NBER WORKING PAPER SERIES

THE EVOLUTION OF PRECEDENT

Nicola Gennaioli
Andrei Shleifer

Working Paper 11265
http://www.nber.org/papers/w11265

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2005

We are grateful to Olivier Blanchard, Filipe Campante, Edward Glaeser, Claudia Goldin, Oliver Hart, Elhanan Helpman, Fausto Panunzi, Richard Posner, Ilia Rainer and especially Louis Kaplow for helpful comments. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

The Evolution of Precedent
Nicola Gennaioli and Andrei Shlefier
NBER Working Paper No. 11265
April 2005
JEL No. K13, K4

ABSTRACT

We evaluate Richard Posner's famous hypothesis that common law converges to efficient legal rules using a model of precedent setting by appellate judges. Following legal realists, we assume that judicial decisions are subject to personal biases, and that changing precedent is costly to judges. We consider separately the evolution of precedent under judicial overruling of previous decisions, as well as under distinguishing cases based on new material dimensions. Convergence to efficient legal rules occurs only under very special circumstances, but the evolution of precedent over time is on average beneficial under more plausible conditions.

Nicola Gennaioli
University of Stockholm
nicola@iies.su.se

Andrei Shleifer
Department of Economics
Harvard University
Littauer Center M-9
Cambridge, MA 02138
and NBER
ashleifer@harvard.edu

1. Introduction

In his Economic Analysis of Law (2003, 1st ed. 1973), Richard Posner raises a question: in a common law system, does judge-made law converge to efficient legal rules? Put differently, do precedents converge to fixed rules, and if so, are these rules efficient? Posner hypothesizes that common law tends toward efficiency. Rubin (1977) and Priest (1977) suggest that disputes involving inefficient legal rules are more likely to be taken to court rather than settled, leading to the replacement of such rules over time. These articles do not focus on how judges actually make decisions. In this paper, we model the evolution of precedents through a series of judicial decisions, and examine its consequences for the convergence and efficiency of legal rules.

The doctrine of *stare decisis*, of deference to precedent, is a crucial feature of common law (e.g., Hayek 1960, Stone 1985, Posner 2003). Respect for precedents gives common law its stability and predictability. In addition, it enables appellate courts to communicate with, and therefore constrain, lower level courts more effectively (Bueno de Mesquita and Stephenson 2002). At the same time, the possibility of judges changing legal rules allows the law to evolve, to adjust to new circumstances, and therefore to become ever more efficient over time. Posner (2003) recognizes that such legal evolution is most effective when judges maximize efficiency. But even when they do not, and differ in their approaches to law, a key evolutionary argument still sees the common law as evolving toward ever better rules. "The eccentricities of judges balance one another. One judge looks at problems from the point of view of history, another from that of philosophy, another from that of social utility, one is a formalist, another a latitudinarian, one is timorous of change, another dissatisfied with the present; out of the attrition of diverse minds there is beaten something which has a constancy and uniformity and average value greater than its component elements" (Cardozo 1921, p. 177). Thus even if judges do not maximize efficiency, evolution selects better legal rules.

To assess these views of the evolution of common law, we present a new model of precedent formation by appellate judges. Our model relies on two assumptions. First, following the legal realism literature (e.g., Cohen 1935, Frank 1930, Radin 1925, Stone 1985) and the theoretical work of Gennaioli (2004), we assume that judges hold biases favoring different types of disputants, and that these biases vary across the population of judges. Frank (1930, p. 28) defines bias as the ideas and beliefs that come from judges' past experiences or philosophies. For example, some judges might believe in literal interpretation of contracts, others in interpreting contracts so as to promote efficiency, and still others in interpreting contracts against the drafter (Posner 2004b). Frank (1930) and Radin (1925) go so far as to say that judges decide the cases backwards: they figure out what outcome is just from their point of view, and then find legal arguments to support their conclusions.

Many legal scholars accept the importance of judicial bias for rulings on politically sensitive issues (e.g., Pinello 1999, Rowland and Carp 1996, Revesz 1997, Sunstein, Schkade, and Ellman 2004). Judges also differ sharply in their sentencing decisions for a given crime (Partridge and Eldridge 1974). But judges may also have preferences over the outcomes of commercial disputes: they may favor the rich or the poor, the government or individuals, insurance companies or accident victims. As Posner (2004a, p. 14) – echoing Frank (1930) -- writes about federal district judges: "But [deciding a particular case in a particular way might increase the judge's utility] by advancing a political or ideological goal, economizing on the judge's time and effort, inviting commendation from people whom the judge admires, benefiting the local community, getting the judge's name in the newspaper, pleasing a spouse or other family member or a friend, galling a lawyer whom the judge dislikes, expressing affection for or hostility toward one of the parties – the list goes on and on." One piece of data on the importance of judicial preferences in commercial disputes, and the consequent unpredictability of judicial decisions, is the sharp share price reactions that companies experience on the dates judges issue decisions (Haslem 2004).

Judicial bias is related to the idea that judges may be swayed by external forces, including political influence, intimidation, or bribes. This alternative assumption has been used to investigate legal systems both historically and in developing countries (Glaeser and Shleifer 2002, Glaeser, Scheinkman, and Shleifer 2003). However, to understand the evolution of common law in a developed modern economy, the assumption of judicial bias appears to be more appropriate.

Second, also following Radin (1925) and Posner (2003, 2004a), we assume that changing precedent is personally costly to judges: it requires extra investigation of facts, extra writing, extra work of persuading colleagues when judges sit in panels, extra risk of being criticized, and so on. "Judges are people and the economizing of mental effort is a characteristic of people, even if censorious persons call it by a less fine name" (Radin 1925, p. 362). The assumption that, other things equal, judges would rather not change the law implies that only the judges who disagree with the current legal rule strongly enough actually change it. Posner (2003, p. 544) sees what he calls "judicial preference for leisure" as a source of stability in the law; we revisit this issue.

Using a model relying on these two assumptions, we examine the evolution of legal rules in the case of a simple tort: a dog bites a man (e.g., Landes and Posner 1987). We consider separately two types of revision of precedents: overruling and distinguishing. By overruling we mean the discarding and replacement of a prevailing legal rule by a new one. Rubin (1977), Priest (1977), and Posner (2003) seem to have this model of precedent change in mind. By distinguishing we mean the introduction of a new legal rule that endorses the existing precedent, but adds a new material dimension to adjudication, and holds that the judicial decision must depend on both the previously established dimension and the new one. Distinguishing cases is perhaps the central mechanism, or leeway, through which the law evolves despite binding precedents (Stone 1985). But the efficiency of this process has not received much analytical attention. So we ask whether the evolution of precedents through either overruling or distinguishing leads to convergence, whether such convergence is to efficient legal rules, and what factors render legal change more efficient.

Our argument is best illustrated with an example, which simplifies the model in significant ways. Consider the evolution of legal rules governing the liability of an owner of a dog that bit a bystander. Suppose for concreteness that there are only two material dimensions of the dispute: the dog's breed (a proxy for its aggressiveness), and whether the bystander provoked it. Suppose further that there are only two kinds of dogs in the world: pit bulls and golden retrievers, and two kinds of provocation: a kick or none. Suppose finally that the first best efficient liability rule calls for liability of all pit bull owners, with or without a provocation, since pit bulls are so dangerous that their owners should efficiently guard against the risks of their biting even after a provocation. However, the first best rule calls for liability of golden retriever owners only in the event of no provocation, since golden retrievers are happy and peaceful animals.

To illustrate our ideas, consider judges distinguishing cases. Suppose that the first case comes along, and that a man is bitten by a golden retriever with no provocation. Suppose that the issue of provocation does not even come up before the judge. However, we have a biased judge, who thinks that dogs are dangerous and unsavory pests, as are their owners, and establishes the rule that all dog owners are liable when dogs bite men. This, as we assumed, is not an efficient rule.

Suppose that, after a while, perhaps with many other cases of dog bites being adjudicated according to the established precedent of strict liability, a case comes up of a golden retriever biting a man who kicked it. Suppose, again for concreteness, that the judge who handles the case is of one of three types: anti-dog like the first judge, efficiency-oriented, or pro-dog, believing that all dogs are quiet pets and only bite men who deserve it. If the judge is anti-dog, he does nothing, and simply lets the precedent stand, without addressing the issue of provocation. In this case, this issue might perhaps be addressed by future judges. Alternatively, the same anti-dog judge can argue that he considered the provocation, but deemed it immaterial, in which case he effectively solidifies the inefficient precedent in which owners of all dogs are liable regardless of provocation, forever.

If the judge is efficiency-oriented, he recognizes that it is a better rule to hold owners of provoked golden retrievers not liable, and so introduces provocation as a new material dimension. This judge writes that the prior court has neglected to consider that sometimes golden retrievers are provoked, in which case it is not efficient to hold their owners liable. This judge clarifies the law entirely: owners of provoked golden retrievers avoid liability, all other dog owners are liable in the event of a bite. This is the case of Posner's efficiency-maximizing judges.

But suppose the judge is a misanthrope. He grabs the opportunity to introduce a material new dimension, and to rule that the precedent applies only to the cases of no provocation. He accordingly revises the legal rule to say that owners of all breeds are liable in the absence of a provocation (so he respects *stare decisis*), but not liable otherwise. With this new rule, not only do the owners of provoked golden retrievers now -- efficiently -- escape liability, but the owners of provoked pit bulls – inefficiently – escape liability as well. In fact, the social cost of the new rule, which wrongly holds the owners of provoked pit bulls not liable, could be much greater than that of the old rule, which wrongly holds the owners of provoked golden retrievers liable. Distinguishing the case and introducing a new material dimension by a biased judge, in this instance, leads the law away from efficiency. And to the extent that *stare decisis* is respected and material dimensions are exhausted by breed and provocation, the inefficient rule is the end of the evolutionary process.

In this discussion, we have not mentioned the judge's cost of changing the precedent relative to his benefit of doing so. If it is somewhat costly to change the legal rule, only the misanthrope may change the original precedent, since his preferred rule is far away from it. In contrast, the efficiency-oriented judge may not be willing to incur a cost just to obtain a small efficiency gain from eliminating the liability of the owners of provoked golden retrievers. Now the result is even worse than before: efficient rule changes do not take place, and only inefficient ones are implemented by extremist judges. Selection works the wrong way.

