. It is assumed that the discretion is greater earlier in the lifespan of a law. Once the law becomes more certain, judicial discretion and influence is less apparent.

A judge has discretion to decide a case that does not fall with the ambit of a previous decision. An alternative way of thinking about this is that if legalist methods fail to provide an answer, the judge falls back to his or her ideological preferences. If the judge is not constrained by precedent, he or she can resolve the dispute in a manner that reflects

⁹² Data from Epstein, Martin, Segal, and Westerland, cited above in note 90.

⁹³ See e.g., Kornhauser, cited above in note 14; Posner, cited above in note 3; Klein, cited above in note 82.

his or her own view of the optimal law. Suppose that the first case drawn involves a driver who could see 243 feet of track when he was in the danger zone. Lewis is selected at random to hear the case. Recall that he is pro-railroad company and has an ideal point of 0 feet. He finds the driver contributorily negligent.

A judge does not have discretion in every case though. Judicial decisions have precedential effect and constrain judges that hear subsequent cases. It is assumed that it is prohibitively costly for a judge to overturn or, more accurately, act inconsistently with an earlier ruling.⁹⁴ Adherence to precedent is costless.

Readers may further be concerned that the precedent value of cases heard by a single judge is low.⁹⁵ The model can, under some assumptions, accommodate the presence of panels. Suppose decisions are made by panels of three judges. If we draw any three judges, the median voice on any panel will be drawn from one of the two blocs—essentially, the median voice on the panel is the single judge in my model. In this respect, the presence of panels has no effect upon the results.⁹⁶ This story does not, however, address some of the important moderating effects of panel composition noted in the introduction, which my model does not capture.⁹⁷

The assumption that judges follow precedent is incorporated into the model in the following way. After Lewis decides that a driver who could see 243 feet was negligent, an implicit rule is created: drivers who have a field of vision that is 243 feet or more will be held to be negligent. That is, not only does the first case cover the specific facts of 243 feet, but also all those cases where the driver could see more than 243 feet. Now, even if Rita hears a case where the driver had a field of vision of 300 feet, she would follow Lewis's precedent and hold the driver negligent. This implicit rule is depicted in **Figure 3**.

⁹⁴ The cost of acting inconsistently with precedent is assumed to be greater than the largest differential in utility for any fact pattern.

⁹⁵ This is commonly true; however, the decisions of single judges may have strong precedent effect. Take, for example, the famous copyright law decision of *Folsom v. Marsh*, 9 F. Cas. 342 (1841), or the decisions of Delaware Chancery Court.

⁹⁶ If the judges' ideal points are uniformly distributed, then there *is* an effect of panels here. The median voice is even less likely to be extreme, and therefore, the law will be even more likely to converge to the median judge's ideal point.

⁹⁷ See references, above at notes 17 to 21, and accompanying text.

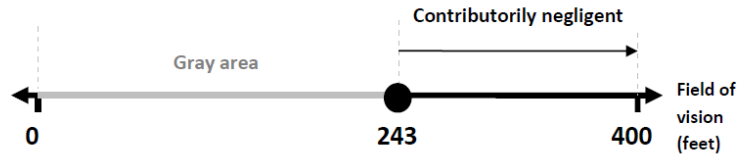


Figure 3: When the driver can see more than 243 feet of the track, he is held to be contributorily negligent. The decision says nothing about cases where the driver could see less than 243 feet. These fact patterns remain in the gray area.

There are, of course, limits to the strength of stare decisis. For the meantime, I shall simply assume that whenever a case falls within the determined area, it is very costly for judges to act inconsistently with such precedent.⁹⁸ I assume that cases falling within the implicitly determined area will never be litigated.⁹⁹ There are (at least) two justifications for this assumption. Drivers and railroad companies may change their behavior in response to changes in the law. For example, after Lewis’s first case, drivers behave more cautiously. An alternative explanation suggests that parties know the law and cases where the driver could see more than 243 feet of the track will never be litigated.

Subsequent cases are, thus, drawn from the gray area. After the first case is decided, the gray area shrinks from all possible fact patterns to scenarios where the driver can see between 0 feet and 243 feet of the track. The first judicial decision does not provide any information as to how cases within this range will be determined. If a case in this range arises, the judge will not be constrained by

⁹⁸ See e.g., Cardozo, cited above in note 1, at 149: “The labor of judges would be increased almost to the breaking point if every past decision could be reopened in every case, and one could not lay one’s own course of bricks on the secure foundation of the courses laid by others who had gone before him.”; Max Radin (1925), ‘The Theory of Judicial Decision: Or How Judges Think,’ 11 *American Bar Association Journal* 357; Justice John Paul Stevens (1983), ‘The Life Span of a Judge-Made Rule,’ 58 *New York University Law Review* 1.

⁹⁹ The non-ergodicity of cases is driven by this assumption that cases are not drawn—or at least are rarely drawn—from the determined or governed areas. If I drop this assumption and allow cases in the governed area to be brought, the following testable hypothesis emerges: judicial behavior becomes less attitudinal and more legalistic over time. Early decisions in the lifespan of a law are more likely to be correlated with judicial preferences; later decisions are more likely to be legalistic. After some time, the bulk of cases would have determinate, predictable outcomes. This corollary appears at some odds with recent empirical literature which suggests that judges have more “discretion” in later cases. See Stefanie A. Lindquist and Frank B. Cross (2005), ‘Empirically Testing Dworkin’s Chain Novel Theory: Studying the Path of Precedent,’ 80 *New York University Law Review* 1156.

precedent. The judge has discretion to decide in accordance with his or her beliefs.

Suppose the next case involves a driver who could see 130 feet of track. Rita is chosen at random to decide the case. She distinguishes this case from the only precedent on point. She notes that a field of vision of 130 feet is a lot less than 243 feet and holds in favor of the driver. This is depicted in **Figure 4**.

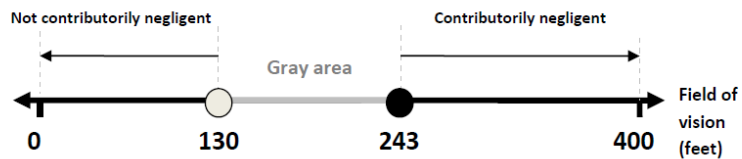


Figure 4: The second case narrows the gray area. The second case holds that if a driver could see 130 feet of track he is not contributorily negligent. Now the gray area shrinks to between 130 feet and 243 feet.

The gray area of the law shrinks over time. As more and more cases are decided, a greater portion of the spectrum of behavior becomes determined. Because of the assumption that judges strictly adhere to the implicit rules generated by precedents, the undetermined area never grows. The accumulation of decisions results in an explicit rule. For example, the law may converge to a bright line rule of 180 feet. Any driver that can see more than 180 feet of track and is hit by a train is contributorily negligent. Any driver that can see less than 180 feet of track is not negligent. This is illustrated in **Figure 5**.

