

JUDICIAL HIERARCHIES AND THE RULE-INDIVIDUAL TRADEOFF

HUGO M. MIALON

PAUL H. RUBIN

JOEL L. SCHRAG¹

ABSTRACT. We analyze decision-making in a simple model of the judicial hierarchy. We assume that trial court judges are more concerned with ex post efficiency with respect to the individuals involved in the cases at hand, and less concerned with ex ante efficiency with respect to the precedents established for society, than are appeals court judges. This implies that the preferred decisions of appeals court judges differ systematically from those of trial court judges. Appeals court judges can enforce their preferred decisions by reversing those of the trial court judges. However, in the model, litigants do not always appeal decisions that would be reversed, both because appeals are costly and because the outcome is uncertain. Consequently, appeals court judges may prefer to enact higher level rules that reduce the discretion of all judges. JEL K1 K4.

August 12, 2005

¹ Mialon and Rubin, Department of Economics, Emory University, Atlanta, GA 30322-2240 (hmialon@emory.edu and prubin@emory.edu); Schrag, Bureau of Economics, Federal Trade Commission, 600 Pennsylvania Avenue NW, Washington, DC 20580 (jschrag@ftc.gov). The authors would like to thank Mark Cohen, Bruce Johnsen, John Lott, Fred McChesney, Richard Posner, and participants at a seminar at George Mason Law School, for helpful comments. The views expressed in this paper are the authors' alone and do not necessarily reflect the opinions of the Federal Trade Commission or of any individual Commissioner.

JUDICIAL HIERARCHIES AND THE RULE-INDIVIDUAL TRADEOFF

I. INTRODUCTION

Analysis of the behavior of judges has proven to be one of the most intractable problems facing law and economics scholars.² This is because opinions are the result of utility maximizing behavior and economists are unable to specify the details of judges' utility functions. Consequently, it has proven difficult to develop a theory explaining why judges make the decisions that we observe, though, as we discuss below, many have tried.

In this paper we argue that a judge's preferred decisions depend at least in part on the judge's location in the hierarchical judicial system. We develop this argument by means of a simple model of the judicial hierarchy in which litigants first take their dispute to a trial court judge, who examines the facts and produces a decision. The litigants then have an opportunity to appeal this decision to a higher court, which either upholds or reverses the lower court's ruling. A central assumption in the model is that the trial court and appeals court judges face different tradeoffs. A trial court judge must weigh the interests of the litigants before him against the interests of the other people in his jurisdiction who will be affected by the precedent that he establishes. An appeals court judge who reviews the lower court judge's decision faces a different tradeoff: she must balance the interests of the litigants against the interests of the people in her entire jurisdiction, which is typically larger than the trial court judge's jurisdiction. The appeals court judge's decision affects more people.

The model predicts that even if both judges have identical preferences, in the sense that they value the effect of their decision on an individual person in the same way, they will still rule differently because the effects of their decisions are different. The trial judge will tend to make decisions that are sympathetic to the interests of the individual litigants, while the appellate judge will give more weight to the interests of the people outside the courtroom who are affected by the precedents that are established. We characterize the tradeoff as between the *individuals* who brought the case and the *rule* that protects the interests of those people who are not directly connected to the case. The prediction is that the appellate judge attaches more weight to rules—and less to individuals—than the trial court judge.

For this reason, the appellate judge will sometimes want to reverse the trial judge's decision. However, her power in this domain is limited because she must rely on the litigants to select cases for appeal. The model assumes that appeals are costly and that the litigants cannot perfectly anticipate whether the trial judge's decision will be reversed. These assumptions imply that the litigants do not in general appeal every decision that the appellate judge would reverse. The appellate judge's ability to create rules that reflect the interests of a broad population is thus itself constrained. Since the higher court may want to reverse the lower court's decision, but may not have the opportunity to do so because litigants do not always appeal decisions that the higher court would reverse, the higher court may prefer to establish a *higher level rule* that limits the trial court judge's ability to choose a decision that is at odds with the appeals court's preferred decision. An example of this occurs in antitrust. The higher level court can promulgate a rule

² “At the heart of economic analysis of law is a mystery that is also an embarrassment: how to explain judicial behavior in economic terms...” (Posner, 1993, at 2).

stating that all vertical restrictions are *per se* illegal, thus limiting the lower court's discretion to decide whether any particular restriction should be permitted.