Of course, a fuller evaluation of the evolution of the precedent requires the consideration of all the different paths of change in the law, as well as a separate treatment of overruling and distinguishing. But two general principles stand out. First, legal change enables judges to reaffirm their own biases, and to undo the biases of their predecessors. Second, such change occurs more often when judges´ preferences are polarized because judges are more likely to be in strong disagreement with the current precedent. Putting these principles to work, we find that, in general, convergence to efficient legal rules occurs under very limited circumstances. With overruling, convergence may not occur at all, and the legal rules may fluctuate between extremes. One exception is the case where efficiency-oriented judges are activist, while heavily biased judges are not, in which case the law does converge to efficiency. With distinguishing, convergence is more likely, but the conditions for full efficiency are implausibly strict. Now judicial bias in not the only force undermining efficiency: because of the sequential order in which new dimensions are introduced into the law, even the rules established by efficiency oriented judges may be suboptimal from the long-run standpoint. Still, distinguishing has the virtue of bringing new data to dispute resolution. When the costs of changing the law are not too high, this informational benefit renders legal evolution on average beneficial, confirming Cardozo's views.

The next section outlines our model of legal precedent. Section 3 describes the efficient legal rules in that model. Section 4 presents a model of judicial overruling of past precedents. Sections 5 and 6 deal with the more interesting case in which judges distinguish cases and introduce new material dimensions into adjudication. Section 7 concludes the paper.

2. A Model of Legal Precedent

There are two parties, *O* and *V*, and a dog. The dog bit victim *V*, who seeks to recover damages from *O*, the dog's owner. The dog was not on a leash, so in order to assess *O*'s liability one should determine whether *O* breached the duty of care (in which case he is liable) or did not (in which case he is not liable).

Let P_{NP} be the probability that the dog bites *V* if *O* does not take precautions (he does not put it on a leash) and P_P the probability that the dog bites *V* if precautions are taken. *O* prefers not to take precautions because he does not want to buy a leash, dislikes limiting the dog's freedom, or simply does not want to sweat to keep the dog quiet. Let *C* be the cost of precautions for *O*.

The Hand formula holds that *O* has a duty of care (is liable) whenever $P_{NP} - P_P \geq C$, i.e., when the reduction in the probability of a bite (weighted by harm, here assumed to equal 1) more than offsets the cost of precautions to *O*. In contrast, *O* has no duty of care (is not liable) if $P_{NP} - P_P < C$, because precautions cost more than they yield.

Many circumstances determine whether *O* was careless. The dog's aggressiveness, the extent to which *V* provoked it, or the place where *O* and his dog walked may all influence the probability of a bite. We assume that two empirical dimensions – the aggressiveness of the dog and *V*´s provocation – determine liability, i.e. constitute *material* dimensions in this legal dispute.

Variable $a \in [0,1]$ measures the dog's aggressiveness. More aggressive dogs have larger values of *a*; a dog with $a = 0$ is extremely peaceful (a golden retriever) and less likely to bite *V* than a dog with $a = 1$ (a pit bull). Variable $q \in [0,1]$, where q stands for *V*´ s quietness , measures the extent to which *V* provoked the dog. If $q = 0$, *V* outrageously provoked the dog; if $q = 1$, *V* was maximally quiet. We assume that *a* and *q* are independently and uniformly distributed over the population of disputes. We further assume that:

$$(1) \qquad P_{NP} - P_P = \begin{cases} \overline{\Delta P} & \text{for } a + q \geq 1 \\ \underline{\Delta P} & \text{for } a + q < 1 \end{cases}$$

where $\overline{\Delta P} > C > \underline{\Delta P}$. Thus, *O* is optimally liable if and only if $a + q \geq 1$. Owners of violent dogs are optimally liable if *V*'s provocation was not egregious, owners of peaceful dogs may still be liable as long as *V* has not provoked them at all ($q = 1$).

In general, the social benefit of the leash is a function $\Delta P(a, q)$ increasing in *a* and *q*. We assume that it only depends on *a*+*q*, and that it "jumps" at *a*+*q*=1. We could allow for more general functions, but our assumptions conveniently clarify the analysis of legal change and its impact on welfare. The first restriction makes *a* and *q* symmetric for determining liability, which allows us to isolate the effect of legal change *per se,* abstracting from the specific nature of the dimension introduced into the law. The second restriction allows us to separate the *probabilities* of the different errors induced by a particular legal rule from their welfare *cost*.

A legal rule in this environment attaches a legal consequence (*O* liable, *O* not liable) to every possible situation, defined as a combination of *a* and *q*. The legal rule specifies all the circumstances (a, q) in which *O* does or does not have a duty of care (i.e. when $P_{NP} - P_P$ is estimated to be greater than C). In other words, a legal rule puts substantive content into Hand's formula by specifying how the incremental probability of an accident must be determined as a function of *a* and *q*. Different legal rules reflect different notions of how $P_{NP} - P_P$ ought to be determined from the empirical attributes of a case.

We restrict the attention to "threshold rules". A simple "threshold rule" uses only one dimension, say *a*, and specifies a threshold *A* such that *O* is held liable if and only if his dog is more aggressive than *A* (i.e., $a \geq A$). A two-dimensional threshold rule – using both *a* and *q* -- is defined by three thresholds A, Q_0 and Q_1 such that *O* is held liable either if $a \leq A$ but $q \geq Q_0$, or if $a > A$ but $q \geq Q_1$. Figure 1 shows a generic two-dimensional threshold rule in the *(a, q)* space:

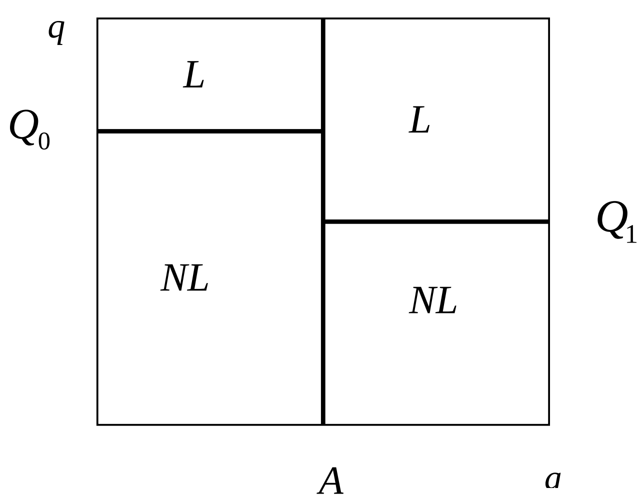

Figure 1.

In Figure 1, *O* is held liable in regions denoted by *L*, but non liable in those denoted by *NL*. Relative to a one-dimensional rule, a two-dimensional rule allows for liability of owners of peaceful dogs ($a \leq A$) whom *V* did not provoke ($q \geq Q_0$), and to hold not liable owners of aggressive dogs ($a > A$) whom *V* provoked egregiously ($q < Q_1$). Note that one should expect $Q_0 \geq Q_1$.

By focusing on threshold rules, we rule out a perfect (or first best) rule, holding *O* liable whenever $a + q \geq 1$. In reality, legal rules often take the form of threshold rules for reasons presumably related to enforcement costs, since they do not require judges to ascertain the exact values of *a* and *q*, but only whether certain thresholds on each of the elements had been crossed. For instance, while under the rule of Figure 1 the knowledge that $q > Q_0$ suffices to hold *O* liable, this is not the case under the perfect rule, which requires a much more precise (and presumably costly) verification of the facts (*q* and *a*).

Before calculating efficient threshold rules, we describe how judges set rules in our model of precedent. When initially no existing rule deals with dog bites, we assume that the only issue that comes up at trial is the aggressiveness of the dog. As a result, the judge adjudicating the dispute for the first time sets the legal rule by choosing the first threshold on *a*, which we call A_1. Owners of dogs more aggressive than A_1 are held liable; owners of dogs less aggressive than A_1 are not.

This specification of judicial decisions is an intermediate way of dealing with precedents. One can alternatively assume that the first judge sets a broad precedent, in which he considers the hypothetical issue of provocation even if does not arise in the specific dispute, and maps out owner liability on the whole (*a*, *q*) space. Under this specification, the law converges immediately, and we cannot talk about judges distinguishing cases; only replacing broad precedents by overruling. One can also imagine a judge setting a very narrow precedent, whereby instead of establishing an aggressiveness threshold for liability, he only makes a decision with respect to the specific breed of dog before him. In this case, there will presumably be a whole collection of narrow judicial decisions, with judges filling in gaps according to their biases, before some threshold aggressiveness level is arrived at. At that point, a judge who is unhappy with the existing cutoff of aggressiveness must change the rule. With such narrow precedents, legal evolution is slower, but the issues we discuss in this paper eventually arise as well.

Once the initial precedent is set, a judge dealing with the same issue later can change the rule. We consider two different models of *stare decisis*. In the first model, which we call overruling, judges discard A_1 and replace it with a new rule A_2. *Stare decisis* only binds in so far as it is costly for the judge to change the precedent. In the second model, which we call distinguishing, the second judge does not assault *stare decisis* with respect to *a*, but can still radically change the law by introducing the additional dimension *q* into adjudication, i.e. by setting Q_0 and Q_1. Effectively, the judge rules that the previous precedent is incomplete and applies to only some of the cases in the (*a*, *q*) space, but not others. To take an extreme example of the power of distinguishing, if the first judge establishes a strict liability rule with $A_1 = 0$, the second judge can reverse it completely by setting $Q_0 = Q_1 = 1$ and eliminating owner liability entirely (by saying that liability exists only when there is absolutely no provocation). In this model, precedent evolves through the introduction by judges of new *material* dimensions (*q* in this case) into the law. The English view of precedent contemplated only distinguishing as a source of legal change, at least until recently. In the United States, overruling coexists with distinguishing. To clarify the core properties of these two strategies of precedent change, as opposed to the judges' choice among them, we consider the cases of overruling and distinguishing separately.

We further assume that, for both overruling or distinguishing, a judge changing the legal rule incurs a personal effort cost *k*, regardless of how he changes the initial precedent. We take *k* to be a fixed cost, independent of the magnitude of precedent change. We could alternatively assume that more radical precedent changes entail higher personal costs. Some of the results of that model are different, but our broad qualitative conclusions continue to hold. We also maintain the view that *stare decisis* prevents the introduction of arbitrary and irrelevant dimensions into the law.

Model timing: *t = 0*: The first judge sets the rule by establishing the aggressiveness threshold A_1. This initial precedent guides adjudication until a judge (if any) changes the rule at some t'. What happens at $t=t'$ and after depends very much on which model we are in.

Overruling: The judge changing A_1 sets a new rule A_2, possibly giving rise to a new round of precedent change. In this model, the issue of provocation never arises.