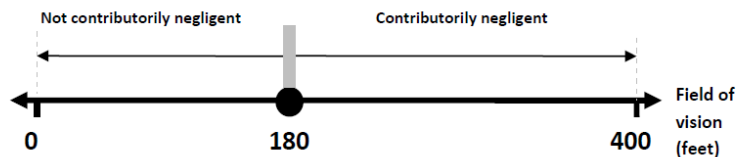


Figure 5: After a number of narrowly-decided and consistent opinions, the law converges to a threshold rule.

B. The solution

In the simulated example above, case-by-case adjudication generated a threshold rule of 180 feet. A rule in the neighborhood of

180 feet, however, was by no means guaranteed. The facts of the first case were drawn at random. The first driver could have seen anywhere between 0 feet and 400 feet. The law may have converged to a different rule depending on the facts of that first case. Further, the first judge in that case was drawn at random. In our example, Lewis held the first driver negligent. If Rita were drawn first the law would converge to a different, more pro-driver rule.

How does case law aggregate the different preferences of Lewis and Rita? The solution is not intuitively obvious. One might hypothesize that if randomly-selected judges narrowly decide cases on the facts, the law will tend to converge toward the mean average of the judges' ideal points. Sometimes Lewis decides disputes, pulling the expected boundary of liability toward 0 feet; sometimes Rita decides the disputes, pushing the expected boundary toward 400 feet. This intuition suggests that the distribution of converged rules is a single-peaked bell-shaped function, with the peak at the midpoint of the two judges' ideal points, 200 feet.

An alternative intuition may suggest the *ex ante* distribution of converged legal rules parallel the *ex ante* facts. Given that there is an equal likelihood of drawing a pro-railroad judge as there is of drawing a pro-driver judge, one might think that the distribution of converged rules would simply mirror the uniform distribution of facts. This hypothesis suggests that the probability that the converged rule will be between 0 and 10 feet is exactly the same as the probability that the converged rule will be between 200 and 210 feet.

Neither of these hypotheses, however, turns out to be correct. These intuitions fail to take into account the powerful effects of path dependence in the law when judges adhere to precedent. The mathematical solution to this model is by no means trivial. It is formally derived in the appendix.¹⁰⁰ For the purposes of the main text it will suffice to merely illustrate the solution and emphasize important features of the distribution of law. The distribution of converged rules is illustrated in **Figure 6**. It shows the probability of case law converging on any candidate rule. The probability density function is U-shaped, blowing up at 0 feet (Lewis's ideal point) and 400 feet (Rita's ideal point).

¹⁰⁰ See Appendix A.

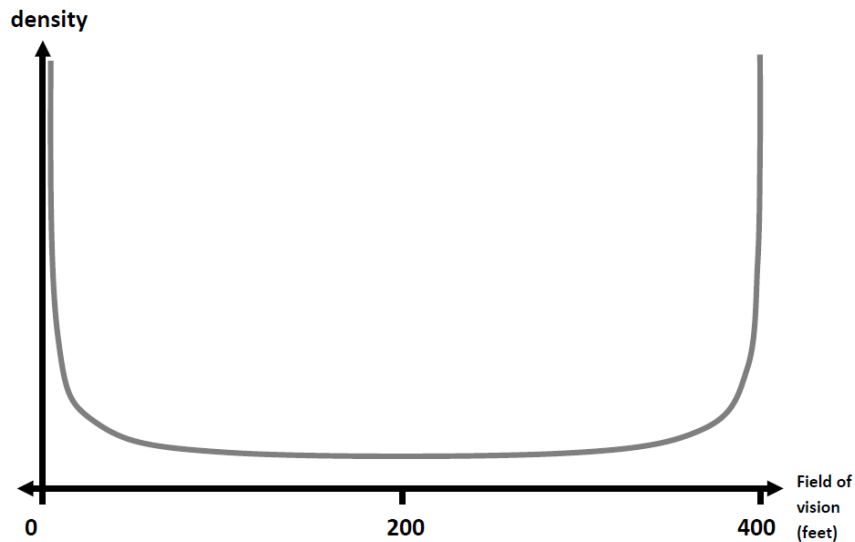


Figure 6: The probability density function of where the converged legal rule. The density function shows that the law is highly likely to converge to either of the ideal points of Lewis (0 feet) or Rita (400 feet).

It is important to point out two key properties of the *ex ante* density function of the converged rule: the expected rule and the variance.

(1) **The expected rule is the mean of the judges' ideal points.** The expected value of the distribution in our example is 200 feet. This is the mean and median of Lewis and Rita's ideal points.

(2) **The variance of the distribution is extremely high.** The density function blows up the judges' respective ideal points, 0 feet and 400 feet.

Although the expected law is 200 feet, the chances of the law converging to the neighborhood of 200 feet are very slim. Contrary to our intuitions, the density is at its lowest at 200 feet. That is, the chance of the law converging to the neighborhood of 200 feet is lower than any other candidate rule. The law is far more likely to end up near either ideal point than in the middle.

The tractable mathematical solution for this particular model allows me to be specific about probabilities. There is a greater likelihood of the law converging to a rule in the range of the lowest 5% or highest 5% than there is of the law converging to a rule in the middle 40%. That is, it is more likely that the rule will be below 20 feet

or above 380 feet than the rule will be between 120 feet and 280 feet.¹⁰¹ The chances of generating a law that reflects one of the judge's ideal points is comparatively *very* high.¹⁰²

Why does case law do such a poor job of compromising and finding the middle ground when judges simply resolve the dispute before the court? First, let's examine why the likelihood of the law converging to the mean of the two judges' ideal points is so low. Think of the unlikely set of circumstances that need to occur in order for the law to converge to the neighborhood of 200 feet. Lewis would need to hear all the cases where the driver could see more than 200 feet and Rita would have to be selected to hear cases where the driver could see less than 200 feet. Since judges are chosen at random, the chance of this occurring is low.

Now let's turn our attention to the two ideal points: 0 feet and 400 feet. Why is it that the likelihood of converging to the neighborhood of one of these points so high? Because the facts of potential cases are uniformly distributed and random cases come before the court, it might seem that the chance of getting a case with "extreme" facts (i.e., the facts of the case are close to either 0 feet or 400 feet) is quite slim. That intuition, however, focuses on the first draw of the urn only. There are multiple chances to draw facts close to one of the judges' ideal points. Even if we get an extreme fact pattern in the first case but the decision is not extreme (e.g., a driver who can see 380 feet is held to be contributorily negligent), we can still converge to the other end of the spectrum (e.g., toward a rule in the neighborhood of 0 feet) in subsequent cases. The chance of *eventually* getting an extreme case is high because there are multiple chances to draw such a case.

