

Stare Decisis: Rhetoric and Substance*

Patricio A. Fernandez
Harvard University

Giacomo A. M. Ponzetto
CREI and Universitat Pompeu Fabra

Revised, April 2010

Abstract

Stare decisis allows common law to develop gradually and incrementally. We show how judge-made law can steadily evolve and tend to increase efficiency even in the absence of new information. Judges' opinions must argue that their decisions are consistent with precedent: this is the more costly, the greater the innovation they are introducing. As a result, each judge effects a cautious marginal change in the law. Alternative models in which precedents are either strictly obeyed or totally discarded would instead predict abrupt large swings in legal rules. Thus we find that the evolution of case law is grounded not in binary logic fixing judges' constraints, but in costly rhetoric shaping their incentives. We apply this finding to an assessment of the role of analogical reasoning in shaping the joint development of different areas of law.

JEL classification: K13, K40

1 Introduction

Case law develops through the rulings of appellate judges bound by stare decisis. This principle constrains courts to render a judgment consistent with previously decided cases, and at the same time empowers them to make law by setting a precedent that will bind future judges. Thus stare decisis serves two seemingly contradictory purposes: on the one hand, it endows case law with consistency and predictability; on the other, it provides for its gradual evolution (Wright, 1939; 1943). Understanding the efficiency properties of this

*E-mail: pfernand@fas.harvard.edu and gponzetto@crei.cat. We are grateful to Ian Ayres, Nicola Genaioli, Oliver Hart, Louis Kaplow, Anthony Niblett, Andrei Shleifer and two anonymous referees for helpful comments. Financial assistance from the Institute for Humane Studies, the Spanish Ministry of Science and Innovation (grants ECO2008-01666 and CONSOLIDER-Ingenio 2010 CSD2006-00016), the Barcelona GSE Research Network and the Generalitat de Catalunya (grant 2009SGR 1157) is gratefully acknowledged.

process is crucial for interpreting the empirical evidence on the economic consequences of legal origins (La Porta et al., 1998; La Porta et al., 2008), which are significantly driven by cross-country differences in the degree of reliance on judicial decisions as a source of law (Beck et al., 2003; 2005; La Porta et al., 2004).

Posner ([1973] 2007) advanced the seminal hypothesis that, through the development of judge-made rules, common law tends to achieve efficient outcomes. Cooter et al. (1979) derived this result from a model in which efficiency-maximizing courts have imperfect information which improves over time. But real-world judges do not necessarily aim at welfare maximization; on the contrary, there is a growing consensus that they hold idiosyncratic preferences and biases (Partridge and Eldridge, 1974; Rowland and Carp, 1996; Revesz, 1997; Pinello, 1999; Klein, 2002; Zywicki, 2003; Sunstein et al., 2004; Posner, 2005a). Gennaioli and Shleifer (2007a) capture this heterogeneity by introducing a theoretical framework to analyze judge-made law on the basis of two assumptions: first, judges have different tastes and ideologies; second, changing legal rules is personally costly for a judge. Furthermore, they assume that case law can only evolve when new information becomes available to the courts. Judges can then distinguish cases from precedents by increasing the informational content of the rule. The improvement in information translates into an increase in the precision of the law, and thus into legal change that is on average beneficial, in spite of judges' biases.

In reality, however, the evolution of case law is constantly ongoing, while discontinuous improvements in the availability of information occur only rarely. Nor does distinguishing require courts to exploit previously unobservable economic information. Judges can distinguish on the basis of purely procedural or conceptual grounds, or of factual considerations that may be deemed legally material despite being devoid of economic relevance for the determination of the efficient rule. Legal realists in particular have emphasized that *stare decisis* is a flexible process and not a hard and fast rule (Holmes, 1881; 1899; Cardozo, 1921; Llewellyn, 1930; 1960; Radin, 1933; Cohen, 1935; Stone, 1946; 1959; 1964; 1969; 1985; Douglas, 1949; Frank, 1949).

In this paper, we incorporate these insights by extending the Gennaioli–Shleifer framework to allow distinguishing in the absence of new information. Our explanation of the evo-

lution of case law focuses on the incentives created by the rhetoric of stare decisis. Judges' decisions are shaped by the duty to justify them, and the doctrine of binding precedent requires this justification to be made in terms of the continuity of the law. Hence, greater rhetorical effort is necessary to argue persuasively in support of a greater departure from precedent. Whereas Gennaioli and Shleifer (2007a) assume that some legal innovations are logically impossible but all the feasible ones are equally costly, we assume instead that all are feasible, but that their cost to the judge is continuously increasing in the extent of the deviation from the inherited rule.

Our model predicts the constant, gradual evolution of case law. Distinguishing is beneficial even when it introduces no new information, because it ensures the inclusion of a variety of perspectives into the law (Cardozo, 1921). In a setting of imperfect information, the process involves long-run randomness and the possibility of errors, but it also induces convergence toward more efficient rules, in accordance with Posner's ([1973] 2007) hypothesis. If perfect information becomes available to the courts, the development of the law remains gradual, but improvement becomes certain, and the first best is eventually achieved.

In the spirit of pragmatism, we argue that rhetorical requirements shape the logical structure of the law, rather than vice-versa (Dewey, 1924). This can help explain more generally the link between how judges think, how they talk, and how they rule. For instance, we consider the role of analogical reasoning in shaping the joint development of different rules. Legal analogies foster adherence to precedent across, as well as within, areas of the law. Thus they make the evolution of all affected rules more predictable and consistent over time. Counterbalancing this benefit, however, is the potential introduction of a long-run bias, as rules tend to be suboptimal compromises between what would be efficient in one area and what would be in another.

We examine and reject competing hypotheses about the constraints imposed by binding precedent. In the spirit of legal formalism, distinguishing might not allow judges to reconsider those empirical dimensions already considered by their predecessors, but only to condition the rule on new ones. Given that judges can introduce dimensions unrelated to economic efficiency, in the long run case law would then establish inefficient degenerate rules. At the opposite extreme, if any arbitrary departure from precedent were feasible at the same

effort cost, distinguishing would be functionally identical to overruling. Case law would then achieve the first best as soon as complete information became available; but it would otherwise exhibit sharp variability, yet no evolutionary tendency. The starkly counterfactual implications of these alternative assumptions bear out the legal realist view of stare decisis embodied in our baseline model.

The next section describes the underlying model of legal rules, which follows Gennaioli and Shleifer’s (2007a) stylized representation of tort law governing liability for accident. Section 3 presents our model of judicial incentives deriving from costly rhetoric, and analyzes the evolution of case law with a variable cost of legal change. Section 4 considers a formalist and an extreme realist view of precedent, assuming a fixed cost of legal change. Section 5 extends the analysis to study the role of legal analogies. Section 6 concludes the paper. The proofs of propositions are in the appendix.

2 The Model of Legal Rules

There are two parties, the offender (tortfeasor) O and the victim V . The former can take precautions at a cost C that reduce the probability of accident from p_N to p_P . Normalizing to unity the harm suffered by the victim in an accident, these precautions are socially optimal if and only if $p_N - p_P > C$. Damages are so high that they induce the tortfeasor to take precautions whenever he is held liable, so the problem simplifies to the finding of liability conditional on observable empirical facts of the case. The conditional probability of accident depends on two attributes $a \in [0, 1]$ and $u \in [0, 1]$, which are independently and uniformly distributed in the population of potential cases. For simplicity,

$$p_N - p_P = \begin{cases} \underline{\Delta} & \text{for } a + u < 1 \\ \overline{\Delta} & \text{for } a + u \geq 1 \end{cases}, \quad (1)$$

where $\overline{\Delta} > C > \underline{\Delta}$ so that precautions are socially optimal if and only if $a + u \geq 1$.

The focus of the framework is on the efficiency of legal rules, understood as their ability to attach the economically appropriate legal consequence to every possible situation (a, u) . Thus, the model focuses on statistical errors, namely cases in which the law mandates an

economically inefficient allocation of liability. On the contrary, it abstracts from legal errors in applying rules to facts: the probability of such judicial mistakes is assumed to be independent of the existing rule. This makes it possible to disregard the misapplication of legal rules: to the extent it occurs, it is merely a source of random noise over which judicial law-making has no influence. Similarly, the analysis abstracts from the parties' decision to litigate and bring cases to court, assuming that litigation occurs with the same frequency regardless of the applicable legal rule.