Distinguishing: The judge changing the rule sets two provocation thresholds Q_0 and Q_1. In this case, the law is permanently fixed, as there are no further material dimensions to introduce.

In Section 4, we study the judges' objectives in changing the law, as well as their costs of doing so. But first, we investigate the efficient – welfare maximizing – rules that provide the normative benchmark for our analysis of legal change and judge made law.

3. Optimal Legal Rules

Legal rules affect social welfare – defined as the sum of *O*'s and *V*'s utility – by changing the precautions taken by dog owners. The likelihood of damages and the fine on *O* when he is found liable shape his decision to put the dog on a leash. We assume that fines are always set high enough to enforce precautions whenever the law dictates that they must be taken. First best welfare, achieved under optimal precautions (i.e., *O* puts the dog on a leash whenever $a+q \geq 1$), is equal to:

(2) $$W^{F.B.} = -(1/2)\underline{\Delta P} - (1/2)C,$$

where the probability of a bite when precautions are taken is normalized to 0. In half the cases, precautions are not efficient and the parties bear the extra risk $\underline{\Delta P}$ of the dog biting the man; in the other half, precautions are efficient and cost *C* to society.

Adjudication cannot achieve such high welfare since threshold rules necessarily induce judicial errors. If *O* is held liable but $a+q<1$, excessive precautions are taken; if *O* is held not liable but $a+q \geq 1$, *O*'s level of care is too low. Let $\Pr(L|NL)$ and $\Pr(NL|L)$ be the probabilities that *O* is erroneously held liable and not liable, respectively, under a particular legal rule. The *loss* of social welfare relative to the first best under this rule is equal to:

(3) $$\Lambda = \Pr(NL|L)\Lambda^{under} + \Pr(L|NL)\Lambda^{over}$$

$\Lambda^{under} = \overline{\Delta P} - C$ is the social cost of under-precautions when *O* is mistakenly held not liable, $\Lambda^{over} = C - \underline{\Delta P}$ is the social cost of over-precautions when *O* is erroneously held liable. In our analysis, these costs of over- and under-precautions are constant, and we focus on how different legal rules affect the likelihood of different mistakes in adjudication.

For concreteness, we assume that under-precautions are the greater evil to avoid:

Assumption 1: $\Lambda^{over} / \Lambda^{under} \equiv \lambda \leq 1$.

Figure 2 represents both the first best. and the one-dimensional rule, in the (a,q) space. Under the first best, *O* is liable above the diagonal but not below. The vertical bold line represents the one-dimensional threshold legal rule, *A*, which holds *O* liable if and only if $a \geq A$.

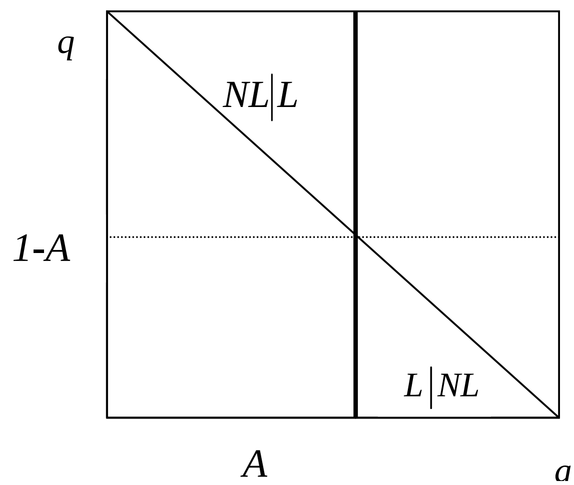

Figure 2.

The one-dimensional threshold rule holds *O* mistakenly liable in region $L|NL$ and mistakenly not liable in region $NL|L$. For a given *A*, the probabilities of these errors are given by $\Pr(L|NL) = (1/2)(1-A)^2$ and $\Pr(NL|L) = (1/2)A^2$. The corresponding loss of social welfare is:

(4) $$\Lambda(A) = (1/2)A^2\Lambda^{under} + (1/2)(1-A)^2\Lambda^{over}$$

If A is the initial precedent, social losses are $\Lambda(A)$ – an average of over and under-precautions costs under the error probabilities that A induces. The larger is A (the more the initial rule favors O), the larger is the loss from under-precautions. Over-precaution costs increase as A gets smaller.

Figure 3 illustrates the two dimensional legal rule with thresholds A, Q_0 and Q_1. Here O is over-punished in region $L|NL$, with area $\Pr(L|NL) = (1/2)\left[(1-Q_0)^2 + (1-A-Q_1)^2\right]$, and under-punished in region $NL|L$, with area $\Pr(NL|L) = (1/2)\left[(A+Q_0-1)^2 + Q_1^{\,2}\right]$.

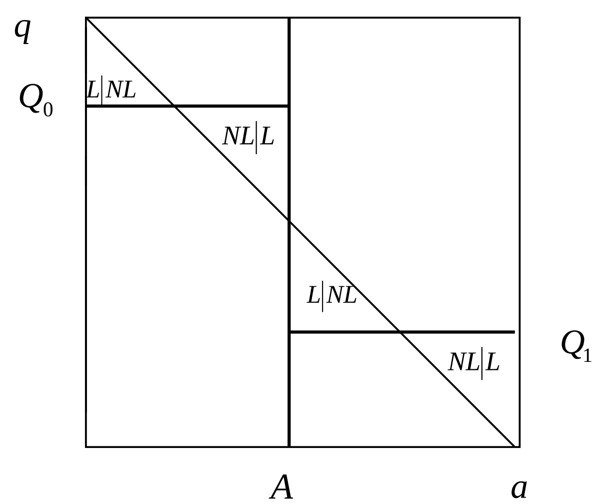

Figure 3.

The social loss from the use of the two-dimensional threshold legal rule is given by:

$$(5)\qquad \Lambda(A,Q_0,Q_1) = (1/2)\left[(A+Q_0-1)^2 + Q_1^{\,2}\right]\Lambda^{under} + (1/2)\left[(1-Q_0)^2 + (1-A-Q_1)^2\right]\Lambda^{over}$$

By minimizing (4) with respect to A and (5) with respect to A, Q_0, Q_1, we find the optimal initial precedent and the optimal two-dimensional rule.

Proposition 1: *i) The optimal one-dimensional legal rule (initial precedent) is given by*

$$A_L = (\Lambda^{over} / \Lambda^{under}) / \left[1 + (\Lambda^{over} / \Lambda^{under})\right].$$

ii) The optimal two-dimensional legal rule is given by

$$A_F = 1/2,\quad Q_{0,F} = (1+A_L)/2,\quad Q_{1,F} = A_L/2.$$

The optimal initial precedent A_L responds to social costs. The larger is the relative cost of over-precautions (the larger is $(\Lambda^{over} / \Lambda^{under})$), the more lenient is the optimal rule (the larger is A_L).

This is also true for the optimal two-dimensional thresholds A_F, $Q_{0,F}$ and $Q_{1,F}$. The optimal legal rule is more lenient toward *O*, the larger is the relative cost of over-precautions ($\Lambda^{over} / \Lambda^{under}$). Going back to Figure 3, if the cost of under-precautions is very large, the optimal Q_O and Q_1 should be small to keep $NL|L$ -- the region where careless owners are held not liable -- small. Conversely, for larger over-precautions cost, the optimal Q_O and Q_1 should be raised so as to reduce the size of $L|NL$, the region where *O* is mistakenly held liable.

The efficiency of a rule generally depends on two factors: its overall imprecision $\Pr(NL|L) + \Pr(L|NL)$, and the ratio of different errors $\Pr(NL|L)/\Pr(L|NL)$. The optimal initial precedent and the optimal two-dimensional rule fare equally well in terms of this second factor (i.e. they induce the same $\Pr(NL|L)/\Pr(L|NL)$), but the two-dimensional rule is more precise, and thus more efficient. $A_F = 1/2$ yields the full benefit of extra information. For extreme A_F (1 or 0), the added dimension *q* is worthless: a single threshold on *q* (Q_0 or Q_1) describes liability over the entire *(a, q)* space, just like in a one-dimensional rule.

With the results of this section in mind, we can move on to study judicial lawmaking under the two postulated forms of *stare decisis*. Our analysis is driven by three main questions. First, we ask whether there is a tendency for the process of precedent change to converge to a decision rule limiting the impact of judicial idiosyncrasies. Second, we scrutinize Posner's proposition that not only do the common law rules converge, but also that they converge to the efficient ones. By efficiency we will mean ex ante efficiency, before judge types are revealed. Third, we ask more broadly whether, aside from the question of convergence, legal change itself is beneficial from the social viewpoint. To this end, we lay out some circumstances under which this is the case.

4. How Judges Shape the Law

Like social welfare, the utility of a judge settling a dispute between *O* and *V* depends on the precision of the rule and on the ratio of different mistakes. However, we assume that a judge's objective diverges from efficiency because of his bias, which reflects his preference for *V* or *O* and induces him to sacrifice optimal precision for a pattern of mistakes more favorable to the preferred party. Specifically, we assume that the utility of judge *j* is given by:

$$U_j = -\beta_{V,j} \Pr(NL|L) - \beta_{O,j} \Pr(L|NL) \tag{6}$$

Judges dislike making mistakes, but they do not dislike the two types of mistakes equally. $\beta_{O,j}$ and $\beta_{V,j}$ ($\beta_{V,j}, \beta_{O,j} \geq 0$) capture the preference of judge *j* for *O* and *V*, respectively: the larger is $\beta_{O,j}$, the more he is eager to hold *O* not liable, the larger is $\beta_{V,j}$, the more he is willing to hold *O* liable.

Under the assumed utility function, judges are unhappy with any mistake they make (albeit differentially for different errors). Thus, if we did not restrict attention to threshold rules and allowed for all two-dimensional rules, even biased judges would pick the first best one (the diagonal). This judicial aversion to making mistakes leads to judicial self-restraint that is crucial for our results: even a judge heavily biased against dog owners would not introduce the most anti-owner liability rule available if this rule leads to mistakes he can avoid, including mistakes favoring bite victims. Such preferences allow us to emphasize – in line with the legal realists – that judicial bias is more problematic in the presence of uncertainty, when judges trade off different errors. But we do not model the kind of favoritism where the judge rules against dog owners even when he knows for sure that they should not be efficiently held liable.