The model highlights the importance of path dependence in case law when judges adhere to precedent; a phenomenon that is well-recognized in the literature.¹⁰³ My model goes a step further though and, in describing the content of the legal rule, illustrates that the case law converges disproportionately toward the respective ideal points of the judges. To be sure, there is *some* compromising of the views; but it is not strong moderation. The likelihood of the law converging to the

¹⁰¹ The probability that the law will converge to the lowest 5% or highest 5% of the domain between the two ideal points is 28.7%. The probability that the law will converge to the middle 40% is around 26.2%.

¹⁰² The probability that the law will converge to the lowest 1% or highest 1% is around 13%. This is ten times as likely the law falling within a domain 1% either side of the midpoint, around 1.3%.

¹⁰³ See, Kornhauser, cited above in note 15; Hathaway, cited above in note 15; and Yeon-Koo Che and Jong Goo Yi (1993), 'The Role of Precedents in Repeated Litigation,' 9 *Journal of Law, Economics, & Organization* 399.

neighborhood of the judges' ideal points is high. I now explore extensions to the baseline model and examine conditions under which we might expect greater moderation.

C. Extensions to the model

i. Other distributions of judges' ideal points

The model makes very simple assumptions about how judicial ideal points are distributed; there are either pro-railroad judges (with an ideal point of 0 feet) or pro-driver judges (with ideal point of 400 feet). In this section, I will highlight the effect of different distributions of ideal points.

The model highlights the importance of variance in judicial ideology, rather than simply focusing on the median or mean ideal point of different judges. Imagine a jurisdiction with just one judge, Michael, who has an ideal point that is the same as the mean and median ideal point as in the two-judge example above, 200 feet. It is trivial to show that, in this world, the law will converge to the Michael's ideal point, 200 feet. In my model with heterogeneous preferences, there is a strong chance that two jurisdictions with exactly the same distribution of judges could end up with very different laws. If the dispersion of ideological preferences is high, then the mean or median of the ideal points will not be a sufficient statistic for informing the content of the legal rule.¹⁰⁴ That is, we may learn very little information about judge-made law from simply looking at the ideology of the median judge on a state court or a circuit. There is some empirical evidence to support this assertion.¹⁰⁵

¹⁰⁴ There is an empirical literature that focuses on the effect of the median judge, but this has primarily looked at the influence of the median judge on the Supreme Court. See e.g., Martin, Quinn, and Epstein, cited above in note 47; Lee Epstein and Tonja Jacobi (2008), 'Super Medians,' 61 *Stanford Law Review* 100. This literature is not directly applicable to the point that I am making here as these papers focus on the influence of the median on one panel that hears all cases. My analysis rather focuses on the median ideal point of many judges who hear cases individually.

¹⁰⁵ In a recent paper, Richard Posner, Andrei Shleifer, and I detail the judicially-created idiosyncratic exceptions to a tort doctrine—the economic loss rule—that different state courts have used. See Anthony Niblett, Richard A. Posner, and Andrei Shleifer (2010), 'The Evolution of a Legal Rule,' 39 *Journal of Legal Studies* (forthcoming). We empirically show that state courts use different idiosyncratic exceptions. That is, the line of liability is drawn differently in different states. Further, we showed that the different lines of liability that state courts drew with respect to the economic loss rule are not correlated with a measure of average political ideology of the judiciary in each state. We use the average PAJID scores from each state supreme court, calculated in Paul Brace, Laura Langer, and Melinda G. Hall (2000), 'Measuring

Readers may be concerned that the key result of my model—the case law does a poor job of finding the middle ground of different views when judges decide cases narrowly—is being driven by the assumption of two “extreme” judges. The key result still holds true if we have two judges with different ideal points that are not located at the very ends of the fact spectrum. Suppose that the facts of potential cases are still distributed uniformly between 0 feet and 400 feet, but Lewis has an ideal point of 100 feet and Rita has an ideal point of 300 feet. Under these assumptions, the *ex ante* distribution of the rule is U-shaped, blowing up at 100 feet and 300 feet. The law converges disproportionately to the respective ideal points, 100 feet and to 300 feet. The case law still has a low probability of converging to the midpoint, 200 feet. This is illustrated in **Figure 7**.¹⁰⁶

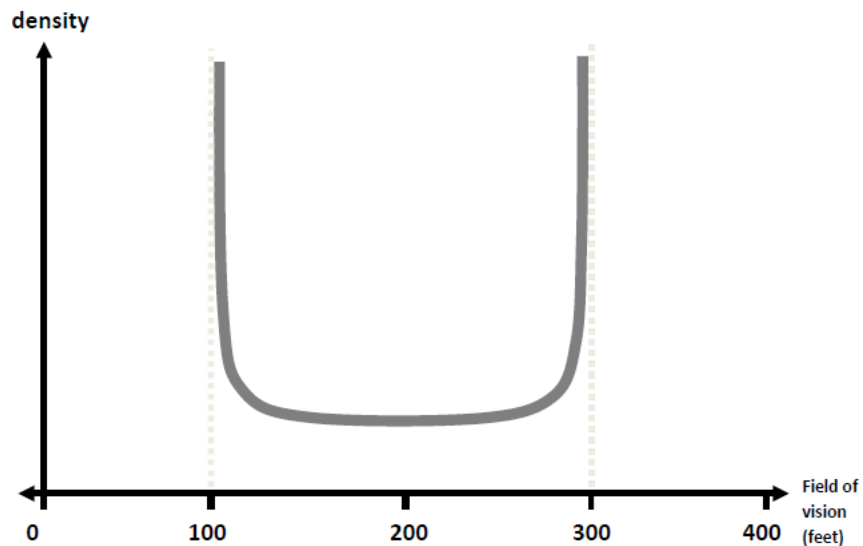


Figure 7: The *ex ante* distribution of the converged rule if there are two blocs of judges with ideal points at 100 feet and 300 feet.

Now let’s introduce more than two judicial ideal points into the story. Take, for example, the scenario where we have three judges: Lewis (ideal point, 0 feet), Michael (ideal point, 200 feet), and Rita

the Preferences of the State Supreme Court Judges,’ 62 *Journal of Politics* 387. We investigate the correlation between the average political ideology of each state court and the incidence of judges circumventing the economics loss rule with well-recognized exceptions. We find no correlation.

¹⁰⁶ The extension to Appendix A describes the formal solution.

(ideal point, 400 feet).¹⁰⁷ The *ex ante* distribution of converged rules again reflect the respective ideal points of the judges. It is illustrated in the left panel of **Figure 8**. The density function blows up at each of 0, 200, and 400 feet. As with before, there is slight moderation, and so the density is a little greater at 200 feet compared to the two outlying ideal points.

Let us now suppose that the judicial ideal points are uniformly distributed with ideal points ranging from 0 feet to 400 feet. This means that there are many, many different judges, each with a slightly different ideal point. The probability of a judge with an ideal point of 400 feet hearing the case is the same as the probability of a judge with an ideal point of 125 feet, or 288 feet, hearing the case. The *ex ante* distribution of the converged rule is now a single-peaked function. The peak is at the mean of the judges' ideal points. Now, the midpoint of the judges' ideal points has the highest probability, not the lowest probability. More starkly, the probability of getting the *most* extreme laws is now zero. The extreme views of judges essentially become nullified in a world where we have all viewpoints represented and judges decide cases narrowly. That is, if we have many different judges representing all different viewpoints, then the Cardozo Theorem appears to hold. This effect is illustrated in the right panel of **Figure 8**.