In the next section we review related literature. In section III, we develop our ideas about the rule-individual tradeoff in more detail. In section IV, we discuss the hierarchical nature of the law. In section V, we describe the model and discuss its equilibrium. In section VI, we identify further implications. Section VII concludes.

II. RELATED LITERATURE

There have to date been several attempts to explain judges' behavior. Posner (1992) has argued that judges seek efficiency, and, more recently (since becoming one), that they seek the pleasures of spectators at plays (Posner, 1993). Whitman (2000), Miceli and Cosgel (1994), Rasmusen (1994), and Kornhauser (1992a,b) have argued that judges' behavior is based on a tradeoff between writing decisions they prefer and the possibility of reversal, either by higher courts or by future judges. Kobayashi and Lott (1994) have argued that judges will want to maximize litigation, and therefore will seek inefficient rules. Except for Posner's argument regarding efficiency and Kobayashi and Lott's argument for inefficiency, these papers do not predict the actual substantive outcome of the judicial process.

Higgins and Rubin (1980), Cohen (1991, 1992), and Levy (2004) explain judicial behavior as an attempt to seek promotion, but the power of these models is limited because circumstances where they can apply are limited. Cohen showed that federal judges seeking promotion were more likely to impose harsher sentences on antitrust defendants in order to seek

a favorable recommendation from the Justice Department, and that judges' decisions on the constitutionality of the Sentencing Commission could be explained by utility maximizing behavior. These results do not generalize in an obvious way to explain more comprehensive aspects of judicial behavior. O'Hara (1993) analyzes why appellate judges choose to follow precedents in the absence of any formal rule mandating adherence to precedent. She explains this behavior as the result of a repeated game between different appellate judges, each of whom follows the precedents established by her colleagues' decisions to avoid triggering a "punishment" phase, in which the precedents established by the errant judge's decisions are ignored by her colleagues.

There are also several papers that have explicitly modeled the judicial hierarchy, and analyzed its efficiency. Spitzer and Talley (2000) study judicial review through the lens of an incomplete information game between a lower court judge, who chooses a ruling, and a higher court judge, who then has the opportunity to audit the lower court's ruling. The choices of litigants are not modeled. In contrast, we model litigant behavior, finding that the higher court cannot always count on litigants to appeal lower court decisions that would be reversed. Shavell (1995, 2004) argues that an appellate process is more efficient than random monitoring of lower court decisions, because it corrects errors by harnessing the private information of litigants, and does so at a low cost since only a subset of the cases are appealed. Cameron and Kornhauser (2004) argue that a three-tier hierarchy (trial court, appellate court, and Supreme Court) is more efficient at correcting errors than a two-tier hierarchy. These papers explain the appellate process in terms of error correction. This differs from our explanation in terms of lawmaking. We argue

that the higher court cares more about establishing ex ante efficient rules, while the lower court cares more about ex post efficiency with respect to the individuals directly involved in the cases.

Daughety and Reinganum (2000) also analyze lawmaking in the appellate process, but their approach is different. In their model, the trial court determines the facts (according to a set of axioms for aggregating evidence), while the appeals court interprets the law (in a Bayesian fashion), choosing its interpretation to conform as closely as possible to that of a yet higher court (the Supreme Court), which knows the correct interpretation of the law. Their model provides, inter alia, an explanation for the legal notion of “harmless error” by a trial court.