In our specification of judicial preferences, a judge's utility depends on the expected outcome arising from the application of a given rule, not from the resolution of a particular case. Such a judge would consider replacing a legal rule he dislikes even if the outcome of the specific case before him is the same under the new rule. A judge cares about having a rule in place that meets his idea of justice, rather than about delivering a desired outcome in a specific dispute before him. This assumption is particularly appropriate for appellate judges, who establish legal rules rather than resolve specific disputes.

The judge is assumed to ignore the possibility that the rule he establishes will be changed in the future. In particular, he does not act strategically with respect to future judges. This assumption can be relaxed, although at the cost of increased analytical complexity, and we believe our basic results would be preserved. One way to justify the present framework is by noting that precedents change relatively rarely, and therefore a judge discounting the future may not put much value on the effect of future legal change.

There is a measure 1 of judges, who can be of three types: share γ of judges are *Unbiased*, with bias $\beta_{O,j} / \beta_{V,j} = \lambda$ reflecting social welfare; the rest is equally divided among *Pro-O*, with bias $\beta_{O,j} / \beta_{V,j} = \lambda\pi$ and *Pro-V*, with bias $\beta_{O,j} / \beta_{V,j} = \lambda / \pi$. Parameter π ($\pi \geq 1$), measures the polarization of judges' preferences: with a higher π, the preferences of *Pro-O* and *Pro-V* judges are more extreme (there is more disagreement among them). We assume that all judges have the same preference intensity and normalize it to 1 ($\beta_{V,j} + \beta_{O,j} = 1,\ \forall j$).

Initial Precedent

The first judge adjudicating a dispute between *O* and *V* establishes the initial precedent. We assume that, in this dispute, the issue of provocation never arises (and the judge does not entertain legal rules taking provocation into account unless that issue arises in the dispute). To resolve this dispute, the judge selects a threshold A to maximize:

(7) $$U_1(A) = -(1/2)\beta_{V,1}A^2 - (1/2)\beta_{O,1}(1-A)^2$$

$\beta_{V,1}$ and $\beta_{O,1}$ parameterize the bias of the initial judge. Define $\beta_1 = \beta_{O,1} / \beta_{V,1}$ as the *Pro-O* bias of this judge. Minimizing the objective above, we find that:

(8) $$A_1 = \frac{\beta_1}{1+\beta_1}$$

The subscript indicates that A_1 is the initial precedent set with *Pro-O* bias β_1. The result is intuitive: the more *Pro-O* is the judge, the more lenient he is (the higher is A_1). A_1 coincides with the efficient initial precedent A_L only if $\beta_1 = \lambda = \Lambda^{over} / \Lambda^{under}$, i.e. if the judge's bias toward *O* reflects the relative social cost of over-precautions.

Under A_1, social losses are given by $\Lambda(A_1)$. Given the variety of judges' preferences, there is no reason to presume that A_1 is set efficiently, i.e. to minimize $\Lambda(A_1)$. If the case ends up in front of a *Pro-O* judge ($\beta_1 > \lambda$), too many aggressive dogs roam and bite with impunity; if instead it ends up in front of a *Pro-V* judge ($\beta_1 < \lambda$) too many peaceful dogs are put on a leash.

Overruling

Depending on β_1, the initial precedent may turn out to be severely inefficient. Still, this bias may be corrected through the *change* of precedent. As Cardozo (1921) might suggest, if the initial rule is very biased in one direction (say *Pro-O*), the successive intervention by a *Pro-V* judge would modify the law by *tempering* its initial bias with the opposite one.

Suppose that precedent A_i is in place, and judge *j* takes the initiative to change the law. He then sets a new threshold A_j, equal to

$$A_j = \frac{\beta_j}{1+\beta_j} \quad (9)$$

where $\beta_j = \beta_{O,j} / \beta_{V,j}$ is the *Pro-O* bias of judge *j*. To see if judge *j* in fact changes the law, we must consider his personal incentive to do so, as judges may be unwilling to bear the effort and other costs of legal change. Judge *j* changes the law only if:

$$U_j(A_j) - U_j(A_i) \geq k \quad (10)$$

i.e. when the cost to the judge of changing the law is smaller than its benefit. Using the judge's utility function, we find that judge *j* overrules the precedent when:

(11) $$\frac{(\beta_i - \beta_j)^2}{(1+\beta_i)^2(1+\beta_j)^2} \geq 2k$$

The smaller is the cost *k*, the higher the chance that a judge changes the law. Clearly, if there are no costs of overruling, the judge prefers to change the law and set the rule A_j that reflects his own bias. Thus, for $k = 0$, judges always overrule precedent, creating expected social losses of $E_j[\Lambda(A_i)]$, where the expectation is taken over all judge types. But how do judges with a positive *k* react to precedent? Since judges regain discretion through overruling, they are more activist when the prevailing legal rule is further away from their preferred one (see the numerator of eq. (11)). This is more likely to be the case if judges´ preferences are more polarized (π is higher). In this way, the extent of disagreement among judges determines the long run configuration of precedent.

The case of $\lambda = 1$ illustrates this intuition. Now there exist two levels of polarization $\pi_1 \leq \pi_2$ such that judicial behavior can be summarized in Table 1. The boxes in the table report the circumstances when a judge *j* changes the legal rule he inherited from *i*. Three patterns of behavior emerge. First, judges never change the initial rule of an adjudicator of the same type. Second, judges´ behavior is symmetric: if judge *j* overrules A_i, then judge *i* overrules A_j. Third, more judges change the law as π increases.

Judge i / *Judge j*	*Pro-O*	*Unbiased*	*Pro-V*
Pro-O	Never	$\pi \geq \pi_2$	$\pi \geq \pi_1$
Unbiased	$\pi \geq \pi_2$	Never	$\pi \geq \pi_2$
Pro-V	$\pi \geq \pi_1$	$\pi \geq \pi_2$	Never

Table 1.

Since a judge changes precedent to set a rule reflecting his preferences, there is no need for him to repudiate a rule established by someone with the same views. Judges are also reluctant to change the law when their views differ little relative to the cost of overruling ($\pi < \pi_1$). At intermediate levels of polarization ($\pi \in [\pi_1, \pi_2)$), *Pro-O* and *Pro-V* judges overrule each other. Unbiased judges stop being passive when $\pi \geq \pi_2$ as they begin to overrule extremists.

The polarization of judicial preferences determines the ultimate configuration of judge-made law. For low polarization ($\pi < \pi_1$), precedent does not change from A_1. When polarization is intermediate ($\pi \in [\pi_1, \pi_2)$), precedent oscillates between *Pro-O* and *Pro-V* rules unless an *Unbiased* judge sets the initial rule, which then becomes permanent. At high levels of polarization, every judge pursues legal change. Overruling is highly problematic for convergence.

Little changes for $\lambda < 1$. Since now *Pro-O judges* disagree with *Unbiased* ones more than the *Pro-V* do, there is a π_3 such that for $\pi \in [\pi_2, \pi_3)$ *Pro-O* judges overrule *Unbiased* judges even if *Pro-V* ones stay passive. Convergence is only achieved for very high *k*'s. In that case, regardless of polarization of judicial preferences, no judge ever changes the law, which remains fixed at A_1.

This result casts doubt on the notion that precedent is a powerful mechanism to constrain judicial arbitrariness. When precedents can be overruled, legal unpredictability is the greatest when judicial preferences are polarized. In a sense, a system of overruling suffers from the very same malady it seeks to cure. Perhaps this point sheds light on the challenges of judicial law-making in politically charged cases, where judicial preferences are highly polarized, and legal evolution itself becomes a source of the very unpredictability it purports to eliminate.

What about the ex ante efficiency of judge-made law? Does the evolution of precedent lead to optimal legal rules? The benchmark here is A_L, the optimal one-dimensional threshold rule we found in section 3. The following proposition explains when overruling leads to optimality:

Proposition 2: *Under overruling, judge made law is efficient if and only if all judges are unbiased.*

In expectation, the law converges to the efficient decision rule A_L only if there is full agreement among judges *and* their views are aligned with efficiency. When some judges are biased, there is a chance that the initial precedent is either set by a *Pro-O* or by a *Pro-V* judge. In either case, the law does not converge to an efficient rule. The contribution of efficiency-seeking judges to the convergence of common law to efficiency is recognized by Posner (2003), although he does not explain just how stringent the conditions for full efficiency are.

Proposition 2 restricts the chances that judge-made law is fully efficient, but it does not in itself prevent overruling of precedents from being beneficial in an evolutionary sense. It is thus important to answer the following question: under overruling of precedents, is there a *tendency* for the law to improve over-time? The next result addresses this question.

Corollary 1: *Under overruling, expected social losses are (weakly) minimized for* $k = +\infty$.

When people take precautions based on the law of the moment, welfare here is independent of *k*. The reason is the model's symmetry: if an efficiency oriented judge overrules a biased precedent, then the biased judge will overrule the efficient precedent, creating no overall tendency toward efficiency. In particular, efficiency is the same when $k = 0$ as when $k = +\infty$: in the absence of legal change, uncertainty over the bias of the initial judge leads to social losses of $E_i\left[\Lambda(A_i)\right]$, the same that prevail when $k = 0$. Despite such irrelevance, there might be some reasons why, under overruling, $k = +\infty$ is preferred. Such values as the predictability of the law or equal treatment may render a bad but stable law preferable to an equally efficient on average but unpredictable law.

The dismal performance of legal evolution under overruling of precedents is due to the symmetry of judges´ behavior: by mutually overruling each other, active judges neutralize their respective impacts on the law. For legal change to be desirable, the odds of moving from a bad to a good rule, should be greater than those of moving in reverse.

When would that be the case? Efficient precedents would be harder to overrule when *Unbiased* judges are more activist than the extremists (have a lower overruling cost k_U). This might be the case, for example, when judicial ability is positively correlated with both unbiasedness and peer respect, which reduces the private cost of legal change:

Proposition 3: *If* $k_U < k$ *, there exists* $\hat{\pi} \le \pi_2$ *such that, for* $\pi \in [\hat{\pi}, \pi_2]$*, judge-made law converges to efficiency.*

When the *Unbiased* judges are more interventionist than the biased ones ($k_U < k$), then not only is there a possibility for legal evolution to improve the law, but Posner's efficiency conjecture also holds. As Proposition 3 shows, the activism of the *Unbiased* judges is not sufficient for the law to converge to full efficiency and a further condition must be met: polarization should not be too extreme. Figure 4 below represents, for a given $k < 1/2$, the set of π and $k_U < k$ where overruling leads to full efficiency and where it does not.