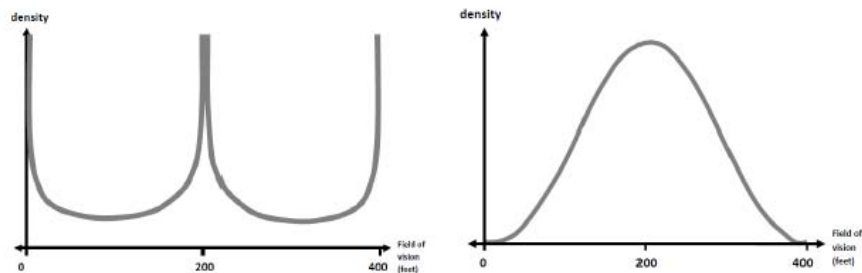


Figure 8: (Left panel) The *ex ante* distribution of the converged rule when there are three judges on the bench with ideal points of 0 feet, 200 feet, and 400 feet. (Right panel) The *ex ante* distribution of the converged rule when the judicial ideal points are uniformly distributed between 0 feet and 400 feet.

The punchline of this analysis is that moderate judges are required if moderate laws are likely to emerge in my model. The eccentricities of the more extreme judges *only* wash away if we introduce moderate judges into the story. Simply placing judges on either side of the

¹⁰⁷ Proofs of these propositions are not provided. The graphs of these figures were generated using simulations.

spectrum and hoping to balance, or counterweight, the eccentricities of the judges, does not work.

ii. Multiple dimensions of facts

In real life, the salient facts of cases are rarely described along one dimension. For example, when the Supreme Court investigates the constitutional validity of state courts' punitive damages awards, they look not only to the ratio of damages to harm, but they also look to the egregiousness of the defendant's behavior. Similarly, when deciding whether non-compete covenants are reasonable or not, state courts look at multiple factors such as the length of time that the covenant covers and the territorial restriction. In the real life contributory negligence cases of *Goodman* and *Pokora*, the court took into account other factors such as the driver's speed, the number of tracks that needed to be crossed, the weather, and the time of day.¹⁰⁸

As already noted, the use of one dimension in the model is a simplification. Each point on the line represents a bundle of facts. In the contributory negligence cases, for example, the bundle may include all the relevant facts, such as the time of day and the number of tracks. The use of one dimension simply suggests that all the bundles of facts can be ranked from the most egregious behavior to the least egregious behavior. At one end of the spectrum, we have the driver who is most likely to be found contributorily negligent; a driver who was very familiar with the one-track crossing, who could see 400 feet of track when the weather was fine and it was the middle of the day. At the other end of the spectrum, we have a driver who is least likely to be found contributorily negligent; a driver who was unfamiliar with the many-track crossing, who could see 0 feet of track during a storm in the middle of the night. In between these extremes, we can imagine a variety of different fact patterns. The assumption of one dimension assumes that these fact patterns are ranked in the same way by all judges. That is, the model can easily incorporate multiple dimensions provided that judges agree on the ordinal ranking of cases or have high correlation between their respective ordinal rankings of fact patterns.

But judges may have very different views of what factors are important. For example, some judges may place greater emphasis on

¹⁰⁸ For example, see *Goodman*: "He had been driving at the rate of ten or twelve miles an hour, but had cut down his rate to five or six miles at about forty feet from the crossing." 275 U.S. at 69. See also, *Pokora*: "The defendant has four tracks on Tenth Street;" 292 U.S. at 100. "Pokora made his crossing in the day time." 292 U.S. at 101.

the field of vision that the driver had; while other judges focus on the number of tracks. Different judges will rank the bundles of facts differently. Under this assumption, we still get convergence to a rule provided that judges adhere to precedent.¹⁰⁹ Convergence, however, is slower under this assumption. The more dimensions that judges deem relevant the slower the rate of convergence to a rule.¹¹⁰ A further complication arises if judges rank bundles of facts differently. Decisions of different judges may appear to be inconsistent if try to impose ordinal rankings upon the bundles of facts.

IV. APPLICATION OF THE MODEL: JUDICIAL APPOINTMENTS

When executives appoint judges they commonly seek to advance political goals through judge-made law.¹¹¹ The statistical evidence on appointees to federal courts is overwhelming. In the 140 years from 1869 to 2008, about 92% of the all 3,200 appointments to the lower federal courts went to candidates affiliated with the party of the appointing president.¹¹²

¹⁰⁹ In the two dimensional model, the rule converges to a line. The line may not, however, be a simple “linear” relationship. See e.g., Cameron M. Cameron and Lewis A. Kornhauser (2005), ‘Modeling Law: Theoretical Implications of Empirical Methods,’ paper presented at the NYU Law School Conference on Modeling Law; and Jonathan P. Kastellac (2010), ‘The Statistical Analysis of Judicial Decisions and Legal Rules with Classification Trees,’ 7 *Journal of Empirical Legal Studies* 202.

¹¹⁰ Gennaioli and Shleifer, cited above in note 11, assume that only relevant dimensions are introduced by judges. In their model, they assume two relevant dimensions. Patricio Fernandez and Giacomo Ponzetto suggest that judges can introduce irrelevant dimensions when distinguishing cases. Patricio A. Fernandez and Giacomo A. M. Ponzetto (2010), ‘Stare Decisis: Rhetoric and Substance,’ *Mimeo*: “In practice. . . nothing ensures that distinguishing occurs only on the basis of those empirical attributes that determine the efficient rule. On the contrary, each court has wide discretion in selecting the elements to be considered legally material.” at 12, following Julius Stone (1946), *The Province and Function of Law*.

¹¹¹ See generally, Sheldon Goldman, *Picking Federal Judges: Lower Court Selection From Roosevelt Through Reagan*; and Lee Epstein and Jeffrey A. Segal (2005), *Advice and Consent: The Politics of Judicial Appointments*. Although, see Sheldon Goldman (1995), ‘Judicial Selection Under Clinton: A Mid-Term Examination,’ 78 *Judicature* 276, discussing President Clinton’s nomination procedure: “there was a determined effort not to screen nominees ideologically. The president told Democratic senators and other officeholders that there should be no ideological screening. Excellence in intellectual ability and judicial temperament were of paramount importance as well as a firm understanding of the role of district courts to follow precedent and to be fair.” at 279.