When the first case is being reviewed by an appellate judge, the only factual issue that comes up through trial is a . Accordingly, the rule established by the judgment is summarized by a threshold A such that the tortfeasor is held liable if and only if $a \geq A$. Imperfect information implies that the rule necessarily induces statistical errors. Liability is imposed on the tortfeasor when this is inefficient (a false positive or type I error) with probability

$$o(A) = \int_A^1 \int_0^{1-a} du da = \frac{1}{2} (1 - A)^2, \quad (2)$$

while no liability is imposed although this would be efficient (a false negative or type II error) with probability

$$v(A) = \int_0^A \int_{1-a}^1 du da = \frac{1}{2} A^2. \quad (3)$$

The social welfare function attaches a cost $\lambda_O > 0$ to inefficient over-precautions and $\lambda_V > 0$ to inefficient under-precautions, and therefore the social loss induced by the rule equals

$$\begin{aligned} \Lambda &= \lambda_O o + \lambda_V v \\ \Rightarrow \Lambda(A) &= \frac{1}{2} [\lambda_O (1 - A)^2 + \lambda_V A^2]. \end{aligned} \quad (4)$$

For the sake of brevity, the relative cost of over-precautions is denoted by $\lambda \equiv \lambda_O/\lambda_V$.

It follows immediately that the optimal one-dimensional legal rule is $A^* = \lambda/(1 + \lambda)$. This reflects the asymmetry in the cost of different errors, as captured by λ : if over-precaution is a greater social concern than under-precaution ($\lambda > 1$) only a minority of tortfeasors shall be found liable ($A^* > 1/2$).

Individual judges have idiosyncratic preferences summarized by their perceived costs of

false positives and false negatives, respectively $\beta_{O,i} > 0$ and $\beta_{V,i} > 0$; the individual utility function is therefore

$$U_i = -(\beta_{O,i}o + \beta_{V,i}v). \quad (5)$$

Each judge is assumed to derive utility from the legal rule that his own decision establishes, and not from those that might be expected in the future, conditional on the further evolution of the law. This assumption is consistent with the idea that judges' primary duty and concern is the adjudication of the concrete dispute before their court, and that judicial law-making is a by-product of this adjudication process.¹ Therefore, the first judge i establishes his preferred rule

$$A = \arg \max_A U_i(A) = \frac{\beta_{O,i}}{\beta_{O,i} + \beta_{V,i}} \equiv \hat{A}_i. \quad (6)$$

All judges have the same preference intensity, normalized so that $\beta_{V,i} + \beta_{O,i} = 1$: hence their favorite one-dimensional rule \hat{A}_i fully characterizes their preferences, which can be expressed by the utility function

$$U_i(A) = -\frac{1}{2} \left[(A - \hat{A}_i)^2 + \hat{A}_i (1 - \hat{A}_i) \right]. \quad (7)$$

The population of judges includes three different types. A fraction γ of judges are unbiased, with welfare-maximizing preferences $\beta_{O,i}/\beta_{V,i} = \lambda \Leftrightarrow \hat{A}_i = A^* \equiv \lambda/(1 + \lambda)$; the remaining $(1 - \gamma)$ comprise equal shares of pro-O judges with preferences $\beta_{O,i}/\beta_{V,i} = \lambda\pi \Leftrightarrow \hat{A}_i = A_O \equiv \lambda\pi/(1 + \lambda\pi)$ and pro-V judges with preferences $\beta_{O,i}/\beta_{V,i} = \lambda/\pi \Leftrightarrow \hat{A}_i = A_V \equiv \lambda/(\lambda + \pi)$. The parameter $\pi \in [1, \infty)$ provides a measure of judicial polarization, namely of the extent of disagreement between judges with opposite biases. Under these assumptions, all judges share an aversion to all errors, and their bias consists in disagreement over the importance of the two types of errors. In particular, the model does not consider judges whose bias is so extreme that they desire the introduction of a rule that is known with certainty to be inefficient.

¹Gennaioli and Shleifer (2007a) also argue that the introduction of a forward-looking strategic motive does not qualitatively affect the results of the model.

3 Judges' Incentives and the Rhetoric of Precedent

Stare decisis requires subsequent judges to abide by the holding of the first court, but it still allows them to refine and limit inherited rules by means of distinguishing, which *Black's Law Dictionary* defines as “not[ing] a significant factual, procedural, or legal difference (in an earlier case), usu[ally] to minimize the case's precedential effect or to show that it is inapplicable” (Garner, 2004:507). This mechanism of legal evolution is personally costly for the distinguishing judge, as emphasized by Gennaioli and Shleifer (2007a).

The effort cost of innovation is rooted in the requirement for judges to provide not only a decision, but also a detailed opinion explaining the reasoning that justifies it. Calabresi (1982:175-176) notes that “the major effective control on courts stems precisely from their duty to explain what they are doing.” At the margin, the choice is either to invest in the crafting of arguments to support a decision, or to “shrink from the very result which otherwise seems good” (Llewellyn, 1960:26).

The justification must be given in terms of adherence to binding precedent. Written opinions are intended to persuade that a decision is correct and consistent with the record of previous decisions. The aim is simultaneously rationalization and legitimation of the court's action (Fisher et al., 1993). Judges need to project the outward appearance of continuity in the law to their peers, the litigants, and society at large (Frank, 1930; Douglas, 1949; Llewellyn, 1960; Fish, 1989; Solan, 1993; Posner, 1990; 1995). Moreover, judges typically believe they have a professional duty to abide by precedent, and engage in attempts to convince themselves they are doing so (Tocqueville [1835] 2000; Llewellyn, 1960; Posner, 2001; 2005b).

Reconciling a judgement with the rhetorical demands of stare decisis requires the more costly effort, the greater the effective deviation from precedent that a judge decides to bring about. This marginal trade-off is clearly perceived by “contemporary judges [who] have insisted that following precedent is not an all-or-nothing choice between blind adherence and total disregard” (Hutchinson, 2005:147). We show that this structure of judicial incentives explains the gradualism observed in the evolution of judge-made law.

Our measure of the magnitude of legal change induced by a decision is the fraction μ_t

of possible cases that are decided differently under the new and the old rule; namely the probability of the set of events (a, u) for which either the old rule assigned liability but the new one does not, or the new rule assigns liability while the old one did not. For simplicity, we assume a quadratic specification of the cost of effort:

$$k(\mu_t) = \frac{1}{2}c\mu_t^2 \text{ for } c > 0. \quad (8)$$

If only one informative dimension a is observable, legal change can only shift the threshold A_t . Hence $\mu_t = |A_t - A_{t-1}|$, leading to a one-dimensional quadratic cost function:

$$k(A_t, A_{t-1}) = \frac{1}{2}c(A_t - A_{t-1})^2 \text{ for } c > 0. \quad (9)$$

Once the second informative dimension u becomes observable, judges acquire the ability to set any two-dimensional rule, which can be described by a function $f(a)$ such that the tortfeasor is held liable if and only if $u \geq f(a)$. Thus judge-made law can create complex balancing tests based on marginal trade-offs between different factors. Arguably the most famous real-world example is Hand's Formula for the assessment of negligence liability, which was explicitly formulated as a continuous algebraic rule in *United States v. Carroll Towing Co.*, 159 F.2d 169 (2d Cir. 1947).

All judges share a strict preference for the welfare-maximizing rule $f^*(a) = 1 - a$. However, the first best is not attained immediately, because changes in the legal rule are costly. The actual development of case law can take many different paths, but both social welfare and the utility function of each judge depend only on the probabilities of false positives o_t and false negatives v_t .

These probabilities also fully define the cost of legal changes once complete information is available. When both a and u are observable, legal innovation can remove any given statistical errors without introducing others. This is what all judges want to do, so enacted changes in the rule coincide with reductions in the probability of error. For such changes, $\mu_t = (o_{t-1} + v_{t-1}) - (o_t + v_t)$. Therefore, if and only if u has become observable, the effort a

judge needs for a desirable deviation from precedent is given by

$$k(o_t, v_t, o_{t-1}, v_{t-1}) = \frac{1}{2}c(o_{t-1} - o_t + v_{t-1} - v_t)^2. \quad (10)$$

For the symmetric case $\lambda = 1$, the following result holds.