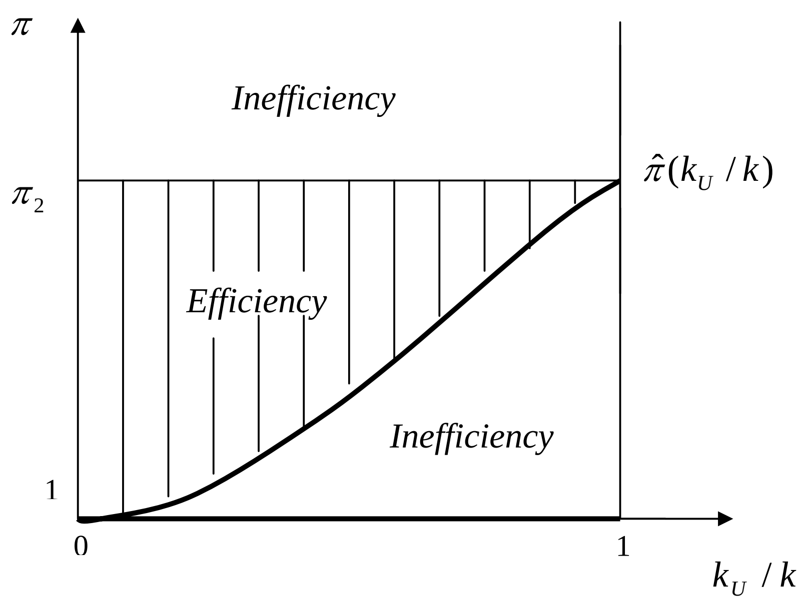

Figure 4.

The vertical axis of the diagram measures preference polarization π, the horizontal axis measures $k_U / k \in [0,1]$, the strength of *Unbiased* judges' comparative advantage in legal change, maximal at $k_U = 0$ and minimal at $k_U = k$. The function $\hat{\pi}(k_U / k)$ shows, for every k_U, the level of polarization above which *Unbiased* judges overrule inefficient precedents. The dashed area and the bold lines identify the set of parameters for which judge-made law converges to efficiency.

Just as in Proposition 2, Posner's conjecture holds – regardless of k_U – when all judges are benevolent ($\pi = 1$). The intriguing feature of Figure 4 is that, in the shaded region above $\hat{\pi}(k_U / k)$, the law converges to full efficiency even if only a few judges are *Unbiased*. Indeed, when $\pi \geq \hat{\pi}$, *Unbiased* judges correct inefficient precedents, and when it is also the case that $\pi \leq \pi_2$, extremists stay passive and do not reverse the efficiency promoting decisions of the *Unbiased* judges.

To summarize, if the behavior of different judges is symmetric, there is no tendency for overruling of precedents to be beneficial and – if judges' preferences are very polarized – convergence itself is unlikely. On the other hand, the greater activism of the *Unbiased* judges effectively leads to a virtuous evolution of the law, to the point of vindicating Posner's conjecture, at least when the polarization of judicial preferences is intermediate.

Our results identify the critical role that efficiency seeking judges play in the convergence and efficiency of judge-made law under overruling. Yet Cardozo's quote suggests that building the law up with incremental dimensions – a feature absent in overruling -- may by itself render the evolution of precedent desirable irrespectively of the type of judges engaging in legal change. Below we investigate this channel of legal evolution.

5. Distinguishing

In the common law tradition, the ability of judges to distinguish cases from previous precedent serves an important constructive role. It allows new information to be considered in adjudication, and thereby enables the law to evolve, to adjust to new circumstances, and to become more precise. Such adaptability of common law is seen by writers from Holmes (1897), to Cardozo (1921), Hayek (1960), Stone (1985), and Posner (2003) as one of the chief virtues of judge-made law. Here we study such a process of distinguishing cases from precedents, and examine its implications for the convergence and efficiency of judge-made legal rules.

The Form of Legal Change and its Welfare Consequences

The utility of a judge who modifies the initial precedent A_1 (we call him judge 2) by introducing the dimension *q* into the legal rule by the choice of thresholds Q_0 and Q_1 is

(12) $$-(1/2)\beta_{V,2}\left[(A_1+Q_0-1)^2+Q_1{}^2\right]-(1/2)\beta_{O,2}\left[(1-Q_0)^2+(1-A_1-Q_1)^2\right]$$

The first term of the expression represents the cost for judge 2 of mistakenly holding *O* not liable (i.e. ruling against *V*), while the second term is the cost for judge 2 of erroneously holding *O* liable. Define $A_2=\beta_2/(1+\beta_2)$. Here A_2 can be interpreted as the ideal threshold on the dog's aggressiveness that would be chosen by judge 2 if he were setting the initial precedent. From first order conditions, we obtain

(13) $$Q_{0,2}(A_1)=1-(1-A_2)A_1$$

(14) $$Q_{1,2}(A_1)=A_2(1-A_1)$$

These reaction functions tell us that some re-equilibrating mechanism is indeed built into precedent, because $Q_{0,2}$ and $Q_{1,2}$ decrease in A_1. Regardless of judge 2´s bias, a more *Pro-O* initial rule induces him to use dimension *q* relatively more in favor of *V*. However, since the extent of the adjustment depends on the bias of the second judge, summarized by A_2, we need to carefully evaluate the welfare impact of legal change through distinguishing before assessing its desirability.

The probabilities of different mistakes after precedent change are

(15) $$\Pr(L|NL)=(1/2)(1-A_2)^2\left[A_1^2+(1-A_1)^2\right];$$

(16) $$\Pr(NL|L)=(1/2)A_2^2\left[A_1^2+(1-A_1)^2\right]$$

These expressions show why, contrary to the common wisdom, judges' biases *do not* balance one another in judge-made law: the ratio of the two errors, $\Pr(NL|L)/\Pr(L|NL)$, is fully determined, through A_2, by the <u>desired</u> bias of the second judge! When judge 2 introduces *q* into adjudication, he discretionally sets $Q_{0,2}$ and $Q_{1,2}$ so as to favor the party he prefers. As a result, there is no presumption that the final configuration of the law is less biased than the initial precedent. Due to the very discretion embodied in distinguishing cases, judge-made law cannot eliminate this first effect of judicial bias: it cannot correct the ratio of different errors. In this sense, the eccentricities of judges do *not* balance one another and judge made law is *not* a solution to the presence of judicial bias.

However, in our model, judicial bias also affects the efficiency of the law by affecting the *overall* likelihood of judicial error, i.e. the law's precision. This channel if very important for distinguishing. Indeed, a very *Pro-O* judge is willing to design a very imprecise rule in order to excuse dog owners, but since he does not want to totally discard the information embodied into the existing legal rule, the waste of information associated with his exercise of discretion is limited. Generally speaking, the threshold A_1 set by judge 1 limits the arbitrariness of judge 2 when introducing *q*, thus improving the precision of the law. The term $\left[A_1^2 + (1 - A_1)^2\right] \leq 1$ in $P(NL|L)$ and $P(L|NL)$ accounts for this second effect of precedent change.

The strength of this second effect depends on the extremism of the first judge. Suppose that judge 2 is extremely *Pro-O* and consider two cases: in the first judge 1 is extremely *Pro-V*, in the second judge 1 is moderate. In the first case, judge 1 only cares about not excusing the owners of any dogs who should be held liable, and therefore sets $A_1 = 0$. Since judge 2 only cares about not holding liable owners of dogs who efficiently should not be, his optimal choice in light of the precedent he faces is to undo the will of judge 1 entirely and set $Q_0 = Q_1 = 1$. According to judge 2, any provocation, no matter how minor, eliminates dog owner's liability. When judge 1 is so extreme, judge 2 is both able *and willing* to move from the regime of strict liability to the regime of virtually no liability by distinguishing the case based on provocation.

Suppose in contrast that judge 1 is moderate, cares about both types of errors, and therefore sets $A_1 = 1/2$. The *Pro-O* Judge 2 still *can* set $Q_0 = Q_1 = 1$, but he does not want to. The reason is that he can set $Q_0 = 1$, and $Q_1 = 1/2$, and this way avoid the error of holding non-liable the owners of unprovoked vicious dogs. He still keeps the area of false liability down to zero, but because he does not like making *any* errors, his decision is more efficient. Judge 1's moderation entails the relative moderation of judge 2. This discussion also shows that our assumption about judicial preferences actually matters; if judge 2 only cared about favoring dog owners without regard for making errors, he would set $Q_0 = Q_1 = 1$ regardless of what judge 1 did before him. To summarize:

Proposition 4: *Judge1's moderation leads to judge 2's moderation*.

How does distinguishing affect social welfare? The tension between the "bias" and "imprecision" effect of distinguishing can be gauged by looking at social losses after q is introduced:

(17) $$\Lambda(A_1, Q_{0,2}, Q_{1,2}) = \left[A_1^2 + (1 - A_1)^2\right]\Lambda(A_2)$$

After judge 2 revises precedent, social loss is a product of two terms. The term $\Lambda(A_2)$ stands for the social loss under the hypothetical assumption that the initial rule is chosen by judge 2. This term captures the bias effect of distinguishing, whereby judges regain their discretion. The term $\left[A_1^2 + (1 - A_1)^2\right]$ captures the "precision" effect of initial precedent, which influences judge 2's optimal exercise of discretion, thus reducing social losses. If the initial threshold on *a*, A_1, were not binding through *stare decisis*, the social loss would be entirely determined by the preferences of judge 2, as reflected by the hypothetical A_2. By comparing (17) with $\Lambda(A_1)$ we find:

Proposition 5: *Legal change through distinguishing cases is beneficial when either of the following conditions is met*

i) $\Lambda(A_1) \geq \Lambda(A_2)$

ii) $\Lambda(A_1) < \Lambda(A_2)$, *but* $\left[A_1^2 + (1 - A_1)^2\right]$ *is small enough*.

Condition *i)* says that distinguishing is always beneficial when the preferences of judge 2 are more efficiency oriented than those of judge 1. Even if this is not the case, condition *ii)* says that distinguishing may still be beneficial if the greater precision induced by the inclusion of q more than offsets the loss from adversely changing the ratio of different mistakes. Put differently, distinguishing is only harmful when two conditions hold simultaneously: judge 2's preferences, if he was hypothetically setting the initial precedent, would yield greater social losses than those of judge 1, *and also*, judge 1's preferences are sufficiently extreme that his initial decision only minimally constrains judge 2's optimal choice. Or, to put this more broadly, legal change through distinguishing is most likely to be detrimental when both judge 1 and judge 2 are extremists, and when judge 2's extremism is more detrimental to social welfare than that of judge 1.