¹¹² See, Epstein and Segal, cited above in note 111, at 26-7 for a summary of data from 1869 to 2004. Epstein and Segal obtain data on party affiliation of lower federal courts were obtained from Deborah J. Barrow, Gary Zuk, and Gerard S. Gryski (1996), *The Federal Judiciary and Institutional Change*; Sheldon Goldman, Elliot Slotnick, Gerard

There exists a view among some legal scholars that newly-appointed judges should offset or counterbalance the biases of currently seated judges. Cass Sunstein, for example, has written: “President Clinton chose two centrist judges for the Supreme Court, Ruth Bader Ginsburg and Stephen Breyer. . . [B]ecause of their centrism, they cannot be seen to as ideological counterweights to Justices Antonin Scalia and Clarence Thomas.”¹¹³ Empirical evidence suggests that the conservative Justices Antonin Scalia and Clarence Thomas have voted in a way that is *more* conservative than the records of their appointing presidents, Ronald Reagan and George H. W. Bush.¹¹⁴ This desire to counterbalance judicial ideology on courts is not limited to the Supreme Court.¹¹⁵

In this section, I use the model described in section III to challenge the idea that appointing ideological counterweights will generate balance in the law. The model suggests that the eccentricities of judges do not balance one another. The model suggests that ideological counterweights are not likely to generate moderate rules on the bench, even if judges decide cases very narrowly.

Take the scenario previewed in the introduction. A moderate executive needs to fill a judicial vacancy. The only other judge on the bench is Rita, a judge whose views clash with the executive. She is too pro-driver for the executive’s liking. If Rita hears all cases looking at

Gryski, Gary Zuk, and Sara Schiavoni (2003), ‘W. Bush Remaking the Judiciary: Like Father Like Son?’ 86 *Judicature* 282; and Sheldon Goldman, Elliot Slotnick, Gerard Gryski, and Sara Schiavoni (2005), ‘W. Bush’s Judiciary: The First Term Record,’ 88 *Judicature* 244. I have updated the data to include George W. Bush’s appointments from 2005-09 using data from Sheldon Goldman, Sara Schiavoni, and Elliot Slotnik (2009), ‘W. Bush’s Judicial Legacy,’ 92 *Judicature* 258.

¹¹³ Cass R. Sunstein (2005), *Radicals in Robes*, at 14.

¹¹⁴ See, Michael Bailey and Kelly H. Chang (2001), ‘Comparing Presidents, Senators, and Justices: Interinstitutional Preference Estimation,’ 17 *Journal of Law, Economics, & Organization* 477, at 495.

¹¹⁵ See e.g., Committee on the Judiciary, United States Senate, ‘The District of Columbia Circuit: The Importance of Balance on the Nation’s Second Highest Court,’ (September 24, 2002); and the comments of Senator Charles E. Schumer, cited above in note 26. Further, take the example of the nomination by Bill Clinton of William Fletcher to the Ninth Circuit. The Republican Senate refused to appoint the liberal Fletcher unless the president appointed Republican Senator Slade Gordon as well. See Sarah Wilson (2003), ‘Appellate Judicial Appointments During the Clinton Presidency: An Inside Perspective,’ 5 *Journal of Appellate Practice and Process* 29, at 43-7. Lee Epstein and Jeffrey Segal noted in 2005 that the Federalist Society prominently lobbied George W. Bush to “counterbalanc[e] what it decries as the ‘orthodox liberal’ ideology that ‘dominates’ the legal community.” Epstein and Segal, above 111, at 52. See also, ‘Blocking Judicial Ideologues,’ *New York Times*, April 27, 2001, at 24A.

drivers' contributory negligence—and even if she decides all the cases narrowly—the law will converge to a threshold rule of 400 feet. This is costly for the centrist appointer who believes the optimal rule is 200 feet. There are two candidates for the vacant judicial position: Lewis and Michael. Lewis has an ideal point of 0 feet. His extreme ideal point appears to be a counterweight to Rita's extreme view of 400 feet. If Lewis is appointed, the average ideal point will be 200 feet, which is the same as the appointer's ideal point. Michael, on the other hand, is a moderate and shares the appointer's ideal point. He believes that the optimal rule is 200 feet.

My model helps determine the content of the law if either judge is appointed. If Lewis is appointed, a judge with an ideal point of 0 feet, the law will converge to a U-shaped distribution with peaks at 0 feet and 400 feet. The likelihood of the law being 200 feet is slim. This is the *ex ante* distribution in section III. It is illustrated in the left panel of **Figure 9**. If Michael is appointed, the *ex ante* distribution of laws will be different. With this distribution of judicial ideal points, there is no chance that the law can converge to a rule below 200 feet. If a case arises with a driver who can see, say, 100 feet, then it matters not whether Rita hears the case or whether Michael hears the case. Both judges believe that this driver is not negligent. This is true for all cases where the driver can see less than 200 feet, precluding the possibility of the law converging to a rule within this range. The *ex ante* distribution of the law is illustrated in the right panel of **Figure 9**.

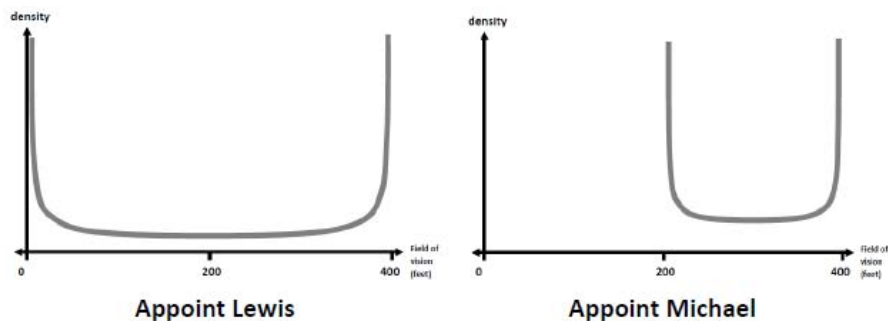


Figure 9: (Left panel) The *ex ante* distribution of laws if Lewis, ideal point 0 feet, is appointed to sit with Rita. (Right panel) The *ex ante* distribution of laws if Michael, ideal point 200 feet, is appointed.

If the executive appoints an extreme judge like Lewis to balance out Rita's extreme views, the mean position of the *ex ante* law is 200 feet. The mean position of the law, however, is not the only factor to be taken into account. The *ex ante* variance of the final position of law is also important. Such variance is costly. If the executive appoints Lewis

it is highly likely that the law will converge to the neighborhood of 0 feet or 400 feet. If, on the other hand, the executive appoints Michael, there is still a chance that Rita's views will dominate and the law will converge to 400 feet. But now there is a strong chance that the law will reflect the centrist's view and a rule in the neighborhood of 200 feet will emerge.