Proposition 1 *While u is unobservable, case law evolves as a first-order autoregressive Markov process, converging to the ergodic distribution $N(A^*, \sigma^2)$. The asymptotic variance (σ^2) is decreasing in the prevalence of unbiased judges ($\partial\sigma^2/\partial\gamma < 0$) and increasing in judicial polarization ($\partial\sigma^2/\partial\pi > 0$). It is always smaller than the variance of judges' preferences, and decreases in the cost of legal innovation ($\partial\sigma^2/\partial c < 0$), vanishing as the latter diverges ($\lim_{c \rightarrow \infty} \sigma^2 = 0$).*

After u becomes observable, case law converges to the first-best efficient two-dimensional rule ($f^(a) = 1 - a$) by gradually eliminating all judicial errors. The average reduction in error per ruling ($E|\Delta_t|$) is decreasing in the prevalence of unbiased judges ($\partial E|\Delta_t|/\partial\gamma < 0$) and increasing in judicial polarization ($\partial E|\Delta_t|/\partial\pi > 0$). It decreases in the cost of legal innovation ($\partial E|\Delta_t|/\partial c < 0$), vanishing as the latter diverges ($\lim_{c \rightarrow \infty} E|\Delta_t| = 0$).*

The first part of the proposition is equivalent to the reduced-form model of case law presented in Ponzetto and Fernandez (2008). As we show in that paper, the result also obtains qualitatively if judges have a forward-looking strategic motive and try to undo the changes that their successors will effect. The asymptotics are the same even if the existing rule influences the parties' incentives to litigate and therefore the opportunity for courts to change the law. Convergence is improved if less efficient rules are more likely to be litigated (as in the models of Priest [1977] and Rubin [1977]), or conversely hindered but not eliminated if more efficient rules induce more litigation (as in Landes and Posner [1979])

Judge-made law develops as a process of incremental change, where each judge marginally moves the rule inherited from precedents in the direction of his own preferences. As a consequence, the legal rule always incorporates, albeit with different weighting, the perspectives of all previous courts as well as the current one. In this evolution, judges' heterogeneous biases tend to balance one another and induce reversion to their mean preference, which coincides with the efficient one-dimensional rule (Cardozo, 1921).

The ergodic distribution is non-degenerate, so that the legal rule never settles immutably on the unbiased one-dimensional rule. Nor can it achieve first-best efficiency in the absence of information about the second informative dimension. Despite the never-ending randomness, which entails the possibility of occasional worsening of case law, the convergence of this stochastic process embodies Leoni's (1961) view of the long-run certainty of judge-made law and Posner's ([1973] 2007) hypothesis of its evolution towards greater efficiency. The ex-ante variance of legal rules decreases over time, possibly to an arbitrarily low level, thereby increasing expected social welfare.

In a context of imperfect information, and therefore of judicial disagreement, judges' polarization is harmful. Social welfare is lower the more numerous and more extreme biased judges are. A high cost of legal innovation is then desirable because it prevents each judge from having an excessive influence on the law, reducing his ability to enact his own preferences, and enforcing instead respect for a slowly evolving tradition that is greater than the individual contributions that built it (Burke, [1790] 1999).

The second part of the proposition addresses instead the case of perfect information. When all informative dimensions have become observable, the first best is feasible. Under the maintained assumptions on the distribution of judicial preferences, all judges agree that the efficient rule is preferable to all alternatives. The law will not only converge to it as a stochastic process, but will certainly reach it as an eventual unchanging end-point. This is a strong form of Posner's efficiency hypothesis, which is rooted in judges' shared commitment to efficiency, as suggested by Landes and Posner (1987).

The cost of legal innovation becomes a burden that prevents immediate adjustment to the optimum, and therefore it is welfare-reducing. Conversely, judicial polarization becomes desirable because it leads to faster innovation, and therefore to faster attainment of the first best. The reversal of the welfare impact of judicial activism when perfect information becomes available highlights the trade-off between stability and adaptability of case law.²

Our model identifies the incentives induced by the rhetoric of *stare decisis* as the mech-

²In Ponzetto and Fernandez (2008) the same trade-off emerges because of changes not in the informational environment, but in the underlying social conditions (e.g., for tort law in the cost of precautions), and therefore in the optimal rule. We show that in a dynamic stochastic setting case law and statutes are complementary in the optimal common-law system: sudden shocks are dealt with by legislation, whereas judges are engaged in the steady marginal revision of existing rules.

anism regulating this trade-off, and the driving force behind the evolution of case law. We thereby marshal analytical support for a conjecture advanced by Stone (1985:103). He suggests that an appellate judge engages in “non-stringent” or “rhetorical” reasoning when seeking a rule and reasons for it that will sincerely appeal to himself and to his “judicial and legal constituencies generally.” This practice is institutionalized by stare decisis, which therefore “drives [the judge] to seek maximum consensus also among his predecessors in time, [while] still leav[ing] a large realm for choicemaking.” At the same time as he looks for rhetorical support for his decision, the judge will choose a rule that is easier to support: thus “the need to justify promotes justifiability.” Herein lies “a built-in check on the instant judge’s individual caprice” that ultimately provides the crucial mediation between the flexibility and the certainty that common law needs to, and does, simultaneously achieve.

4 Counterfactual Models of Distinguishing

We have argued that the significance of stare decisis belongs to the realm of pragmatic rhetorical expediency. On the contrary, a long-standing tradition of legal formalism conceives of distinguishing as a process confined within strict logically defined boundaries (Goodhart, 1930). Such a view is consistent with three key assumptions made by Gennaioli and Shleifer (2007a) about the process of legal evolution.

In their model, all permissible legal innovation requires an invariant effort cost $k \geq 0$ from the judge deviating from precedent. However, only certain legal changes are feasible. Distinguishing must introduce into legal consideration a new empirical dimension b , and stare decisis only lets the judge choose two thresholds B_0 and B_1 that respect the precedent A in the sense that liability is imposed if and only if $a < A$ and $b \geq B_0$, or $a \geq A$ and $b \geq B_1$. Finally, the new dimension must provide previously unexploited economic information about the optimal allocation of liability: b must coincide with u .

The last assumption implies that in their model case law can evolve only when new relevant facts become observable to the judge. Naturally, this may be due to technological advances. But the availability of information also depends on the principle that courts have power to rule only on the facts that have come up during the trial. In the domain

of product liability, for instance, *Boyle v. United Technologies Corp.*, 487 U.S. 500 (1988), established the “government contractor defense,” which immunizes federal contractors from liability for manufacturing design defects when the government has approved reasonably precise specifications. The rule could only be formulated once the Supreme Court heard a product-liability case whose facts allowed the manufacturer to show that its design conformed to government specifications. Similarly, after strict product liability had been imposed upon manufacturers and retailers, the Supreme Court of Illinois declined to impose it on sellers of used products in *Peterson v. Lou Bachrodt Chevrolet Co.*, 329 N.E.2d 785 (Ill. 1975). This distinction could only be introduced in adjudicating a suit filed against a used-car dealership.

In practice, however, 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 (Llewellyn, 1930; Stone, 1946; 1964). Many legal distinctions are at best dubiously grounded in objective efficiency considerations. E.g., recent decisions such as *Rousseau v. K.N. Construction, Inc.*, 727 A.2nd 190 (RI 1999), established that the rules governing tort liability in construction cases are different for commercial plaintiffs and for homeowners.³ Starker and more important, manufacturers’ strict liability to the consumer was gradually introduced through a series of rulings that created exceptions for specific products such as soap, hair dye, dog food, and fish food.⁴

To capture this phenomenon, we need to recognize that a judge can distinguish on the grounds of a dimension that is independent of both a and u , and therefore contains no statistical information: $b \sim U[0, 1]$. Then the resulting rule determines errors

$$o(A; B_0, B_1) = \frac{1}{2} [(1 - B_0) A (2 - A) + (1 - B_1) (1 - A)^2] \quad (11)$$

³Niblett et al. (Forthcoming) outline a broader group of “idiosyncratic exceptions” to the general liability rule in this domain.

⁴Respectively *Kruper v. Procter & Gamble Co.*, 113 N.E.2d 605 (Ohio App. 1953); *Graham v. Bottenfield’s, Inc.*, 176 Kan. 68, 269 P.2d 413 (1954); *McAfee v. Cargill, Inc.*, 121 F. Supp. 5 (S.D. Cal. 1954); *Midwest Game Co. v. M.F.A. Milling Co.*, 320 S.W.2d 547 (Mo. 1959). Prosser (1960) provides a classic account of the erosion of the requirement of privity in implied warranty actions.

and

$$v(A; B_0, B_1) = \frac{1}{2} [B_0 A^2 + B_1 (1 - A^2)], \quad (12)$$

under the maintained assumption that distinguishing sets a pair of thresholds B_0 and B_1 for dimension b , conditional on the existing threshold A for dimension a .

The assumption of so rigid a constraint on distinguishing makes it a blunt instrument of change. The new dimension b cannot be used to overrule precedent covertly and obtain an outcome equivalent to an arbitrary shift in A . All that can be achieved is the equivalent of the most extreme shift. By setting $B_0 = B_1 = 0$, universal liability is established: the equivalent of $A = 0$. By setting $B_0 = B_1 = 1$, liability is completely eliminated: the equivalent of $A = 1$. Any other use of the uninformative dimension would merely add unnecessary randomness to the legal rule, and no judge wishes to introduce such noise.

Sufficiently biased judges prefer the extreme rules to the original rule set by a judge with the opposite bias, or even to the efficient one-dimensional rule. All-or-nothing rules, however, are the least precise and include the least efficient possible. Hence case law would eventually collapse if judges had both idiosyncratic biases and the power to change the law through uninformative distinctions, and yet were bound by this strictly formalistic notion of respect for precedent. In an infinite-horizon framework, the following result obtains.