III. INDIVIDUALS VERSUS RULES

In many legal disputes, there is a conflict between efficiency with respect to the parties actually involved in the particular dispute at issue and efficiency with respect to the decision rule. That is, a decision rule that is efficient ex ante will often not be efficient ex post. This is a natural result of sunk costs: the costs of efficient accident avoidance, for example, are sunk after an accident has occurred and so are not relevant ex post. Epstein (1993, at 557-8) suggests that this situation is common: “It will therefore be a frequent occurrence that the ideal set of rules, measured from the ex ante perspective, will often produce a set of outcomes that will be regretted by the participants to the transaction ex post.” However, a decision in one case, while ex post with respect to that case, will be ex ante with respect to all succeeding cases.

Consider a contractual situation where A is the least cost avoider of breach, so that if A is liable ex ante the risk of breach is minimized. However, B is the least cost insurer if breach

occurs. Ex ante costs are minimized by making A liable because the savings from avoiding inefficient breach by placing liability on A are greater in total than the savings from efficiently adapting to breach by placing liability on B. But ex post, once breach has occurred, it will be cheaper to place liability on B because B can most efficiently bear this cost. Thus, an efficient rule would hold A liable, but efficiency with respect to the parties after breach has occurred (that is, after there is a legal dispute) would hold B liable.

Consider a tort where the injured is ex ante the least cost preventer. However, if an injury occurs, then the marginal utility of wealth for the victim increases to compensate for medical costs, lost wages, and other pecuniary losses (but not to compensate for “pain and suffering”). Thus, an efficient rule would hold the victim liable in order to provide proper incentives for care in the future. On the other hand, efficiency with respect to the parties in the particular case would find the injurer liable in order to provide efficient compensation.

Consider a patent. Once something has been invented, there is no benefit from patent protection, and some social costs. Nonetheless, the patent laws provide large social benefits ex ante in that they create incentives for future inventions. The tradeoff in a patent case is purely an ex ante-ex post tradeoff, with enforcement providing ex ante benefits but ex post costs. A similar analysis would apply in some vertical antitrust cases between the upstream and downstream producers, as in a franchise case. The benefit of enforcing the franchise contract is in terms of future franchisees and other potential franchisers; in many cases, short run efficiency would require not enforcing the contract.

A rule oriented decision indicates a concern for a greater number of people than an individual oriented decision. And since individuals are affected by rulings immediately and rules

take longer to have an impact, an individual oriented decision also indicates a higher discount rate (shorter time horizon) than a rule oriented decision.³ Within the judicial system, judges at differing levels may be expected to differ systematically with respect to these dimensions. We now turn to the hierarchical nature of legal rules.

IV. LAW AS HIERARCHY

There is another sense in which the legal system is hierarchical. Rules as well as courts are hierarchical. Thus, in order to change a given legal rule it will often first be necessary to change some higher level rule. Indeed, this hierarchy of rules is what allows a hierarchy of courts to exist: a higher level court will change a higher level rule and then lower level courts will choose lower level rules to implement the change. Because of this hierarchical structure, each decision of a higher level court will be more influential than each decision of a lower level court. The hierarchies of rules and of courts are related. Higher level courts will announce higher level rules, and lower level courts will interpret these rules in particular cases. The higher level rules are more general than the rules announced by lower level courts, and each higher level decision will, through the establishment of a precedent binding on many lower level rules and courts, affect more individuals than will each lower level rule (Posner, 1993).

As an example, consider vertical restrictions in antitrust. A higher level decision is whether to treat these relationships under a rule of reason or as per se illegal. If the courts go from treating such relationships as per se illegal to treating them under the rule of reason, then

³ Note that we do not argue that the lower level judge is influenced by seeing the actual parties to the dispute.

many subsidiary lower level decisions are required. For each type of vertical relationship, a determination is necessary: Is this particular relationship per se or rule of reason? (Vertical price fixing of minimum prices is still per se illegal; exclusive distribution territories are viewed under a rule of reason.) If it is decided that some relationship is to be treated under the rule of reason, then further even lower level decisions are required: What are the standards for illegality for this relationship under the rule of reason? Each rule becomes a constraint for the lower level rules. Trial courts then make the lowest level decision: Did this firm in this case violate this rule?