I assume that, for a centrist executive, extreme rules to the left are just as costly as extreme rules to the right. Deviations away from his ideal point in either direction are equally costly. That is, the cost functions are *symmetric*.¹¹⁶ I also assume that the cost functions are continuous functions, minimized at the executive's ideal point, and monotonically increasing as the law deviates from the ideal point. Under these conditions, it is better for the executive to appoint Michael, the moderate, to the bench. The mathematical description of this problem is reserved for the appendix.¹¹⁷

My approach has taken the view that the executive's objective in appointing judges is *outcome* driven. That is, the executive cares only about the content of the legal rule. There are, of course, other factors that weigh into the appointment decision. For example, the executive is constrained by other political players who have the power to block judicial nominees, such as the Senate.¹¹⁸ Since extremists are more likely to be blocked, the argument for appointing a moderate judge is strengthened.¹¹⁹ On the other hand, the executive must appease the

¹¹⁶ A *symmetric* cost function implies that the executive with an ideal point of 200 feet experiences the same cost under a rule of 100 feet as he does with a rule of 300 feet. Further, the executive is punished equally under rules of 0 feet and 400 feet.

¹¹⁷ See Appendix B. The result, that an appointer with moderate views should not select judges to provide a counterweight to extreme judges, is not driven by the shape of the appointer's cost function. There are some assumptions about the appointer's cost function—such as symmetry, continuousness, and monotonicity—that do limit this result, however.

¹¹⁸ Article II, Section 2, paragraph 2 of the United States Constitution states: "[The President] shall have Power, by and with the *Advice and Consent* of the Senate, to make Treaties, provided two thirds of the Senators present concur; and he shall nominate, and by and with the Advice and Consent of the Senate, shall appoint Ambassadors, other public Ministers and Consuls, Judges of the supreme Court, and all other Officers of the United States, whose Appointments are not herein otherwise provided for, and which shall be established by Law: but the Congress may by Law vest the Appointment of such inferior Officers, as they think proper, in the President alone, in the Courts of Law, or in the Heads of Departments" (my emphasis).

¹¹⁹ See e.g., discussions of the failed nominations by President Ronald Reagan of Robert Bork and Douglas Ginsburg to the Supreme Court in Stephen L. Carter (1994), *The Confirmation Mess*; and Jan Crawford Greenburg (2007), *Supreme Conflict*.

party base and various interest groups which *may* lead toward nominations of more extreme candidates.¹²⁰

I now explore five extensions to this example. The punchline of all five extensions is that the main result on judicial appointments does not change: the centrist executive may be better off appointing the moderate judge and not seeking to counterbalance the biases of sitting judges.

There are, of course, many considerations that an executive will take into account when appointing judges. For example, the executive may wish to overturn the precedents of existing or earlier judges. In the context of the Supreme Court, for example, the Reagan and Bush administrations were hoping to appoint judges who would roll back some prominent Warren and Burger-era precedents, such as *Roe v. Wade*.¹²¹ More recently, there has been some discussion over whether Obama nominees will roll back the decisions of the Rehnquist and Roberts courts.¹²² This is not a recent phenomenon. In the Lincoln-Douglas debates of 1858, then-Senate candidate Abraham Lincoln promised to appoint judges that would overrule the *Dred Scott* decision.¹²³ This is the basis for the first extension: what if judges can overrule previous decisions? Does this change the conclusion over which should be appointed? The answer is no.

If judges are able to costlessly overturn precedent—or at least costlessly act inconsistently with precedent¹²⁴—then executives should still appoint judges who have similar preferences to themselves. A centrist executive, for example, would prefer judges to overturn extreme precedents in favor of moderate laws rather than have extreme precedents overturned in favor of equally extreme rulings at

¹²⁰ See discussion in Ch. 1 of Epstein and Segal, cited above in n 111. Crawford Greenburg, cited above in n 119, has an interesting discussion of the factors underpinning judicial nominations for, amongst others, Bork, Ginsburg, and Anthony Kennedy (ch. 2), David Souter (ch. 4), and Clarence Thomas (ch. 5).

¹²¹ *Roe v. Wade*, 410 U.S. 113 (1973). On the question of courts, judges, and other lawmakers factoring into presidential elections and campaigns, see William G. Ross (2002), 'The Role of Judicial Issues in a Presidential Campaign,' 42 *Santa Clara Law Review* 391.

¹²² See e.g., Patrick Healy, 'Seeking to Shift Attention to Judicial Nominees', *New York Times*, October 5, 2008.

¹²³ *Dred Scott v. Sandford*, 60 U.S. (19 How.) 393 (1857). On the Lincoln-Douglas debates, see Harold Holzer (ed.) (1993), *The Lincoln-Douglas Debates*.

¹²⁴ A recent empirical study tracks inconsistent decisions in disputes involving the unconscionability doctrine in California. A key finding of that paper is that ideology of different judges is a significant driver of inconsistency between decisions. See Anthony Niblett (2010), 'Tracking Inconsistent Judicial Behavior,' *Mimeo*.

the other end of the spectrum.¹²⁵ By appointing an offsetting judge in this world, the center cannot hold. The law does not converge in a world of costless, or low cost, overturning.¹²⁶

A second extension one may consider is the effect of judges issuing broader rulings. If judges issue broad rulings (and these rulings are respected by judges in subsequent cases), this merely strengthens the claim for appointing the moderate judge. If a judge, upon hearing the first case, simply sets a rule that comports with his or her ideal point, then it is in the interests of the centrist executive to appoint Michael rather than Lewis. If Michael is appointed there will be a 50 per cent chance of a rule of 200 feet; if Lewis is appointed there is a 0 per cent chance of such a rule.

A third extension explores the effect of appointing a judge in a situation when there are existing precedents that will be respected, rather than a legal *tabula rasa*. For example, let's suppose that the first case was decided by Lewis and he held that the driver who could see 243 feet of track contributorily negligent. But then Lewis retires. At the time of the new appointment the gray area of the law is the domain between 0 feet and 243 feet. Here, the issue is not overruling but rather operating within the constraints of existing precedents. It is still the case that the executive should appoint a judge who shares his ideal point rather than a judge who will offset Rita's pro-driver preferences. Even if the precedent precludes any chance that the law can converge on the executive's ideal point, it is still in the interest of the executive to appoint a judge with similar preferences.

A fourth extension considers the possibility that the ideology of judges is not perfectly known at the time of appointment. For example, a nominated judge, perceived to be a moderate at the time of appointment, may turn out to be either left or right of the center when sitting. President Eisenhower regretted the appointments of Earl Warren and William J. Brennan, Jr. to the Supreme Court; as did President Truman of the appointment of Tom Clark. Empirical evidence suggests that notable "ideological shifts" of judges such as Harry Blackmun, John Paul Stevens, Sandra Day O'Connor, and David Souter

¹²⁵ Saul Brenner and Marc Stier find that extreme judges are more likely to vote against the precedent than centrist judges. Saul Brenner and Marc Stier (1996), 'Retesting Segal and Spaeth's *Stare Decisis* Model,' 40 *American Political Science Review* 1036. As the authors note in footnote 2 at 1037, however, "[t]his finding is partly tautological, for strong ideologues achieve their high or low scores, in part, because of their failure to conform to precedents that are in conflict with their preferences."