To illustrate how distinguishing can be harmful, we make the following

Assumption 2: *Judge 1 is Pro-V, with bias* $\beta_1 < \lambda$, *and judge 2 is Pro-O, with bias* $\beta_2 > \lambda$.

Together with Assumption 1 (which posits that under- precautions are socially costlier than over-precautions), Assumption 2 tells us that the bias of judge 1 is more efficiency oriented than that of judge 2, so that $\Lambda(A_1) < \Lambda(A_2)$. Under Assumption 2, condition *i)* of Proposition 5 is violated. In this case, distinguishing *may* be harmful to society as it just represents a way for judge 2 to excuse careless owners of very aggressive dogs, whom he is fundamentally sympathetic to, by finding *V*´s provocation. Such an excuse may be so costly to society as to undermine the desirability of legal change through distinguishing cases altogether.

Overall, our analysis suggests two points. First, in a system of precedent, the desirability of distinguishing and the efficiency of judge-made law depend on judicial bias, particularly on the bias of the last judge who changes legal rules. Legal precedent does not balance the different opinions of judges, and its ultimate configuration may be severely inefficient if an "anti-efficiency" judge sets it. Second, the effectiveness of precedent in constraining judges depends on its initial configuration: the more biased is the initial rule, the more likely is that the introduction of further empirical dimensions is biased as well. The precision of the law exhibits a strong path dependency.

However, just as we saw with *overruling*, we cannot properly evaluate a system of precedent before determining which judges are likely to change the law.

Judicial Activism and Distinguishing

By comparing the utility judge *j* derives from retaining A_i completely with the utility he obtains by introducing his preferred thresholds Q_{0j} and Q_{1j} into the law (for A_i given), we find that judge *j* distinguishes A_i when:

$$\frac{\beta_i^2 + \beta_j^2}{(1+\beta_i)^2(1+\beta_j)^2} \geq 2k \tag{18}$$

The smaller is the cost *k*, the greater the chance that a judge changes the law. For $k = 0$, judges always distinguish precedents (and introduce Q_{0j} and Q_{1j}). Notice the difference between (18) and the condition for *overruling*. Just as with overruling, a greater disagreement between judges *j* and *i* leads the former to distinguish A_i more often. However, now there is also an informational gain associated to distinguishing. This gain is stronger for moderate judges ($\beta_j = 1$) who care most about the precision of the law. Such a gain may induce a judge to distinguish even a precedent set by a predecessor with identical preferences.

To evaluate the properties of distinguishing, we must characterize the activism of different judges, in particular their proclivity to change a precedent set by a judge they disagree with. As in Section 3, the polarization of judges´ preferences determines the final configuration of judge-made law. This can be seen clearly for $\lambda = 1$. Suppose for simplicity that *k* is high enough that judges do not change their preferred initial precedent. Then there exist two thresholds $\tilde{\pi}_1 \leq \tilde{\pi}_2$, such that for $\pi < \tilde{\pi}_1$, judges never introduce *q* into the law and the initial precedent (solely based on *a*) stays in place forever. At intermediate polarization ($\tilde{\pi}_1 \leq \pi < \tilde{\pi}_2$), *Pro-O* and *Pro-V* judges distinguish each others´ precedents and the law converges to a two dimensional legal rule unless an *Unbiased* judge sets the initial precedent. Finally, when polarization is high ($\pi \geq \tilde{\pi}_2$) *Unbiased* judges stop being passive, distinguish precedents set by extremists (and vice-versa), and the law always converges to a two-dimensional legal rule.

Under distinguishing, the law *always* converges, at least if there are a finite number of empirical dimensions germane to defining a transaction and – which is essentially the same – if the nature of transactions does not change over time. Such long-run predictability hinges on the assumption that judges cannot introduce irrelevant dimensions into the law. In other words, the *materiality* of a dimension is a physical characteristic that even the most biased judges cannot subvert. A *stare decisis* doctrine constraining judges to modify the current precedent only by enriching the empirical content of the law is successful in assuring convergence.

6. The Properties of Judge-Made Law under Distinguishing

The result that under distinguishing the law converges does not imply that distinguishing is an effective constraint on the arbitrariness of judges. Indeed, all judges now use the same legal rule in the long run, but the rule may be very biased because it was set by a biased judge.

This consideration makes it imperative to evaluate when, under distinguishing, judge-made law converges to the optimal two-dimensional rule we analyzed in Section 3. We then find:

Proposition 6: *There exists a* $\tilde{k} > 0$ *such that, under distinguishing judge-made law converges to the efficient two-dimensional rule if and only if all judges are unbiased,* $\Lambda^{over} = \Lambda^{under}$*, and* $k \leq \tilde{k}$.

In the same spirit as Proposition 2, Proposition 6 says that, when judges distinguish cases, a population of fully unbiased judges is necessary for judge-made law to converge to the efficient two-dimensional rule. However, two *additional* conditions must now be fulfilled.

First, it is essential that judges be interventionist enough to introduce q into the law: this is the condition $k \leq \tilde{k}$. If all judges are unbiased but $k > \tilde{k}$, the law – just as in the case of overruling – starts out at the optimal one-dimensional rule A_L and simply stays there.

Second, it must be that $\lambda = 1$ ($\Lambda^{over} = \Lambda^{under}$), i.e., the optimal rule is unbiased in terms of the ratio of different errors. Recall that Proposition 1 says that, in the optimal two-dimensional rule,

$Q_{0,F}$ and $Q_{1,F}$ yield the optimal ratio of errors as a function of $\Lambda^{over} / \Lambda^{under}$, whereas $A_F = 1/2$ maximizes the precision of the law independently of $\Lambda^{over} / \Lambda^{under}$. Under distinguishing, even with two unbiased judges ruling sequentially, the emerging two-dimensional rule is suboptimal: the initial precedent is set at A_L and not at 1/2 because the first judge disregards the effect of his choice on long run law. Forward looking behavior on this judge's part would not remove this inefficiency unless he is infinitely patient. Indeed, even if the *Unbiased* judges distinguish cases more often than do the extremists ($k_U < k$), we do not expect the law to converge to efficiency except in the knife-edge case where the goal of maximizing the precision of the law and the goal of biasing it optimally coincide, i.e. when $\Lambda^{over} = \Lambda^{under}$. In that case, an *Unbiased* judge sets $A_1 = 1/2$ and, provided that an *Unbiased* judge also introduces q into the law, a fully efficient rule emerges.

The ambitious normative benchmark used in this section (the optimal two-dimensional rule) makes it hard for judge-made law to be efficient, but it also reveals the greater potential of legal evolution under distinguishing. The "technological" advantage of distinguishing over overruling arises from the greater precision that new material dimensions bring into the law.

To fully evaluate the performance of distinguishing, we need to ask whether, regardless of the long-run efficiency of the law, there is a *tendency* for legal change to improve it. In Section 5, we saw that, when under-precautions are relatively costly, a *Pro-O* judge distinguishing a *Pro-V* precedent may reduce social welfare. Yet, to properly evaluate legal change, we must consider all possible paths of legal evolution. This leads to the following result:

Corollary 2: *Under distinguishing, the level of k that minimizes social losses accommodates some legal change. In particular,* $k = 0$ *is always socially preferable to* $k = \infty$.

This result stands in stark contrast to **Corollary 1**, which showed that, when judges behave symmetrically, under overruling the best outcome is achieved for $k = +\infty$. In that case, frequent legal change prevented the law from converging, leading to unpredictability.

Legal change is beneficial here because the introduction of q into the law brings an informational benefit that *on average* overpowers the cost of bias. How this can be the case is best seen by looking at expected social losses when $k = 0$. Now every judge distinguishes away the initial precedent A_i, leading to second period expected loss of $\left[A_i^2 + (1 - A_i)^2\right]E_j\left[\Lambda(A_j)\right]$, which – as we saw in **Proposition 5** – can be larger or smaller than the period 1 loss $\Lambda(A_i)$. Along some paths (e.g. starting from an efficient A_i) legal change is likely to be bad, but it is good along others.

Averaging the period 2 loss across initial precedents, we find the *ex-ante* loss:

$$E_i\left[A_i^2 + (1 - A_i)^2\right]E_j\left[\Lambda(A_j)\right] \tag{23}$$

The *ex-ante* loss under *no* legal change is instead $E_i\left[\Lambda(A_i)\right]$. Since $E_i\left[A_i^2 + (1 - A_i)^2\right] \leq 1$, legal change is in expectation beneficial and $k = 0$ is socially preferred to $k = \infty$. The fact that the benefit of legal change is more readily available when judges are activists yields our final result:

Proposition 7: *There exists a* $\bar{k}$ *>0 such that, for* $k \leq \bar{k}$ *distinguishing of precedents is on average beneficial.*

This result vindicates at least partly Cardozo's intuition for the presence of a "technological" force driving the evolution of precedent toward efficiency despite the vagaries of individual judges. When judges embrace legal change (as in the case of *k = 0*), their biases "wash out" and the net gain for the law comes from the more accurate information (greater number of empirical dimensions) embodied into legal rules. This creates a *tendency* for the law to become more efficient over time. This effect is absent when judges overrule prior precedents, since massive legal change brings no new data to dispute resolution, and is only a source of unpredictability.

7. Conclusion.

When does the evolution of judge-made law through precedent change lead to efficient legal rules? When does such evolution improve the law on average? We addressed these questions in a legal-realist model, in which deciding judges face opportunities to either overrule the precedent or distinguish it from the case before them, but may be both biased and averse to changing the law.

We found that, for both overruling and distinguishing, the conditions for ultimate efficiency of judge-made law are implausibly stringent. Indeed, with overruling, legal rules need not even converge. One key case of convergence to full efficiency under overruling occurs when efficiency-oriented judges are more activist than the biased ones, and the preferences of the biased judges are not too polarized. Yet even though full efficiency is hard to attain, in the case of distinguishing there is a strong presumption that the process of legal change raises welfare as it improves the informational quality of judicial decision making, at least when the cost of changing the law is low. We also found that disagreement among judges is likely to undermine the quality of legal change, suggesting that common law is likely to be least efficient in areas where judges strongly disagree.

The model in this paper is a first step in the analysis of judge-made law, and omits several important aspects of legal evolution. First, unlike the previous research, we focus on decision-making by judges, and neglect the selection of disputes for judicial resolution rather than settlement. It is far from clear, however, that such selection improves the quality of law, since it may be the combination of extremist litigants and judges that leads to legal change.