Proposition 2 *Suppose that for a fixed cost $k \geq 0$ a judge can introduce a new dimension b independent of a and u , and establish a rule that imposes liability if and only if $a < A$ and $b \geq B_0$, or $a \geq A$ and $b \geq B_1$, for $B_0, B_1 \in [0, 1]$.*

There exists a value $\bar{k} > 0$ such that for any cost $k \in [0, \bar{k})$ if polarization is greater than a finite threshold $\bar{\pi}(k)$ then an uninformative dimension is introduced with probability one in the long run, and the legal rule becomes either universal liability or no liability, discarding all available information.

No matter what the initial rule is, given a sufficiently low but positive cost of distinguishing k and a sufficiently high but finite level of polarization π , there are judges who prefer extreme rules. Over time, all types of judge almost surely adjudicate a case, and therefore have the opportunity to introduce an uninformative dimension b . Hence it will eventually be introduced, even if this reduces social welfare. Then the absolute respect for

precedent on the dimensions that were part of the ratio decidendi of a previous decision prevents any subsequent judge from undoing the damage and recovering the legal relevance of the observable informative dimension a .⁵ In the long run, case law becomes more extreme than the preferences of any judge, because the straitjacket of precedent exacerbates judicial extremism.

This bleak prediction runs counter to empirical observation. The evolution of case law produces rules that are typically far from extreme, and that make use of available information, albeit without necessarily achieving perfectly efficient outcomes. A growing body of econometric evidence shows the country-level correlation between the importance of judge-made law and economic success (Beck et al., 2003; 2005; La Porta et al., 2004; La Porta et al., 2008). Proposition 2 thus highlights the counterfactual implications of a rigid formalist view of stare decisis.

The inadequacy of such a conception is equally borne out by the practice of common law. Judges are clearly able to escape the grip of theoretically binding precedents without explicitly overruling them. Famous judicial decisions attain their landmark status because, with hindsight, they are seen as turning points. Yet the judges writing these decisions typically stress their consistency with stare decisis. Product liability in negligence to a remote seller is now considered to originate from *MacPherson v. Buick Motor Co.*, 111 N.E. 1050 (N.Y. 1916). Cardozo’s decision, however, does not highlight innovation but continuity with the principle of *Thomas v. Winchester*, 6 N.Y. 397 (1852). The further shift from negligence to strict liability is commonly associated with *Escola v. Coca-Cola Bottling Co.*, 150 P.2d 436 (Cal. 1944), whose decision ostensibly applies the ancient doctrine of *res ipsa loquitur*. *Greenman v. Yuba Power Products, Inc.*, 377 P.2d 897 (Cal. 1962) eliminates the requirement of privity for implied warranties by appealing to a tradition of liability for unwholesome food products dating back to the Middle Ages.

Jurisprudence has long recognized that “distinguishing a precedent to death” (Posner, 1996:373) is common and can take many forms (Douglas, 1949; Llewellyn, 1960; Stone, 1946;

⁵Even observability of the second material dimension u would not change the result. If it becomes observable after uninformative distinguishing has occurred, the first dimension a has already been lost and cannot be recovered: the same will eventually happen to the second. If a two-dimensional rule is established before uninformative distinguishing, the result still obtains qualitatively, although the thresholds \bar{k} and $\pi(k)$ are more stringent, because biased judges are destroying a more precise two-dimensional rule.

1964; 1985; Summers and Eng, 1997). It is possible because the judge who decides a case cannot fix unambiguously its ratio decidendi; instead, the rule he had the power to establish is determined by later courts (Cardozo, 1921; Allen, 1927; Radin, 1933; Cohen, 1935; Frank, 1949; Montrose, 1957; Llewellyn, 1960; Dias, 1985; Posner, 1990; Garner, 2004). As a consequence, distinguishing can confine the authority of a precedent to its particular facts, however narrowly construed (Llewellyn, 1930; Stone, 1964; Cross and Harris, 1991).

Formally, when a judge introduces a new dimension, the resulting increase in the dimensionality of the problem allows him to claim that the previous rule applies to an arbitrarily small portion of the infinitely larger space that he is now mapping. He can condition his modified rule not only on the two categories established for the previous dimension by existing law, but also on a finer new partition of his own making. Rather than merely choosing two thresholds (B_0, B_1) given A , judges acquire the ability to create continuously variable rules $B(a)$ such that the tortfeasor is held liable if and only if $u \geq B(a)$.

This implies that distinguishing can achieve the same outcomes as overruling. The additional assumption that all feasible legal innovation requires the same effort cost $k \geq 0$ yields the stronger implication that distinguishing and overruling are identical. This radical interpretation is consistent with the perspective of the critical legal studies movement, whose exponents denounce the arbitrariness of judicial law-making (Unger, 1986; Fish, 1989; Kennedy, 1997).

Gennaioli and Shleifer (2007b) present the model of overruling when only one informative dimension a is observable. Any change in the law implements exactly the preferences of the single judge effecting it. For a sufficiently low (but positive) cost k and a sufficiently high (but finite) level of polarization π , the legal rule fluctuates incessantly between the bliss points of different judges. This volatility has no welfare effect, since it is equivalent ex ante for the law to be chosen either by a different random judge each period, or by a single random judge for all periods.

We can additionally show that as soon as the second informative dimension u becomes observable, its introduction achieves first-best efficiency.

Proposition 3 *Suppose that the second informative dimension u becomes observable, and let any judge who changes the existing legal rule incur a fixed cost $k \geq 0$.*

There exists a value $\bar{k} > 0$ such that for any cost $k \in [0, \bar{k})$, if polarization is greater than a finite threshold $\bar{\pi}(k)$ then u is introduced with probability one in the long run, and the first-best legal rule is established.

When both informative dimensions are observable, the first best can be implemented by the optimal two-dimensional rule $B(a) = 1 - a$. Moreover, this is the two-dimensional rule that all judges want to implement, because in the Gennaioli–Shleifer framework judges agree on the goal of efficiency. Disagreement and biases persist only so long as ignorance does: when u is unobservable, there is a trade-off between reducing false positives and false negatives, and different judges have different preferences in this regard. Once u is observed, the first-best rule become feasible, and it is strictly preferred by every judge.

Legal change is efficiency-increasing in expectation. Changes based on uninformative dimensions have no expected impact on social welfare, while those based on the second informative dimension maximize efficiency. As in proposition 1, judicial polarization is good in the long run because it provides sharper incentives for judges to incur the private cost of distinguishing and achieve the benefits (both private and social) of a fully efficient rule. On the other hand, judicial polarization is detrimental in the short run, before the second informative dimension becomes observable, because it leads to more biased one-dimensional rules being established some fraction of the time by biased judges (Gennaioli and Shleifer, 2007b).

The pattern of legal change described by proposition 3, however, does not provide a realistic picture of the evolution of common law. Under the assumption of a fixed cost of innovation, the law jumps from one rule to another, completely unrelated to the one it replaces. Efficiency only increases through one sudden jump to the first best, which is exogenously triggered by the arrival of perfect information.

In reality, instead, “gradual or incremental change is the dominant form of change in a decentralized system of judge-made law” (Landes and Posner, 1979:270). Such a mechanism of careful and predictable marginal innovation drives the welfare-increasing development of case law that the legal realist tradition emphasizes (Holmes, 1897; Cardozo, 1921; [1932] 1947; Radin, 1925; Frank, 1930; Llewellyn, 1930; 1960; Posner, [1973] 2007). Proposition 1 accounts for this process by recognizing that all legal innovation is not equally costly, and

that each judge faces a marginal trade-off between bringing legal rules into closer alignment with his own preferences and having to expend personal effort in order to do so.

The history of the law of product liability reflects this incrementalism. “[T]he last half of the nineteenth century witnessed a steady, but limited, erosion of [the] privity limitation as exceptions were created” (Epstein, 2004:651). The decision in *MacPherson v. Buick* surveys and builds upon this pre-existing trend in rulings. The privity requirement was gradually eroded in the twentieth century, from *MacPherson v. Buick* through *Escola v. Coca-Cola* to the Restatement (Second) of Torts (Prosser, 1960; 1966). Advocacy of strict liability was a minority view in 1944, but judicial opinion gradually shifted in its favour until, by 1965, it had prevailed.

This pattern disproves what Cross and Harris (1991:52) characterize as “the extreme realist position [...] that our judges are capable of the grossest hypocrisy,” a view that fully equates distinguishing with overruling.⁶ In reality, the fundamental difference between the two is that distinguishing does not immediately efface precedent but steadily erodes it (Cardozo, 1921; Douglas 1949; Summers 1997). Rather than being the sudden and discontinuous choice of a single judge, “developments in the [legal] landscape come about bit by bit, as a result of the actions of many courts” (Calabresi, 1982:224).