The process of legal change at the lower court level is the working out of the implications of the new higher court rule. This process will be time consuming and the results may not be ex ante obvious. At each level, parties to the dispute can force the higher level court to check that the rulings of the lower level court are consistent by appealing the decision. The higher level court can indicate that the lower level court is correct, by affirming, or that the lower level court has misinterpreted its rules, by reversing.

V. THEORETICAL MODEL

To understand the conflict that arises in the judicial hierarchy, we analyze the choices of both a trial court judge and the appeals court judge who reviews the trial court's decisions. The trial court judge J_T decides the cases that arise in his jurisdiction, which corresponds to the geographic area where his rulings serve as precedents unless they are overturned on appeal. After J_T rules in a case, the losing litigant chooses whether to appeal the decision to the appeals court judge J_A , who either upholds or reverses the decision of the lower court. The decision by J_A

serves as a precedent in all jurisdictions over which she has authority, including the jurisdiction in which the case arose.

In period 1, a case arises and the litigants present their evidence to J_T , who examines the facts and determines the strength of the plaintiff's case.⁴ J_T 's assessment of the strength of the plaintiff's case is summarized by the variable e . Higher values of e mean that J_T believes that the defendant has a stronger case or, equivalently, that the plaintiff has a weaker case. We assume that both J_T and J_A interpret the facts of the case in the same way, so e represents their common assessment of the merits of the case. We make this assumption to avoid the trivial conclusion that there is a conflict among judges simply because they disagree about what the facts mean. While judges may disagree in some instances, such an explanation provides no systematic predictions about how judges' choices will differ.

We suppose that the litigants have incomplete information about e . They have common prior beliefs about e , which are represented by the probability distribution function $g(e)$ defined on the interval $[0, \infty)$. This distribution is common knowledge. We assume that litigants observe a common signal of e , which we denote by s . The signal s is distributed on the interval $[0, \infty)$ according to the density function $f(s|e)$. We assume that this density satisfies the monotone likelihood ratio property, so that the ratio $f(s|e')/f(s|e'')$ is increasing in s if and only if $e' > e''$. This assumption means that a high signal s is more likely if the true value of the judge's assessment of the strength of the case e is high. Given their prior and signal, litigants form

posterior beliefs $g(e|s) = \frac{g(e) f(s|e)}{\int_0^{\infty} g(e) f(s|e) de}$ according to Bayes' rule.

In period 2, after observing the evidence e , J_T makes a decision $d \in \{y, n\}$, that is, decides whether ($d = y$) or not ($d = n$) to accept the plaintiff's argument. The judge makes the decision to maximize his utility. While the judge's welfare is without doubt a function of many factors, our simplifying assumption is that his utility only depends on his evaluation of the effect of his decision on the welfare of the litigants who are directly involved in the case and the welfare of the rest of the population that will be affected by the precedent established by the decision. Furthermore, we assume that J_T 's utility is separable in these two effects and therefore can be summarized by the following utility function:

$$(1) \quad u_T(e, d) = u_L(e, d) + u_P(e, d),$$

where $u_L(e, d)$ is J_T 's valuation of the effect of his decision on the litigants' welfare, and $u_P(e, d)$ is his valuation of the effect of the precedent established by his decision, given the evidence e .