Second, we have ignored the important fact that judges make decisions in panels, which could in principle moderate polarization of their views, and lead to better law. However, as shown by Revesz (1997) and Sunstein et al. (2004), panels sometimes lead to the convergence of member views to the bias of the majority, rather than to a moderate compromise. Collective decision making does not then reduce polarization, so crucial to the efficiency of legal change.

Third, we have presented an extremely limited model of judicial leeways, in which only one verifiable material dimension can be added to the judicial consideration of a dispute. In reality, there are many such dimensions and, moreover, some of them include complex issues such as causality or knowledge. According to Stone (1985), the flexibility of language offers appellate judges tremendous leeway in distinguishing cases and rewriting the law. This leeway may offer considerable benefits when the law evolves toward efficiency, but it can also slow down legal change, or turn it in bad directions, when used by judges uninterested in efficiency.

Fourth, we have focused on judicial discretion in making new laws under the assumption that the facts of the case are verifiable. However, as argued by Frank (1930, 1951), Stone (1985) and Posner (1990), judges can also manipulate their interpretation of the facts, by emphasizing some aspects of the evidence and neglecting others, thereby reaching the outcomes they desire through fact-discretion rather than changes in the law. Such fact-discretion in itself many undermine the efficiency of the law, but is also likely to slow the pace of legal change, as judges choose to "work on" the facts rather than to rewrite precedents. For this reason, fact-discretion is one of the crucial challenges in the analysis of legal evolution.

As a final note, we emphasize that ours is a theoretical analysis of the propositions that legal change in a system or precedent is beneficial, and that the law converges to efficiency. Posner's hypotheses, however, are empirical propositions, and as such cannot be rejected by theory. We have tried to develop several testable implications of our analysis, which might make it possible to identify the areas of the law where Posner's hypothesis is more likely to hold. These hypotheses may be easier to verify empirically than the broad propositions about the efficiency of common law.

8. Proofs.

Proof of Proposition 1. The optimal one-dimensional threshold rule A_L is defined as

$$A_L = \arg\min_{A\in[0,1]} (1/2)A^2\Lambda^{under} + (1/2)(1-A)^2\Lambda^{over}$$

If $\Lambda^{over}, \Lambda^{under} > 0$, the objective function is convex and $A_L = (\Lambda^{over}/\Lambda^{under})/\left[1+(\Lambda^{over}/\Lambda^{under})\right]$ is found by solving the f.o.c. ($A_L\Lambda^{under} - (1-A_L)\Lambda^{over} = 0$). Notice that $A_L \in [0,1]$. The optimal two-dimensional threshold rule (A_F, Q_O, Q_1) is defined as

$$(A_F, Q_0, Q_1) = \arg\min_{A,Q_0',Q_1'\in[0,1]^3} (1/2)\left[(A+Q_0'-1)^2 + (Q_1')^2\right]\Lambda^{under} + (1/2)\left[(1-Q_0')^2 + (1-A-Q_1')^2\right]\Lambda^{over}$$

Again, $\Lambda^{over}, \Lambda^{under} > 0$ ensures that the above objective function is convex in $\left(A, Q_0^{'}, Q_1^{'}\right)$ (its Hessian is positive definite). Thus, solving the first order conditions for (A_F , Q_O , Q_1), namely

$$\partial\Lambda/\partial A = (A_F + Q_0 - 1)\Lambda^{under} - (1 - A_F - Q_1)\Lambda^{over} = 0$$

$$\partial\Lambda/\partial Q_0 = (A_F + Q_0 - 1)\Lambda^{under} - (1 - Q_0)\Lambda^{over} = 0$$

$$\partial\Lambda/\partial Q_1 = Q_1\Lambda^{under} - (1 - A_F - Q_1)\Lambda^{over} = 0$$

yields $A_F = 1/2$, $Q_O = (1 + A_L)/2$, $Q_1 = A_L/2$. Notice that (A_F , Q_O , Q_1) $\in [0,1]^3$.♠

Proof of Proposition 2. Judge *j* overrules A_i with A_j when $U_j(A_j) - U_j(A_i) \geq k$. A_j minimizes the expression for social losses where $(\beta_{V,j}, \beta_{O,j})$ replaces $(\Lambda^{under}, \Lambda^{over})$. Judge *j* overrules when

$$f_{j,i} \equiv \frac{\left(\beta_i - \beta_j\right)^2}{(1+\beta_i)^2(1+\beta_j)^2} \geq 2k$$

Under these judicial preferences, judicial behavior is symmetric: if π is such that *j* overrules A_i it is also the case that *i* overrules A_j. Such symmetric behavior in turn implies that unless $\gamma = 1$ or $\pi = 1$ (i.e., all judges are unbiased), there exists no *k* such that precedent converges to A_L.♠

Proof of Corollary 1. Let us analyze judges´ behavior by studying the functions $f_{i,j}(\pi)$. Their symmetric behavior considerably restricts the number of cases that we need to look at.

a) *Pro-O (* $\beta_i = \pi\lambda$ *); Pro-V (* $\beta_i = \lambda/\pi$ *).* Now $f_{O,V}(\pi) = f_{V,O}(\pi) = \dfrac{\lambda^2(\pi^2 - 1)^2}{(1+\pi\lambda)^2(\pi+\lambda)^2}$.

b) *Pro-O (* $\beta_i = \pi\lambda$ *); Unbiased (* $\beta_i = \lambda$ *).* Now $f_{O,U}(\pi) = f_{U,O}(\pi) = \dfrac{\lambda^2(\pi - 1)^2}{(1+\pi\lambda)^2(1+\lambda)^2}$.

c) *Pro-V (* $\beta_i = \lambda/\pi$ *); Unbiased (* $\beta_i = \lambda$ *).* Now $f_{V,U}(\pi) = f_{U,V}(\pi) = \dfrac{\lambda^2(\pi - 1)^2}{(\pi+\lambda)^2(1+\lambda)^2}$.

The function $f_{j,i}(\pi)$ increases in π, starting at $f_{i,j}(1) = 0$. Moreover, $f_{O,V}(\infty) = 1$, $f_{O,U}(\infty) = 1/(1+\lambda)^2$, $f_{V,U}(\infty) = \lambda^2/(1+\lambda)^2$. Call $\pi_{j,i} \in [1,+\infty) \cup \{+\infty\}$ the level of π above which *j* overrules *i* (and vice-versa). We define $\pi_{i,j} = +\infty$ when there exist no π such that *i* and *j* overrule each other (for instance $\pi_{U,O} = +\infty$ if $k > 1/(1+\lambda)^2$). Notice that for $\lambda \leq 1$, $\pi_{O,V} \leq \pi_{O,U} \leq \pi_{V,U}$. If $k > 1/2$, precedent does not change from its initial configuration; otherwise there will be legal change provided π is large enough and the law does not converge. At $k = 0$, judges always overrule precedents. In this case, expected social losses are given by

$E_i[\Lambda(A_i)] \equiv \alpha_V \Lambda(A_V) + \alpha_B \Lambda(A_B) + \alpha_O \Lambda(A_O)$, where A_V, A_B, A_O are the preferred rules of *Pro-V*, *Benevolent* and *Pro-O* judges, respectively. If $k \to +\infty$, the initial precedent sticks forever, and expected social losses are $E_i[\Lambda(A_i)]$. Thus, $k = 0$ and $k = +\infty$ are equally desirable from a welfare standpoint. This is also the case for intermediate *k's*, as the symmetry in judicial behavior makes different patterns of legal change irrelevant for social welfare: the effect of judge *j* on law A_i set by judge *i* is exactly neutralized by the effect of judge *i* on the law set by judge *j* (A_j). Finally, for the purpose of Table 1 and its discussion, re-label $\pi_1 = \pi_{O,V}, \pi_2 = \pi_{O,U}, \pi_3 = \pi_{V,U}$.♠

Proof of Proposition 3. Efficiency is attained if *Unbiased* judges set A_L without being overruled by extremists. This happens if $\pi < \pi_{O,U}$, which is necessary for A_L to stay in place. The law can converge to A_L either if *Unbiased* judges overrule all extremists, i.e. when $\pi \geq \overline{\pi}_{V,U}$ where $f_{V,U}(\overline{\pi}_{V,U}) = k_U$, or if *Unbiased* judges overrule only *Pro-O* ones, but at the same time *Pro-O* overrule *Pro-V*. This event happens when $\pi \geq \overline{\pi}_O \equiv \max[\pi_{O,V}, \overline{\pi}_{O,U}]$ ($f_{U,O}(\overline{\pi}_{O,U}) = k_U$). Call $\hat{\pi} = \min[\overline{\pi}_O, \overline{\pi}_{V,U}]$. Then, if $\pi \in [\hat{\pi}, \pi_{O,U})$, the law converges to A_L. Clearly, it is always the case that $\hat{\pi} \leq \pi_{O,U}$. In order to draw figure 4, suppose that $0 < k < 1/2(1+\lambda)^2$, k fixed. Define now the variable $k_U / k \in [0,1]$. $\hat{\pi}(k_U / k)$ starts at $\hat{\pi}(0) = 1$, ends at $\hat{\pi}(1) = \pi_{O,U} > 1$ and it increases continuously in k_U / k.♠

Proof of Proposition 4. We say that judge 2 is more moderate under $A_1^{'}$ than under A_1 if, given his bias β_2 – fully translated in the ratio between different errors under the two-dimensional rule – the imprecision of the law (i.e. the sum of different errors) is smaller under $A_1^{'}$. Judge 2´s moderation is maximized at $A_1 = (1/2)$, when judge 1 is maximally moderate, and minimized at $A_1 = 0$ or $A_1 = 1$, (when judge 1 is fully biased).♠

Proof of Proposition 5. After A_1, judge *2* minimizes the expression for social losses where $(\beta_{V,j}, \beta_{O,j})$ replace $(\Lambda^{under}, \Lambda^{over})$ for given A_1. Under the new (two-dimensional) rule, expected social losses are $\Lambda(A_1, Q_{0,2}, Q_{1,2}) = [A_1^2 + (1-A_1)^2]\Lambda(A_2)$, as opposed to the level $\Lambda(A_1)$ under A_1. Proposition 5 follows directly by comparing these two magnitudes.♠