Our analysis thus bears out formally Posner’s (2008:230) rejection of both legal formalism and extreme realism as “inadequately descriptive of judicial behavior.” The rhetoric of *stare decisis*, far from being empty, constitutes its substance. What might seem a paradox of legal theory admits a natural economic interpretation, because rhetoric is a costly activity that judges economize on, even at the expense of other goals.

5 Legal Analogies and Judge-made Law

The evolution of case law derives from nuanced rhetorical incentives rather than from clear-cut logical boundaries. The costs faced by judges in explaining and justifying their decisions

⁶A model of abrupt changes in the law based on a fixed cost of legal innovation provides a better representation of statute-writing by a legislature than of decision-making by appellate judges (Ponzetto and Fernandez, 2008). Pound (1913) and Hayek (1973) contend that in common law legislation is necessary to obtain rapid adaptation of legal rules, which is alien to the nature of case law.

thus shape both the evolution and the structure of the law. Our model can shed some light on the latter as well. We consider the use of analogical reasoning, which is widely recognized as the characteristic mode of legal reasoning (Levi, 1949; Raz, 1979; Posner, 1990; 2008; Brewer, 1996; Weinreb, 2005). Like respect for precedent, legal analogy is a rhetorical device that confers legitimacy to courts' role as legislators and at the same time induces "cautious, incremental judicial legislating" (Posner, 2006:773).

We can interpret analogical reasoning as a form of stare decisis that refers to precedents in factually different but conceptually related areas of law. To take a case that has been repeatedly cited, *Adams v. New Jersey Steamboat Co.*, 45 N.E. 369 (N.Y. 1896), shows that the relationship of a transport operator to its passengers was framed by the court as analogous to that of an innkeeper to his guests. The decision thus finds liability for a steamboat company where none existed for a railroad, arguing that "a steamer ... is, for all practical purposes, a floating inn" and referring to precedents on the rule of responsibility applicable to innkeepers.

A full-fledged model of the conceptual categories competing for judges's attention, and of the resulting cost of justification of their decisions, remains beyond the scope of this paper. We simply assume that two areas of law have been co-categorized in legal discourse, and that judges must exert costly effort to justify deviations both from the existing rule governing the matter before their court, and from precedent in the co-categorized area. Given existing rules (A_{t-1}, Z_{t-1}) governing the two issues, the cost function for a decision changing the rule on issue A becomes

$$k(A_t, A_{t-1}, Z_{t-1}) = \frac{c}{2} \left[\left(1 - \frac{\alpha}{2}\right) (A_t - A_{t-1})^2 + \frac{\alpha}{2} (A_t - Z_{t-1})^2 \right], \quad (13)$$

where the parameter $\alpha \in [0, 1]$ provides a measure of the power of analogical thinking. The cost of changes in the issue- Z rule is defined symmetrically.

Assuming that the distribution of judicial preferences has no average bias ($E(\hat{A}_j) = A^*$ and $E(\hat{Z}_j) = Z^*$) and the same finite variance on both issues ($Var(\hat{A}_j) = Var(\hat{Z}_j) < \infty$), we can prove the following result.

Proposition 4 *Let two areas of the law (A and Z) be co-categorized by analogical reasoning*

($\alpha > 0$). Then case law in the two areas evolves as a first-order vector autoregressive Markov process, converging to an ergodic joint normal distribution with asymptotic moments $E(A) = A^* + \delta(Z^* - A^*)$, $E(Z) = Z^* + \delta(A^* - Z^*)$, $Var(A) = Var(Z) = \theta\sigma^2$, and $Cov(A, Z) = \rho\theta\sigma^2$.

Legal analogies introduce positive correlation in the two rules ($\rho > 0$) and reduce the variance of each rule ($\theta < 1$). If the two efficient rules are different, legal analogies introduce long-run bias in the law ($\delta \in (0, 1/2)$).

All three effects are increasing in the power of analogical thinking ($\partial\rho/\partial\alpha > 0$, $\partial\theta/\partial\alpha < 0$ and $\partial\delta/\partial\alpha > 0$) and in the cost of legal innovation ($\partial\rho/\partial c > 0$, $\partial\theta/\partial c < 0$ and $\partial\delta/\partial c > 0$).

Analogical reasoning creates not only a rhetorical, but also a substantive link between two economically distinct areas of the law. The evolution of each can no longer be analyzed independently from the other, and intuitively they tend to move in tandem ($\rho > 0$).

Moreover, the proposition identifies a cost and a benefit of legal analogy. The rhetorical connection between distinct areas generates a bias distorting the law away from efficiency in the long run, unless the two areas are economically identical despite their legal differences ($\delta > 0$). In their effort to achieve rhetorical consistency, judges bring extraneous considerations to bear on either issue. Both rules become suboptimal compromises between what would be efficient in one area and what would in the other.

On the other hand, co-categorization makes each rule more predictable and consistent over time ($\theta < 1$), much as stare decisis does (Raz, 1979). The mechanism is the same underpinning proposition 1. The need for rhetorical justification induces judges to write decisions consistent with the consensus of their peers. Analogical reasoning magnifies the beneficial effect of this process by extending the group of peers whose consensus is sought, which comes to include all judges who ruled on the related area as well as on the one currently before the court.

Because of this trade off, the efficiency properties of analogical reasoning cannot be assessed unambiguously. Consistent with Posner's (1990; 2006; 2008) analysis, the key is the connection between analogy and policy relevance. The normative implications of our model thus concern the higher and more complex level of the formation of categories of analogy. There are significant benefits to be reaped from analogical connections between areas of the

law that are similar from the point of view of efficiency, but also potentially large costs from economically pointless co-categorization.⁷

6 Conclusion

How can the principle of stare decisis make common law stable and certain, but simultaneously flexible and constantly evolving? The solution to this puzzle has escaped all attempts to construct a formalist definition of binding precedent. Consistent with the insights of pragmatism and of legal realism, we have given instead an economic answer based on judicial incentives.

Appellate judges have idiosyncratic policy preferences but must exert costly effort to change the law. Any departure from precedent is logically possible, but we focus on the rhetorical requirements imposed by stare decisis. Common-law judges must support their decisions with an opinion presenting them as consistent and continuous with those of their predecessors. The persuasive effort necessary to hide or justify departures from precedent is increasing in the extent of the innovation effectively introduced. This trade off puts each judge at the margin of the balance between the stability and the evolution of the law. The rhetoric of stare decisis is its substance.

More extreme models generate counterfactual empirical predictions. If distinguishing did not allow judges to review empirical dimensions already considered by precedent, the ability to introduce economically irrelevant dimensions would eventually lead case law to establish with certainty extreme, inefficient rules. If distinguishing were effectively identical to overruling, the law would achieve the first best as soon as judges acquired perfect information, but would otherwise be sharply volatile and lack any evolutionary properties.

We have proposed instead a legal realist model in which distinguishing can be used to overrule, but is kept in check by an increasing effort cost of legal innovation. Thus case law develops gradually, and reflects the preferences of all past judges as well as the instant

⁷In general, for any two legal issues A and Z , given any finite difference in the respective efficient rules ($|A^* - Z^*| \in (0, \infty)$), long-run social welfare is increased by an arbitrary small analogical connection ($\alpha \simeq 0$) between the two. Starting from no connection, the marginal gain from increased consistency is of a higher order than the marginal loss from increased bias.

one. Its evolution converges towards greater efficiency and predictability, according to the intuitions of Burke, Cardozo, Leoni, and Posner. Legal change increases social welfare: in expectation when no new information becomes available; with certainty once all relevant facts are observable. Judge-made law will then eventually reach the first best.

Rhetorical requirements beyond *stare decisis* create incentives that shape judicial decisions. As a first step towards a broader analysis, we have considered the role of legal analogies, which generate adherence to precedent across, as well as within, areas of the law. They thereby increase the consistency of legal evolution at the cost of a potential long-run bias.

We have not attempted to provide an in-depth account of the social conventions underpinning judges' rhetorical incentives. It remains for further analysis of judicial decision-making to assess the role of legal education, public scrutiny, peer pressure, and the organizational structure of the judiciary in generating such incentives.

Our account does not deny the importance of logic in legal reasoning. Rather, it reverses the causal link between logic, practice and rhetoric. Judges will craft their decisions and opinions so as to facilitate the task of rationalizing and legitimizing them. It is precisely this process of justification that results in the logical structure of the law.