We make the following assumptions about J_T 's preferences:

$$(A1) \quad \frac{\partial u_L(e, y)}{\partial e} < 0, \quad \frac{\partial u_P(e, y)}{\partial e} \leq 0$$

$$(A2) \quad \frac{\partial u_L(e, n)}{\partial e} > 0, \quad \frac{\partial u_P(e, n)}{\partial e} \geq 0$$

Assumption A1 means that J_T receives a lower payoff from accepting the plaintiff's argument when e is higher, both because his valuation of the litigants' benefit from this ruling decreases in e —corresponding to the idea that higher values of e are stronger evidence in favor of the defendant's case—and because the precedential value of the ruling is non-increasing in e .⁵

⁴ To simplify the model we do not analyze the defendant's decision to commit the infraction or the plaintiff's decision to sue the defendant, nor do we consider the possibility that the litigants will settle their case. These issues are not central to our main question.

⁵ The assumption that $\partial u_P(e, y)/\partial e \leq 0$ permits the possibility that the precedential value of ruling in favor of the plaintiff does not depend on the evidence specific to any particular case.

Assumption A2 means that the judge receives a higher payoff from rejecting the plaintiff's argument when e is higher, for analogous reasons.

Because J_T 's utility function is strictly monotone in e for each of the two possible decisions that he could make, his problem is equivalent to choosing a "cutoff" value of e , say e^* , such that for all $e \leq e^*$ the judge accepts the plaintiff's argument and for all $e > e^*$ the judge rejects it. In the following lemma we describe the solution to J_T 's problem, corresponding to his best reply to the other players' actions.

Lemma 1. J_T 's best reply e^* satisfies:

$$\begin{aligned} &< 0, e^* = 0 \\ u_L(e^*, y) + u_P(e^*, y) - u_L(e^*, n) - u_P(e^*, n) &= 0, e^* \in [0, \infty). \\ &> 0, e^* = \infty \end{aligned}$$

Proof. Proofs of each of the formal statements are presented in the appendix.

We assume that J_T chooses the e^* given in Lemma 1 if he is indifferent about what to choose because the litigants always appeal his decision.

In period 3, after J_T observes e and announces his decision, the losing litigant chooses an action $a \in \{p, q\}$, that is, chooses whether to proceed with a costly appeal ($a = p$) or quit ($a = q$). If he chooses to proceed, then both litigants pay a cost c and the appeals court judge J_A hears the case. The losing litigant's incentive to appeal depends on the likelihood that J_A will reverse J_T 's decision. The plaintiff receives a payoff of $\pi_p > 0$ when the case is decided in his favor and zero

otherwise, while the defendant receives a payoff of $\pi_D > 0$ when the case is decided in his favor and zero otherwise. These payoffs do not include the cost of an appeal, c .

Whether J_A reverses J_T 's decision depends on J_A 's preferences, to which we now turn. As before, we suppose that J_A 's utility only depends on his evaluation of the effect of his ruling $d \in \{y, n\}$ on both the litigants' welfare and the welfare of the rest of the population that will be affected by the precedent established by the decision. Now, however, the judge's decision affects not only the population of the jurisdiction in which the case arose, but also the population of all jurisdictions over which she has authority. Suppose that there are $N > 1$ identical jurisdictions. Therefore, J_A 's utility function is given by:

$$(2) \quad u_A(e, d) = u_L(e, d) + Nu_p(e, d).$$

Here, our assumption is that J_A and J_T value the effect of a decision on the litigants and the rest of the population in the same way. Their utility only differs because J_A must account for the effect of her decision on more people; J_A has the opportunity to affect more people's welfare than J_T .

As in J_T 's case, J_A 's problem is equivalent to choosing a "cutoff" value of e , say \hat{e} , such that for all cases in which the evidence $e \leq \hat{e}$, J_A rules in favor of the plaintiff and for all cases in which the evidence $e > \hat{e}$, J_A rules in favor of the defendant. In the following lemma we describe the solution to J_A 's problem, corresponding to his best reply \hat{e} .