Proof of Proposition 6. Judge *j* distinguishes precedent A_i by introducing *q* into the law when $U_j(A_i, Q_{0,j}, Q_{1,j}) - U_j(A_i) \geq k$. $Q_{0,j}, Q_{1,j}$ minimize, for given A_i, the expression for social losses where $(\beta_{V,j}, \beta_{O,j})$ replaces $(\Lambda^{under}, \Lambda^{over})$. Then, *j* distinguishes A_i when

$$h_{j,i}(\pi) \equiv \frac{\beta_i^2 + \beta_j^2}{(1+\beta_i)^2(1+\beta_j)^2} \geq 2k$$

Notice that judicial behavior is symmetric: if π is such that *j* distinguishes A_i, then it is also the case that *i* distinguishes A_j. By the logic of Proposition 2, such symmetry implies that unless $\gamma = 1$, the bias of the law is not efficient. With distinguishing, $\gamma = 1$ is not enough for efficiency. If $\gamma = 1$ and *k=0* the law converges to $A = \lambda/(1+\lambda), Q_0 = 1 - Q_1, Q_1 = \lambda/(1+\lambda)^2$, which is efficient only if $\lambda = 1$. Finally it must be that judges introduce *q* into the law, i.e. that $k \leq \tilde{k} \equiv (1/2)h_{U,U}$. ♠

Proof of Corollary 2. Let us analyze judges´ behavior by studying the functions $h_{i,j}(\pi)$. Given symmetry, we must only consider the following cases

a) *Pro-O ($\beta_i = \pi\lambda$); Pro-O ($\beta_i = \pi\lambda$).* Now $h_{O,O}(\pi) \equiv 2\dfrac{\pi^2\lambda^2}{(1+\pi\lambda)^4}$.

b) *Pro-V ($\beta_i = \lambda/\pi$); Pro-V ($\beta_i = \lambda/\pi$).* Now $h_{V,V}(\pi) \equiv 2\dfrac{\lambda^2}{(\pi+\lambda)^4}$.

c) *Unbiased ($\beta_i = \lambda$); Unbiased ($\beta_i = \lambda$).* Now $h_{U,U}(\pi) \equiv 2\dfrac{\lambda^2}{(1+\lambda)^4}$.

d) *Pro-O ($\beta_i = \pi\lambda$); Pro-V ($\beta_i = \lambda/\pi$).* Now $h_{O,V}(\pi) \equiv h_{V,O}(\pi) \equiv \dfrac{\lambda^2(\pi^4+1)}{(1+\lambda\pi)^2(\pi+\lambda)^2}$.

e) *Pro-O ($\beta_i = \pi\lambda$); Unbiased ($\beta_i = \lambda$).* Now $h_{O,U}(\pi) \equiv h_{U,O}(\pi) \equiv \dfrac{\lambda^2(\pi^2+1)}{(1+\pi\lambda)^2(1+\lambda)^2}$.

f) *Pro-V ($\beta_i = \lambda/\pi$); Unbiased ($\beta_i = \lambda$).* Now $h_{V,U}(\pi) \equiv h_{U,V}(\pi) \equiv \dfrac{\lambda^2(1+\pi^2)}{(\pi+\lambda)^2(1+\lambda)^2}$.

If $j \neq i$, $h_{j,i}(\pi)$ is increasing, $h_{j,i}(1) = h_{U,U}$, $h_{O,V}(\infty) = 1$, $h_{O,U}(\infty) = 1/(1+\lambda)^2$, $h_{V,U}(\infty) = \lambda^2/(1+\lambda)^2$. $h_{j,j}(\pi)$ is decreasing, $h_{j,j}(1) = 2\lambda^2/(1+\lambda)^2$, $h_{O,O}(\infty) = h_{V,V}(\infty) = 0$. Some rankings in the $h_{j,i}(\pi)$ are: $h_{V,U}(\pi) = \min_{j\neq i} h_{j,i}(\pi)$, $h_{V,V}(\pi) = \min_j h_{i,i}(\pi)$, $h_{O,U}(\pi) \geq \max_i h_{i,i}(\pi)$. Disagreement tends to be a stronger incentive for distinguishing than information ($\max_{j\neq i} h_{i,j}(\pi) \geq \max_i h_{i,i}(\pi)$). As π gets sufficiently large (and for $\lambda = 1$), the ranking becomes $h_{O,V} \geq h_{O,U} \geq h_{V,U} \geq h_{U,U} \geq h_{O,O} \geq h_{V,V}$. In the spirit of **Proposition 2**, define $\overline{\pi}_{j,i} \in [1,+\infty) \cup \{+\infty\}$ (above $\overline{\pi}_{j,i}$ j distinguishes i and vice-versa. Under the above ranking $\overline{\pi}_{O,V} \le \overline{\pi}_{O,U} \le \overline{\pi}_{V,U} \le \overline{\pi}_{U,U} \le \overline{\pi}_{O,O} \le \overline{\pi}_{V,V}$ (for every k). By reducing k social welfare need not go up. If k is set such that at the current level of π, $\overline{\pi}_{O,U} < \pi < \overline{\pi}_{V,U}$, social losses are $\hat{L} = (1/2)(1-\gamma)\theta_O [x_U \Lambda_U + x_V \Lambda_V] + \gamma\theta_U \Lambda_O + (1/2)(1-\gamma)\theta_V \Lambda_O$, with $\theta_i = A_i^2 + (1-A_i)^2$, $\Lambda_i = \Lambda(A_i)$, $x_U = 2\gamma/(1+\gamma), x_V = (1-\gamma)/(1+\gamma)$. We could have $\hat{L} \ge E_i(\Lambda_i)$, i.e. legal change reduces welfare. Yet, for some k legal change raises welfare. If $\overline{\pi}_{O,V} < \pi < \overline{\pi}_{O,U}$ legal change is good, as $(1/2)(1-\gamma)\theta_O \Lambda_V + \gamma\Lambda_U + (1/2)(1-\gamma)\theta_V \Lambda_O \le E_i(\Lambda_i)$. Clearly, $k = 0$ is preferred to $k = \infty$, as $E_i(\theta_i)E_i(\Lambda_i) \le E_i(\Lambda_i)$. For the discussion in the text, relabel $\overline{\pi}_{O,V} = \tilde{\pi}_1; \overline{\pi}_{O,U} = \tilde{\pi}_2$.♠

Proof of Proposition 7. Define $\overline{k} = h_{i,j}(1) = 2\lambda^2/(1+\lambda)^4$. If $k \le \overline{k}$, judges $j \ne i$ and $j = i = U$, distinguish A_i (they would do it at $\pi = 1$). For $j = i$, $i = O,V$ judges can either be active or not. If *Pro*-*O* become active welfare goes down ($\Lambda_O \ge E(\Lambda_i)$). For $k \le \overline{k}$, legal change is good if it is good when only *Pro*-*O* are active. If $\Lambda_V \ge E(\Lambda_i)$ (i.e. the activism of *Pro*-*V* reduces welfare), then legal change is good for $k \le \overline{k}$, as adding to the activism of *Pro*-*O* the harmful one of *Pro*-*V* leads to the same as $k = 0$. If $\Lambda_V < E(\Lambda_i)$ one finds after some algebra that legal change is good if: $[(1/2)(1-\gamma)\theta_O + \gamma\theta_U + (1-\gamma)\theta_V - 1]E(\Lambda_i) \le (1/2)(1-\gamma)^2 \theta_V \Lambda_V$. This certainly holds if legal change is good when $\Lambda_V = E(\Lambda_i)$. This is indeed the case because if $\Lambda_V = E(\Lambda_i)$, social losses are the same as under $k = 0$.

Bibliography.

Bueno de Mesquita, Ethan, and Mathew Stephenson, 2002, "Informative Precedent and Intrajudicial Communication," *American Political Science Review* 90 (4), 1-12.

Cardozo, Benjamin, 1921, *The Nature of the Judicial Process*, New Haven: Yale University Press.

Cohen, Felix, 1935, Transcendental Nonsense and the Functional Approach, *Columbia Law Review* 35, 809-849.

Frank, Jerome, 1930, *Law and the Modern Mind*, New York: Brentano's.

Frank, Jerome, 1951, *Courts on Trial*, Princeton, N.J.: Princeton University Press.

Gennaioli, Nicola, 2004, Contracting in the Shadow of the Law, Mimeo, Harvard University.

Glaeser, Edward, Jose Scheinkman, and Andrei Shleifer, 2003, The Injustice of Inequality, *Journal of Monetary Economics* 50, 199-222.

Glaeser, Edward, and Andrei Shleifer, 2002, Legal Origins, *Quarterly Journal of Economics* 117, 1193-1230.

Haslem, Bruce, 2004, Managerial Opportunism During Corporate Litigation, *Journal of Finance*, forthcoming.

Hayek, Friedrich, 1960, *The Constitution of Liberty*, Chicago: the University of Chicago Press.

Holmes, Oliver Wendell, 1897, The Path of the Law, *Harvard Law Review* 10, 457-478.

Landes, William, and Richard Posner, 1987, *The Economic Structure of Tort Law*, Cambridge, MA: Harvard University Press.

Partridge, Anthony, and William Eldridge, 1974, *The Second Circuit Sentencing Study*, Federal Judicial Center.

Pinello, Daniel, 1999, Linking Party to Judicial Ideology in American Courts: A Meta-analysis, *The Justice System Journal* 3, 219-254.

Posner, Richard, 1990, *Cardozo: A Study in Reputation*, Chicago: the University of Chicago Press.

Posner, Richard, 2003 (1^{st} ed. 1973), *The Economic Analysis of Law*, 6^{th} ed., Boston: Little Brown.

Posner, Richard, 2004a, Judicial Behavior and Performance: an Economic Approach, Mimeo.

Posner, Richard, 2004b, The Law and Economics of Contract Interpretation, Mimeo.

Priest, George, 1977, The Common Law Process and the Selection of Efficient Rules, *Journal of Legal Studies* 6, 65-82.

Radin, Max, 1925, The Theory of Judicial Decision: or How Judges Think, *American Bar Association Journal* xx, 357-362.

Revesz, Richard, 1997, Environmental Regulation, Ideology, and the D.C. Circuit, *Virginia Law Review* 83, 1717-1772.

Rowland, C.K., and Robert Carp, Politics and Judgment in Federal District Courts, Lawrence: the University Press of Kansas.

Rubin, Paul, 1977, Why is the Common Law Efficient?, *Journal of Legal Studies* 6, 51-63.

Stone, Julius, 1985, *Precedent and the Law: Dynamics of Common Law Growth*, Sydney: Butterworths.

Sunstein, Cass, David Schkade, and Lisa Michelle Ellman, 2004, Ideological Voting on Federal Courts of Appeals: a Preliminary Investigation, *Stanford Law Review* 90, 301-354.