A Appendix

A.1 Proof of Proposition 1

When u is unobservable, judge j faced with precedent A_{t-1} sets a rule

$$\begin{aligned} A_t &= \arg \max_{A \in [0,1]} \{U_j(A) - k(A, A_{t-1})\} \\ &= \arg \min_{A \in [0,1]} \left\{ \left(A - \hat{A}_j \right)^2 + c(A - A_{t-1})^2 \right\} \\ &= \frac{c}{1+c} A_{t-1} + \frac{1}{1+c} \hat{A}_j. \end{aligned} \quad (\text{A1})$$

By the properties of an AR(1) process, judge-made law converges to the ergodic distribution

$$A \sim N \left(E \left(\hat{A}_j \right), \frac{1}{1+2c} \text{Var} \left(\hat{A}_j \right) \right), \quad (\text{A2})$$

as long as the invariant distribution of judges' preferences has finite variance.

For $\lambda = 1$, a share γ of judges have unbiased preferences $A^* = 1/2$, a share $(1 - \gamma)/2$ have pro-O preferences $A_O = \pi/(1 + \pi)$, and a share $(1 - \gamma)/2$ have pro-V preferences $A_V = 1/(1 + \pi)$. This distribution of judges' preferences has expectation

$$E \left(\hat{A}_j \right) = \frac{1}{2} = A^* \quad (\text{A3})$$

and variance

$$\begin{aligned} \text{Var} \left(\hat{A}_j \right) &= E \left(\hat{A}_j^2 \right) - \left[E \left(\hat{A}_j \right) \right]^2 \\ &= \frac{\gamma}{4} + \frac{1-\gamma}{2} \frac{\pi^2+1}{(1+\pi)^2} - \frac{1}{4} \\ &= \frac{1-\gamma}{4} \left(\frac{\pi-1}{\pi+1} \right)^2. \end{aligned} \quad (\text{A4})$$

Thus judge-made law converges to the ergodic distribution $A \sim N(A^*, \sigma^2)$ with asymptotic variance

$$\sigma^2 = \frac{1-\gamma}{4(1+2c)} \left(\frac{\pi-1}{\pi+1} \right)^2, \quad (\text{A5})$$

such that

$$\frac{\partial \sigma^2}{\partial \gamma} = -\frac{1}{4(1+2c)} \left(\frac{\pi-1}{\pi+1} \right)^2 < 0, \quad (\text{A6})$$

$$\frac{\partial \sigma^2}{\partial \pi} = \frac{1-\gamma}{(1+2c)} \frac{\pi-1}{(\pi+1)^3} > 0, \quad (\text{A7})$$

and

$$\frac{\partial \sigma^2}{\partial c} = -\frac{1-\gamma}{2(1+2c)^2} \left(\frac{\pi-1}{\pi+1} \right)^2 < 0, \quad (\text{A8})$$

with $\lim_{c \rightarrow \infty} \sigma^2 = 0$.

After u becomes observable, judge j faced with a precedent inducing errors $(o_{t-1}, v_{t-1}) \geq 0$

chooses a new rule inducing errors

$$\begin{aligned}
(o_t, v_t) &= \arg \max_{\substack{o \in [0, o_{t-1}] \\ v \in [0, v_{t-1}]}} \{U_i(o, v) - k(o, v, o_{t-1}, v_{t-1})\} \\
&= \arg \min_{\substack{o \in [0, o_{t-1}] \\ v \in [0, v_{t-1}]}} \left\{ \beta_{O,i} o + \beta_{V,i} v + \frac{1}{2} c (o_{t-1} - o + v_{t-1} - v)^2 \right\} \\
&= \arg \max_{\substack{o \in [0, o_{t-1}] \\ v \in [0, v_{t-1}]}} \left\{ \hat{A}_i (o_{t-1} - o) + (1 - \hat{A}_i) (v_{t-1} - v) - \frac{1}{2} c (o_{t-1} - o + v_{t-1} - v)^2 \right\}.
\end{aligned} \tag{A9}$$

Each judge can reduce the type of error he is most concerned with. A pro-O judge reduces only the probability of false positives, setting $o_t = o_{t-1} - A_O/c$ and $v_t = v_{t-1}$ for all $o_{t-1} \geq A_O/c$. A pro-V judge reduces only the probability of false negatives, setting $o_t = o_{t-1}$ and $v_t = v_{t-1} - (1 - A_V)/c$ for all $v_{t-1} \geq (1 - A_V)/c$. An unbiased judge reduces the probability of any type of error so that $o_t + v_t = o_{t-1} + v_{t-1} - 1/(2c)$.⁸

The average reduction in error per ruling is

$$E|\Delta_t| \equiv E(o_{t-1} - o_t + v_{t-1} - v_t) = \frac{1}{c} \left[\frac{\gamma}{2} + (1 - \gamma) \frac{\pi}{\pi + 1} \right], \tag{A10}$$

such that

$$\frac{\partial E|\Delta_t|}{\partial \gamma} = -\frac{1}{2c} \frac{\pi - 1}{\pi + 1} < 0 \tag{A11}$$

and

$$\frac{\partial E|\Delta_t|}{\partial \pi} = \frac{1 - \gamma}{c(\pi + 1)^2} > 0. \tag{A12}$$

If all false positives have been eliminated, then all judges reduce the probability of false negatives by an amount $(1 - \hat{A}_j)/c$. If all false negatives have been eliminated, then all judges reduce the probability of false positives by an amount \hat{A}_j/c . Thus the average reduction in error per ruling becomes

$$E|\Delta_t| = \frac{1}{2c} \text{ such that } \frac{\partial E|\Delta_t|}{\partial \gamma} = \frac{\partial E|\Delta_t|}{\partial \pi} = 0 \tag{A13}$$

once either type of error has been eliminated.

A.2 Proof of Proposition 2

Judge j 's utility from the rule A, B_0, B_1 is

$$\begin{aligned}
U_j(A, B_0, B_1) &= -\hat{A}_j o(A; B_0, B_1) - (1 - \hat{A}_j) v(A; B_0, B_1) \\
&= -\frac{1}{2} \left[\hat{A}_j + A (A - 2\hat{A}_j) B_0 + (1 - A) (1 + A - 2\hat{A}_j) B_1 \right],
\end{aligned} \tag{A14}$$

which is maximized by:

1. universal liability ($B_0 = B_1 = 0$) if $\hat{A}_j < A/2$;

⁸We do not explicitly compute the single ruling that sets the probability of either type of error to zero.

2. the existing precedent ($B_0 = 1, B_1 = 0$) if $A/2 < \hat{A}_j < (1 + A)/2$;
3. no liability ($B_0 = B_1 = 1$) if $\hat{A}_j > (1 + A)/2$.

In the knife-edge cases $\hat{A}_j = A/2$ or $\hat{A}_j = (1 + A)/2$ the judge is indifferent between all values respectively of B_0 and of B_1 . Given a cost of change $k > 0$, this indifference is always resolved in favour of the preservation of the status quo.

In general, when only dimension a is observable and the legal rule is therefore A , a judge j replaces it with universal liability ($B_0 = B_1 = 0$) if and only if

$$\hat{A}_j < \frac{1}{2} - \frac{1 - A}{2} - \frac{k}{A} \equiv \underline{\Omega}(A), \quad (\text{A15})$$

or with no liability ($B_0 = B_1 = 1$) if and only if

$$\hat{A}_j > \frac{1}{2} + \frac{A}{2} + \frac{k}{1 - A} \equiv \bar{\Omega}(A). \quad (\text{A16})$$

Thus a pro-O judge with preferred rule $A_O = \lambda\pi/(1 + \lambda\pi)$ introduces an uninformative dimension when $A_O > \bar{\Omega}(A_{t-1})$, and a pro-V judge with preferred rule $A_V = \lambda/(\lambda + \pi)$ introduces an uninformative dimension when $A_V < \underline{\Omega}(A_{t-1})$.

If $A_O > \bar{\Omega}(A^*)$ and $A_V < \underline{\Omega}(A_O)$ then a pro-O judge changes the efficient rule A^* , and a fortiori a pro-V rule A_V ; while a pro-V judge changes a pro-O rule A_O . For all $k < (1 - A^*)^2/2 \leq 1/2$ these conditions hold strictly as $\pi \rightarrow \infty$ (which implies $A_O = 1$ and $A_V = 0$) and by continuity for sufficiently large finite values of π .

If $A_O > \bar{\Omega}(A_V)$ and $A_V < \underline{\Omega}(A^*)$ then a pro-V judge changes the efficient rule A^* , and a fortiori a pro-O rule A_O or the no liability rule; while a pro-O judge changes a pro-V rule A_V or the universal liability rule. For all $k < (A^*)^2/2 \leq 1/2$ these conditions hold strictly as $\pi \rightarrow \infty$ and by continuity for sufficiently large finite values of π .

Over an infinite horizon, all types of judge decide a case at least once with probability one, and therefore an extremist rule is eventually established if $k < (\max\{A^*, 1 - A^*\})^2/2$ and π is sufficiently high.