¹²⁶ The results on overturning are consistent with the findings of Gennaioli and Shleifer, cited above in note 12.

have been documented.¹²⁷ Overall, however, the empirical evidence suggests that, with respect to the Supreme Court at least, the ideological divergence between appointing presidents and appointed judges are only slight. The voting behavior of judges and the political ideology of the nominating president is strongly correlated.¹²⁸ The data, however, also suggest that the correlation drops for justices with more than ten years service.¹²⁹ At the court of appeal level, the positive correlation between the president's ideology and the voting behavior of judges is slightly weaker.¹³⁰

In my model, the effects of an ideological shift depend on whether ideological shifts are unbiased or not. If judiciary ideology can shift both ways over time—and the expected shift is zero—then the effect is neutral. Anecdotal evidence may suggest that it is more likely that justices become more liberal over time. Justices Stevens and Souter are notable examples. If this bias were both true and endemic in all judges, then the executive may wish to offset this bias by appointing judges who are further to the right. In my example, the argument for appointing the moderate judge is even stronger. However, the empirical evidence does not necessarily bear this out, suggesting that some judges become more liberal on the bench than expected and others become more conservative than expected.¹³¹

¹²⁷ See Epstein and Segal, cited above in note 111, at 119-20; and Martin, Quinn, and Epstein, cited above in note 47, at 1310 who illustrate the ideological shift in the voting behavior of Sandra Day O'Connor over time (from more conservative to more liberal.)

¹²⁸ Epstein and Segal, cited above in note 111, for example show that the correlation between liberal votes cast by the justice (percent) and the president's ideology from Presidents Eisenhower through to President Clinton is +0.64 (Justices Warren through to Breyer), at 131. Epstein and Segal use data from Keith Poole on presidential ideology (see e.g., Nolan N. McCarty and Keith T. Poole (1995), 'Veto Power and Legislation: An Empirical Analysis of Executive and Legislative Bargaining from 1961-1986,' 11 *Journal of Law, Economics, & Organization* 282); and Harold J. Spaeth's data on Supreme Court voting, available at: www.as.uky.edu/polisci/ulmer-project/UlmerProject/databases.htm.

In 1985, Laurence Tribe suggested that the problem of ideological shift is a myth, because "in areas of particular and known concern to a President, Justices have been loyal the ideals and perspectives of the men who have nominated them." See Laurence Tribe (1985), *God Save This Honorable Court*, at 60.

¹²⁹ Epstein and Segal, cited above in note 111, show that the correlation drops to +0.49 for this subset of Supreme Court justices.

¹³⁰ See e.g., Cass R. Sunstein, David Schkade, and Lisa M. Ellman (2004), 'Ideological Voting on the Federal Courts of Appeals,' 90 *Virginia Law Review* 301.

¹³¹ Epstein and Segal, cited above in note 111, at 125 compare the perceived ideology of the judge at the time of nomination to the voting behavior of each judge on the Supreme Court. The data used are from Jeffrey A. Segal and Albert D. Cover dataset

Fifth, what if there are other candidates besides Lewis and Michael? The model is simplistic in assuming that the executive's choice set consists only of Lewis and Michael. Restricting the choice set of judges in this way adduces the bold claim that centrists should not counterbalance the biases of sitting judges. Suppose now that the pool includes potential judges with all ideal points. Under these conditions, it may be optimal for the executive to appoint a judge whose preferences *slightly* offset the preferences of an existing judge—but it is not in the executive's best interests to fully offset or counterbalance these biases. Suppose, for example, that the executive has a "linear" cost function. That is, as the law diverges further from the executive's ideal point of 200 feet, his costs increase at a constant rate. It is twice as costly if the rule is 300 feet than if the rule is 250 feet. Under these assumptions, the executive would lean toward a judge who has an ideal point of about 160 feet.¹³² The functional form of the executive's cost function does affect the selection of the judge now; but under all sensible cost functions, my model suggests that it is still the case that the law better reflects a centrist appointer's view if he does not try to completely offset the ideology of extreme judges.

V. CONCLUSION

This paper has examined how standards converge to rules when judges decide cases narrowly and respect precedent. The model determines the content of these legal rules if judges have different ideological preferences over the position of the law. The model suggests that the domain of the Cardozo Theorem is perhaps more modest than previously thought. Case law may not compromise and reconcile different positions of judges. If there are two groups of judges, the law will be unlikely to converge to the midpoint of the judges' respective ideal points; the biases of judges are unlikely to wash out.

If the executive wishes for judges to deliver moderate laws, then the executive should appoint moderate judges. It is unlikely that the extreme views of judges will be washed out by simply appointing judges with equally extreme views at the opposite end of the spectrum.

that creates a score of perceived ideology for each supreme court justice since 1953 using newspaper editorials at the time of the nomination (see e.g., Charles D. Cameron, Albert D. Cover, and Jeffrey A. Segal (1990), 'Senate Voting on Supreme Court Nominees: A Neo-Institutional Model,' 84 *American Political Science Review* 525) and Harold J. Spaeth's website on voting behavior, cited above in note 128.

¹³² See the extension to Appendix B.

The model suggests that resolving disputes narrowly and not setting broad rules does little to negate the effects of judges with extreme views. Deciding cases narrowly is not a necessary, nor a sufficient, condition for generating centrist laws. Even though judges who decide cases narrowly are constrained by the facts of the dispute before them, the law will tend to reflect one of the extreme views if putatively offsetting extreme judges are appointed.

APPENDIX

A.

Distribution of the converged law in the baseline model.

I show how the law converges when facts initially are uniformly distributed with mass one, that is, $f \sim U[0,1]$. Judges have the ideal points $j \in \{0,1\}$ with $Pr(j = 0) = \frac{1}{2}$. To keep in line with our baseline model, we can think of $j = 0$ as Lewis, and $j = 1$ as Rita.

Before any cases have been decided the undetermined region is $[0,1]$. Define X as the final position of the law, and let $H(x) = Pr\{X \leq x\}$ be the distribution function with density $h(x) = H'(x)$. For a fixed x , after the first case there are four possibilities:

Possibility 1: The drawn facts in case 1 were less than the candidate rule x and Lewis was the judge: $f_1 \leq x$ and $j_1 = 0$. In this case, $X \leq x$. This happens with probability, $x/2$.

Possibility 2: The drawn facts in case 1 were greater than the candidate rule x and Rita was the judge: $f_1 \geq x$ and $j_1 = 1$. In this case, $X \geq x$. This happens with probability, $(1-x)/2$.

Possibility 3: The drawn facts in case 1 are greater than our candidate rule x and Lewis was the judge. Let's call the drawn facts s . Here, $f_1 = s \geq x$ and $j_1 = 0$. In this case, the current interval is $[0, s]$ and given this, the probability that $\{X \leq x\}$ is $H\left(\frac{x}{s}\right)$.