Lemma 2. J_A 's best reply \hat{e} satisfies:

$$\begin{aligned} &< 0, \hat{e} = 0 \\ u_L(\hat{e}, y) + Nu_p(\hat{e}, y) - u_L(\hat{e}, n) - Nu_p(\hat{e}, n) &= 0, \hat{e} \in [0, \infty). \\ &> 0, \hat{e} = \infty \end{aligned}$$

There will be a conflict between J_T and J_A whenever $\hat{e} \neq e^*$, as is usually the case. To see this, we rewrite the condition in Lemma 2 as:

$$(3) \quad \begin{array}{l} < 0, \hat{e} = 0 \\ u_L(\hat{e}, y) + u_P(\hat{e}, y) - u_L(\hat{e}, n) - u_P(\hat{e}, n) + (N-1)(u_P(\hat{e}, y) - u_P(\hat{e}, n)) = 0, \hat{e} \in [0, \infty) . \\ > 0, \hat{e} = \infty \end{array}$$

Comparing with the condition in Lemma 1, which summarizes the solution of J_T 's problem, it follows that the first four terms in (3) sum to zero when evaluated at $e^* \in (0, \infty)$. Therefore, we will have $\hat{e} \neq e^*$ whenever $u_P(e^*, y) - u_P(e^*, n) \neq 0$. In the following proposition we describe how \hat{e} and e^* are related.

Proposition 1. $u_P(e^*, y) - u_P(e^*, n) > 0 \Rightarrow \hat{e} \geq e^*$ and $u_P(e^*, y) - u_P(e^*, n) < 0 \Rightarrow \hat{e} \leq e^*$.

The proposition states that J_A will more inclined to rule in favor of the plaintiff than J_T ($\hat{e} \geq e^*$) when the interests of people other than the litigants are served by ruling in the plaintiff's favor. The appellate judge affects a larger population through her decision, and therefore she is less attentive to the litigants' interests; she cares more about establishing a rule that reflects the interests of people other than the litigants.

We now derive the litigants' best replies. After J_T decides whether or not to accept the plaintiff's argument the litigants use Bayes' rule to update their beliefs about the true value of e . After observing $d \in \{y, n\}$, the litigants' updated beliefs are:

$$(4) \quad g(e|s, y) = \begin{cases} \frac{g(e)f(s|e)}{\int_0^{e^*} g(e)f(s|e)de}, & e \leq e^* \\ 0, & \text{else} \end{cases} \quad \text{and} \quad g(e|s, n) = \begin{cases} \frac{g(e)f(s|e)}{\int_{e^*}^{\infty} g(e)f(s|e)de}, & e > e^* \\ 0, & \text{else} \end{cases} .$$

The plaintiff wants to appeal a decision by J_T to reject the plaintiff's argument if he believes that it is sufficiently likely that J_A will reverse J_T 's decision. The plaintiff's assessment of this probability depends on both his belief about the true value of the evidence e and the judges' strategies, e^* and \hat{e} . Suppose that $d = n$. If $\hat{e} > e^*$, then the plaintiff's net expected payoff from pursuing an appeal is given by:

$$(5) \quad \frac{\int_{e^*}^{\hat{e}} g(e) f(s|e)de}{\int_{e^*}^{\infty} g(e) f(s|e)de} \pi_p - c ,$$

where the quotient in (5) is the plaintiff's belief that J_A will reverse J_T . The plaintiff will pursue an appeal if the expression given in (5) is non-negative. J_A never reverses a decision by J_A in favor of the defendant if $\hat{e} < e^*$, and therefore in this case the plaintiff does not appeal a decision by J_T in favor of the defendant.

Similarly, the defendant wants to appeal a decision by J_T in favor of the plaintiff if he believes that it is sufficiently likely that J_A will reverse the decision. Suppose that $d = y$. If $\hat{e} < e^*$, then the defendant's net expected payoff from pursuing an appeal is given by:

$$(6) \quad \frac{\int_{\hat{e}}^{e^*} g(e) f(s|e)de}{\int_0^{e^*} g(e) f(s|e)de} \pi_D - c .$$

The defendant will pursue an appeal if the expression given in (6) is non-negative. J_A never reverses a decision by J_T in favor of the plaintiff if $\hat{e} > e^*$, and therefore in this case the defendant does not appeal a decision by J_T in favor of the plaintiff.