A.3 Proof of Proposition 3

Distinguishing to the first-best rule happens if the second informative dimension u is observable and a judge with preferences \hat{A}_j faces a one-dimensional precedent A such that

$$2k \leq \hat{A}_j(1 - A)^2 + (1 - \hat{A}_j)A^2 = \hat{A}_j - 2A\hat{A}_j + A^2. \quad (\text{A17})$$

The bluntest incentive is to change the unbiased one-dimensional rule A^* , and:

1. if $A^* < 1/2$ the unbiased one-dimensional rule A^* is replaced with the first-best two-dimensional rule by a pro-O judge, provided that $k < (1 - A^*)^2/2$ and $\hat{A}_O > [2k - (A^*)^2]/(1 - 2A^*)$;
2. if $A^* = 1/2$ the unbiased one-dimensional rule A^* is replaced with the first-best two-dimensional rule by any judge, provided that $k < 1/8$;

3. if $A^* > 1/2$ the unbiased one-dimensional rule A^* is replaced with the first-best two-dimensional rule by a pro-V judge, provided that $k < (A^*)^2/2$ and $\hat{A}_V < [(A^*)^2 - 2k] / (2A^* - 1)$.

Over an infinite horizon, all types of judge decide a case at least once with probability one after the second informative dimension has become observable; therefore the first-best rule is adopted in the long run provided that either

$$\lambda \in (0, 1) \text{ and } k < \frac{1}{2(1+\lambda)^2} \text{ and } \pi \geq \frac{2(1+\lambda)^2 k - \lambda^2}{\lambda [1 - 2(1+\lambda)^2 k]}, \quad (\text{A18})$$

or

$$\lambda = 1 \text{ and } k < \frac{1}{8} \text{ and } \pi \geq 1, \quad (\text{A19})$$

or

$$\lambda \in (1, \infty) \text{ and } k < \frac{\lambda^2}{2(1+\lambda)^2} \text{ and } \pi \geq \frac{\lambda [2(\lambda+1)^2 k - 1]}{\lambda^2 - 2(\lambda+1)^2 k}. \quad (\text{A20})$$

A.4 Proof of Proposition 4

Given precedents (A_{t-1}, Z_{t-1}) , a judge a whose decision at time t sets the new rule on issue A sets

$$\begin{aligned} A_t &= \arg \max_{A \in [0,1]} \{U_a(A) - k(A, A_{t-1}, Z_{t-1})\} \\ &= \arg \min_{A \in [0,1]} \left\{ \left(A - \hat{A}_a \right)^2 + c \left[\left(1 - \frac{\alpha}{2} \right) (A - A_{t-1})^2 + \frac{\alpha}{2} (A - Z_{t-1})^2 \right] \right\} \\ &= \frac{c}{1+c} \left[\left(1 - \frac{\alpha}{2} \right) A_{t-1} + \frac{\alpha}{2} Z_{t-1} \right] + \frac{1}{1+c} \hat{A}_a. \end{aligned} \quad (\text{A21})$$

Simultaneously, an independently drawn judge z whose decision determines legal evolution on issue Z sets

$$\begin{aligned} Z_t &= \arg \max_{Z \in [0,1]} \{U_z(Z) - k(Z, Z_{t-1}, A_{t-1})\} \\ &= \frac{c}{1+c} \left[\left(1 - \frac{\alpha}{2} \right) Z_{t-1} + \frac{\alpha}{2} A_{t-1} \right] + \frac{1}{1+c} \hat{Z}_z. \end{aligned} \quad (\text{A22})$$

Co-categorized areas of the law thus evolve jointly as the VAR(1) process

$$\begin{cases} A_t = \frac{c}{1+c} \left[\left(1 - \frac{\alpha}{2} \right) A_{t-1} + \frac{\alpha}{2} Z_{t-1} \right] + \frac{1}{1+c} \hat{A}_a \\ Z_t = \frac{c}{1+c} \left[\left(1 - \frac{\alpha}{2} \right) Z_{t-1} + \frac{\alpha}{2} A_{t-1} \right] + \frac{1}{1+c} \hat{Z}_z \end{cases}, \quad (\text{A23})$$

which converges to an ergodic normal distribution with expectation

$$\begin{cases} E(A) = A^* + \frac{1}{2} \frac{\alpha c}{1+\alpha c} (Z^* - A^*) \\ E(Z) = Z^* + \frac{1}{2} \frac{\alpha c}{1+\alpha c} (A^* - Z^*) \end{cases} \quad (\text{A24})$$

and second moments

$$\begin{cases} Var(A) = Var(Z) = \frac{1}{2} \sigma^2 \left[1 + \frac{1+2c}{(1+2c-\alpha c)(1+\alpha c)} \right] \\ Cov(A, Z) = \frac{1}{2} \sigma^2 \left[1 - \frac{1+2c}{(1+2c-\alpha c)(1+\alpha c)} \right] \end{cases}, \quad (\text{A25})$$

where σ^2 is the asymptotic variance of each rule when $\alpha = 0$.

The correlation coefficient is

$$\rho \equiv \frac{Cov(A, Z)}{\sqrt{Var(A)Var(Z)}} = \frac{(2 - \alpha) \alpha c^2}{2(1 + 2c) + (2 - \alpha) \alpha c^2}, \quad (\text{A26})$$

such that

$$\frac{\partial \rho}{\partial \alpha} = \frac{4(1 - \alpha)(1 + 2c)c^2}{[2(1 + 2c) + (2 - \alpha)\alpha c^2]^2} > 0 \quad (\text{A27})$$

and

$$\frac{\partial \rho}{\partial c} = \frac{4(2 - \alpha)\alpha c(1 + c)}{[2(1 + 2c) + (2 - \alpha)\alpha c^2]^2} > 0. \quad (\text{A28})$$

The long-run bias is captured by

$$\delta \equiv \frac{E(A) - A^*}{Z^* - A^*} = \frac{E(Z) - Z^*}{A^* - Z^*} = \frac{1}{2} \frac{\alpha c}{1 + \alpha c}, \quad (\text{A29})$$

such that

$$\frac{\partial \delta}{\partial \alpha} = \frac{c}{2(1 + \alpha c)^2} > 0 \quad (\text{A30})$$

and

$$\frac{\partial \delta}{\partial c} = \frac{\alpha}{2(1 + \alpha c)^2} > 0. \quad (\text{A31})$$

The impact of analogical thinking on the asymptotic variance is captured by

$$\theta \equiv \frac{Var(A)}{\sigma^2} = \frac{Var(Z)}{\sigma^2} = \frac{1}{2} \left[1 + \frac{1 + 2c}{1 + 2c + (2 - \alpha)\alpha c^2} \right], \quad (\text{A32})$$

such that

$$\frac{\partial \theta}{\partial \alpha} = -\frac{(1 - \alpha)(1 + 2c)c^2}{[1 + 2c + (2 - \alpha)\alpha c^2]^2} < 0 \quad (\text{A33})$$

and

$$\frac{\partial \theta}{\partial c} = -\frac{(2 - \alpha)\alpha c(1 + c)}{[1 + 2c + (2 - \alpha)\alpha c^2]^2} < 0. \quad (\text{A34})$$

References

- Allen, Carleton K. 1927. *Law in the Making*. Oxford: Clarendon Press.
- Beck, Thorsten, Asli Demirgüç-Kunt, and Ross Levine. 2003. “Law and Finance: Why Does Legal Origin Matter?,” 31 *Journal of Comparative Economics* 653-675.
- . 2005. “Law and Firms’ Access to Finance,” 7 *American Law and Economic Review* 211-252.
- Brewer, Scott. 1996. “Exemplary Reasoning: Semantics, Pragmatics, and the Rational Force of Legal Argument by Analogy,” 99 *Harvard Law Review* 923-1028.
- Burke, Edmund. [1790] 1999. *Reflections on the Revolution in France*, in *Select Works of Edmund Burke*, vol. 2. Indianapolis: Liberty Fund.
- Calabresi, Guido. 1982. *A Common Law for the Age of Statutes*. Cambridge, Mass.: Harvard University Press.
- Cardozo, Benjamin N. 1921. *The Nature of the Judicial Process*. New Haven, Conn.: Yale University Press.
- . [1932] 1947. “Jurisprudence,” in M.E. Hall, ed., *Selected Writings of Benjamin Nathan Cardozo*. Reprint. New York: Fallon Publications.
- Cohen, Felix S. 1935. “Transcendental Nonsense and the Functional Approach,” 35 *Columbia Law Review* 809-849.
- Cooter, Robert, Lewis Kornhauser, and David Lane. 1979. “Liability Rules, Limited Information, and the Role of Precedent,” 10 *Bell Journal of Economics* 366-373.
- Cross, Rupert, and J. W. Harris. 1991. *Precedent in English Law*. Oxford: Clarendon Press.
- Dewey, John. 1924. “Logical Method and Law,” 10 *The Cornell Law Quarterly* 17-27.
- Dias, Reginald W. M. 1985. *Jurisprudence*. London: Butterworths.
- Douglas, William O. 1949. “*Stare Decisis*,” 49 *Columbia Law Review* 735-758.
- Epstein, Richard A. 2004. *Cases and Materials on Torts*. 8th edition. New York: Aspen.
- Fish, Stanley. 1989. *Doing What Comes Naturally: Change, Rhetoric, and the Practice of Theory in Literary and Legal Studies*. Durham, N.C.: Duke University Press.