Possibility 4: The drawn facts in case 1 are less than our candidate rule x and Rita was the judge: $f_1 = s \leq x$ and $j_1 = 1$. In this case, the current interval is $[s, 1]$ and given this, the probability that $\{X \leq x\}$ is $H\left(\frac{x-s}{1-s}\right)$.

Using this information, the cumulative distribution function can be written:

$$H(x) = \frac{x}{2} + \frac{1}{2} \int_x^1 H\left(\frac{x}{s}\right) ds + \frac{1}{2} \int_0^x H\left(\frac{x-s}{1-s}\right) ds$$

Differentiating with respect to x and using $H(0) = 0$ and $H(1) = 1$, the density function satisfies:

$$h(x) = \frac{1}{2} \int_0^x h\left(\frac{x-s}{1-s}\right) \frac{1}{1-s} ds + \frac{1}{2} \int_x^1 h\left(\frac{x}{s}\right) \frac{1}{s} ds$$

Using the change of variables $y = \frac{x-s}{1-s}$ in the first integral and $y = \frac{x}{s}$ in the second, the equation can be written:

$$h(x) = \frac{1}{2} \int_0^x \frac{h(y)}{1-y} dy + \frac{1}{2} \int_x^1 \frac{h(y)}{y} dy$$

To solve this, we take the derivative with respect to x once more:

$$h'(x) = \frac{1}{2} \cdot \frac{h(x)}{1-x} - \frac{1}{2} \cdot \frac{h(x)}{x} = \frac{1}{2} h(x) \left\{ \frac{1}{1-x} - \frac{1}{x} \right\} = \frac{1}{2} h(x) \left\{ \frac{2x-1}{x(1-x)} \right\}$$

It follows that:

$$\frac{h'(x)}{h(x)} = \frac{1}{2} \cdot \frac{2x-1}{x(1-x)}$$

$$[\ln h(x)]' = \frac{1}{2} \cdot \frac{2x-1}{x(1-x)}$$

Therefore:

$$\ln h(x) = \int_a^x \frac{2t-1}{t(1-t)} dt + c$$

The solution that satisfies both this equation and $\int_0^1 h(x) dx = 1$ is:

$$h(x) = \frac{1}{\pi \cdot \sqrt{x(1-x)}}$$

This distribution has the following features:

- (1) $H(0.5) = 0.5$
- (2) $h(x)$ blows up at $x = 0$ and $x = 1$.

In order to generate the function used in the graphical example in section III, the facts are distributed uniformly over 0 feet to 400 feet, with the two judges having ideal points of 0 (Lewis) and 400 (Rita). It follows that the distribution is:

$$h(x) = \frac{1}{\pi \cdot \sqrt{x(400-x)}}$$

Distribution of the converged law if the judges are not at 0 or 1.

For the more general case where facts are distributed uniformly over $[a, b]$, but judges are drawn from a distribution where $j \in \{c, d\}$ where $a \leq c \leq d \leq b$ and $\Pr(j = d) = \frac{1}{2}$.

The law cannot converge to a position between 0 and c , since all judges agree that such facts do not warrant liability. Similarly, the law cannot converge to a position between d and b as all judges agree this behavior should attract liability. The only area of interest is the area between c and d . When the facts are drawn from this area, cases will be decided differently depending on the judge.

The proof therefore mimics the proof in A , with the exception that the lower bound is c instead of 0, and the upper bound is d instead of 1.

The *ex ante* distribution of the law is:

$$h(x) = \frac{1}{\pi \cdot \sqrt{(x-c)(d-x)}}$$

B.

Cost functions and the costs associated with appointing judges

The appointer has the cost function $c(x)$, which is symmetric, continuous, and minimized at 200. Formally, $c'(200) = 0$, $c'(x) > 0$ for $x > 200$, and $c'(x) < 0$ for $x < 200$.

Assume Rita is on the bench. If the only judge on the bench is Rita, the law will converge to a rule of 400 feet. The two options are to appoint: (1) Michael (ideal point 200 feet); or (2) Lewis (ideal point 0 feet). I call the two distributions Michael's curve and Lewis's curve for convenience. There exists a point $p \in (200, 400)$ where the following holds:

$$\frac{2}{\pi \cdot \sqrt{p(400-p)}} = \frac{1}{\pi \cdot \sqrt{(200-p)(400-p)}}$$

That is, Lewis's curve multiplied by a scalar of 2 crosses Michael's curve.

Since $h'(n) < 0$ for Michael's curve over $n \in (200, p)$, and $h'(n) > 0$ over the same domain for Lewis's curve, it holds that for any $n \in (200, p)$:

$$2 \int_{200}^n \frac{1}{\pi \sqrt{x(400-x)}} dx < \int_{200}^n \frac{1}{\pi \sqrt{(x-200)(400-x)}} dx$$

That is, for any candidate $n \in (p, 400)$ the area under Michael's curve is greater than twice the area under Lewis's curve between 200 and p .

Consequently for all $n \in (p, 400)$:

$$2 \int_n^{400} \frac{1}{\pi \sqrt{x(400-x)}} dx > \int_n^{400} \frac{1}{\pi \sqrt{(x-200)(400-x)}} dx$$

The area under Michael's curve is less than twice the area under Lewis's curve between p and 400.

Since the area under both curves is equal to 1, it holds that:

$$\begin{aligned} \int_{200}^p \frac{1}{\pi \sqrt{(x-200)(400-x)}} dx - 2 \int_{200}^p \frac{1}{\pi \sqrt{x(400-x)}} dx = \\ 2 \int_p^{400} \frac{1}{\pi \sqrt{x(400-x)}} dx - \int_p^{400} \frac{1}{\pi \sqrt{(x-200)(400-x)}} dx \end{aligned}$$

Using this result and since $c'(x)$ is increasing in x over the domain, $x \in (200, 400)$, it holds that:

$$2 \int_{200}^{400} \frac{c(x)}{\pi \sqrt{x(400-x)}} dx > \int_{200}^{400} \frac{c(x)}{\pi \sqrt{(x-200)(400-x)}} dx$$

That is, the cost to the executive is lower from appointing Michael (RHS) than from appointing Lewis (LHS).

Drawing the optimal judge from a large pool

Assume that the cost function of the appointer is linear. That is, the utility function is simply the absolute value of the difference between the actual converge rule and the ideal point of the appointer. In accordance with the example in the text of the paper, the ideal point of the appointer is 200 feet.

The linearity and symmetry of the cost function suggests that a rule of 300 feet is twice as costly as a rule of 250 feet, but equally costly as a law of 100 feet. The utility function is:

$$c(x) = |x - 200|$$

Assume Rita is already on the bench. The optimal appointee has ideal point, j^* , where j^* is defined:

$$j^* = \arg \min \int_j^{400} \frac{|x - 200|}{\pi \sqrt{(x - j)(400 - x)}} dx$$

The solution is approximately 161.16 feet.