- Fisher, William M., Morton J. Horwitz, and Thomas A. Reed. 1993. "Legal Reasoning," in W. M. Fisher, M. J. Horwitz, and T. A. Reed, eds., *American Legal Realism*. New York: Oxford University Press.
- Frank, Jerome. 1930. *Law and the Modern Mind*. New York: Brentano's.
- . 1949. *Courts on Trial*. Princeton, N.J.: Princeton University Press.
- Garner, Bryan A., ed. 2004. *Black's Law Dictionary*. 8th edition. St. Paul, Minn.: West Group.
- Gennaioli, Nicola, and Andrei Shleifer. 2007a. "The Evolution of Common Law," 115 *Journal of Political Economy* 43-68.
- . 2007b. "Overruling and the Instability of Law," 35 *Journal of Comparative Economics* 309-328.
- Goodhart, Arthur L. 1930. "Determining the Ratio Decidendi of a Case," 40 *Yale Law Journal Review* 161-183.
- Hayek, Friedrich A. 1973. *Law, Legislation and Liberty, Volume 1: Rules and Order*. Chicago, Ill.: The University of Chicago Press.
- Holmes, Oliver Wendell, Jr. 1881. *The Common Law*. Boston, Mass.: Little, Brown & Co.
- . 1897. "The Path of the Law," 10 *Harvard Law Review* 457-478.
- . 1899. "Law in Science and Science in Law," 12 *Harvard Law Review* 443-463.
- Hutchinson, Allan C. 2005. *Evolution and the Common Law*. New York: Cambridge University Press.
- Kennedy, Duncan. 1997. *A Critique of Adjudication: Fin de Siècle*. Cambridge, Mass.: Harvard University Press.
- Klein, David. 2002. *Making Law in the United States Courts of Appeals*. Cambridge: Cambridge University Press.
- La Porta, Rafael, Florencio Lopez-de-Silanes, Cristian Pop-Eleches, and Andrei Shleifer. 2004. "Judicial Checks and Balances," 112 *Journal of Political Economy* 445-470.

- La Porta, Rafael, Florencio Lopez-de-Silanes, and Andrei Shleifer. 2008. "The Economic Consequences of Legal Origins," 46 *Journal of Economic Literature* 285-332.
- La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer, and Robert W. Vishny. 1998. "Law and Finance," 106 *Journal of Political Economy* 1113-1155.
- Landes, William M., and Richard A. Posner. 1979. "Adjudication as a Private Good," 8 *Journal of Legal Studies* 235-284.
- . 1987. *The Economic Structure of Tort Law*. Cambridge, Mass.: Harvard University Press.
- Leoni, Bruno. 1961. *Freedom and the Law*. Princeton, N.J.: Van Nostrand.
- Levi, Edward H. 1949. *An Introduction to Legal Reasoning*. Chicago, Ill.: University of Chicago Press.
- Llewellyn, Karl N. 1930. *The Bramble Bush: On Our Law and its Study*. New York: Oceana Publications.
- . 1960. *The Common Law Tradition: Deciding Appeals*. Boston, Mass.: Little, Brown & Co.
- Montrose, J. L. 1957. "The *Ratio Decidendi* of a Case," 20 *The Modern Law Review* 587-595.
- Niblett, Anthony, Richard Posner, and Andrei Shleifer. Forthcoming. "The Evolution of a Legal Rule," *Journal of Legal Studies*.
- Partridge, Anthony, and William Eldridge. 1974. *The Second Circuit sentencing study: a report to the judges of the Second Circuit*. Washington, D.C.: Federal Judicial Center.
- Pinello, Daniel. 1999. "Linking Party to Judicial Ideology in American Courts: A Meta-analysis," 3 *The Justice System Journal* 219-254.
- Ponzetto, Giacomo A.M., and Patricio A. Fernandez. 2008. "Case law versus Statute Law: An Evolutionary Comparison," 37 *Journal of Legal Studies* 379-430.
- Posner, Richard A. [1973] 2007. *The Economic Analysis of Law*. [1st edition, 1973] 7th edition. New York: Aspen Publishers.

- . 1990. *The Problems of Jurisprudence*. Cambridge, Mass.: Harvard University Press.
- . 1995. *Overcoming Law*. Cambridge, Mass.: Harvard University Press.
- . 1996. *The Federal Courts: Challenge and Reform*. Cambridge, Mass.: Harvard University Press.
- . 2001. *Frontiers of Legal Theory*. Cambridge, Mass.: Harvard University Press.
- . 2005a. “Judicial Behavior and Performance: an Economic Approach,” 32 *Florida State University Law Review* 1259-1279.
- . 2005b. “The Supreme Court 2004 Term: A Political Court,” 119 *Harvard Law Review* 31-102.
- . 2006. “Book Review: Reasoning by Analogy,” 91 *Cornell Law Review* 761-774.
- . 2008. *How Judges Think*. Cambridge, Mass.: Harvard University Press.
- Pound, Roscoe. 1913. “Courts and Legislation,” 7 *The American Political Science Review* 361-383.
- Priest, George. 1977. “The Common Law Process and the Selection of Efficient Rules,” 6 *Journal of Legal Studies* 65-82.
- Prosser, William L. 1960. “The Assault upon the Citadel (Strict Liability to the Consumer),” 69 *Yale Law Journal* 1099-1148.
- . 1966. “The Fall of the Citadel (Strict Liability to the Consumer),” 50 *Minnesota Law Journal* 791-848.
- Radin, Max. 1925. “The Theory of Judicial Decisions: or How Judges Think,” 11 *American Bar Association Journal* 357-362.
- . 1933. “Case law and stare decisis: Concerning *Präjudizienrecht in Amerika*,” 199 *Columbia Law Review* 199-212.
- Raz, Joseph. 1979. *The Authority of Law: Essays on Law and Morality*. Oxford: Clarendon Press.
- Revesz, Richard. 1997. “Environmental Regulation, Ideology, and the D.C. Circuit,” 83 *Virginia Law Review* 1717-1772.

- Rowland, C. K., and Robert Carp. 1996. *Politics and Judgment in Federal District Courts*. Lawrence, Kan.: University Press of Kansas.
- Rubin, Paul. 1977. "Why is the Common Law Efficient?," 6 *Journal of Legal Studies* 51-63.
- Solan, Lawrence M. 1993. *The Language of Judges*. Chicago, Ill.: The University of Chicago Press.
- Stone, Julius. 1946. *The Province and Function of Law*. Sydney: Associated General Publications.
- . 1959. "The *Ratio* of the *Ratio Decidendi*," 22 *The Modern Law Review* 597-620.
- . 1964. *Legal Systems and Lawyers' Reasonings*. Stanford, Calif.: Stanford University Press.
- . 1969. "1966 and All That! Loosing the Chains of Precedent," 69 *Columbia Law Review* 1162-1202.
- . 1985. *Precedent and the Law: Dynamics of Common Law Growth*. Sydney: Butterworths.
- Summers, Robert. 1997. "Precedent in the United States (New York State)," in N. McCormick and R. Summers, eds., *Interpreting Precedents*. Aldershot: Ashgate Publishing.
- Summers, Robert, and Sven Eng. 1997. "Departures from Precedent," in N. McCormick and R. Summers, eds., *Interpreting Precedents*. Aldershot: Ashgate Publishing.
- Sunstein, Cass, David Schkade, and Lisa Michelle Ellman. 2004. "Ideological Voting on Federal Courts of Appeals: a Preliminary Investigation," 90 *Stanford Law Review* 301-354.
- Tocqueville, Alexis de. [1835] 2000. *Democracy in America*. Reprint. Chicago, Ill.: University of Chicago Press.
- Unger, Roberto M. 1986. *The Critical Legal Studies Movement*. Cambridge, Mass.: Harvard University Press.
- Weinreb, Lloyd L. 2005. *Legal Reason: The Use of Analogy in Legal Argument*. Cambridge: Cambridge University Press.

Wright, Robert A. 1939. *Legal Essays and Addresses*. Cambridge: Cambridge University Press.

———. 1943. “Precedents,” 8 *The Cambridge Law Journal* 118-145.

Zywicki, Todd. 2003. “The Rise and Fall of Efficiency in the Common Law: A Supply-Side Analysis,” 97 *Northwestern University Law Review* 1551-1633.