Lemma 3. Suppose that $f(s|e)$ satisfies the monotone likelihood ratio property. Then if $\hat{e} > e^*$,

$$\frac{\int_{e^*}^{\hat{e}} g(e) f(s|e) de}{\int_{e^*}^{\infty} g(e) f(s|e) de} \text{ is decreasing in } s. \text{ If } \hat{e} < e^*, \frac{\int_{\hat{e}}^{e^*} g(e) f(s|e) de}{\int_0^{e^*} g(e) f(s|e) de} \text{ is increasing in } s.$$

Lemma 3 implies that the plaintiff's problem is equivalent to choosing a "cutoff" value of s , say $s_P(d)$, such that for all $s \leq s_P(d)$ the plaintiff appeals J_T 's decision $d = n$, choosing action $a = p$, and for all $s > s_P(d)$ he does not appeal, that is, chooses $a = q$. Similarly, the defendant's problem is equivalent to choosing a "cutoff" $s_D(d)$ such that for all $s \geq s_D(d)$ the defendant appeals the ruling $d = y$, choosing $a = p$, and for all $s < s_D(d)$, he chooses $a = q$. In the following lemma we describe the litigants' best replies.

Lemma 4. If $\hat{e} > e^*$, the litigants' best replies are $s_D(d) = \infty$ for all $d \in \{y, n\}$, $s_P(y) = 0$, and

$$s_P(n) \text{ satisfying } \begin{cases} \int_{e^*}^{\hat{e}} g(e) f(s_P(n)|e) de > 0, & s_P(n) = \infty \\ \int_{e^*}^{\infty} g(e) f(s_P(n)|e) de - \pi_P - c = 0, & s_P(n) \in (0, \infty) \\ \int_{e^*}^{\infty} g(e) f(s_P(n)|e) de < 0, & s_P(n) = 0 \end{cases}$$

If $\hat{e} < e^*$, the litigants' best replies are $s_P(d) = 0$ for all $d \in \{y, n\}$, $s_D(n) = \infty$, and

$$s_D(y) \text{ satisfying } \begin{cases} \int_{\hat{e}}^{e^*} g(e) f(s_D(y)|e) de > 0, & s_D(y) = 0 \\ \int_{e^*}^{\hat{e}} g(e) f(s_D(y)|e) de - \pi_D - c = 0, & s_D(y) \in (0, \infty) \\ \int_0^{e^*} g(e) f(s_D(y)|e) de < 0, & s_D(y) = \infty \end{cases}$$

Lemmas 1, 2, and 4 characterize the model's equilibrium.

Proposition 2. The Bayesian Nash equilibrium of the game exists and is given by the vector $(e^*, \hat{e}, s_P(d), s_D(d); g(e|s, d))$.

In the equilibrium described in Proposition 2 the litigants will not in general appeal every decision made by J_T that J_A would reverse, both because appeals are costly and because the litigants are incompletely informed about how J_A would rule. Suppose, for example, that J_T rules in favor of the defendant. Although the plaintiff would like J_A to reverse the decision, the plaintiff will not appeal if she observes a high value of s , since in this case she is pessimistic about her chances of winning. Because s is only a noisy signal of e the plaintiff may forego an appeal that would ultimately succeed. Clearly the litigants are less inclined to appeal as the cost of appeals c rises. Thus, we conclude that the appeals court judge cannot act as a perfect monitor of the lower court, since she must rely on the litigants to select cases for appeal.

A consequence of this decentralized review process is that some cases will be decided in ways that are contrary to the wishes of the judges at the top of the hierarchy. This is unavoidable as long as the trial court and appeals court judges decide cases using different evidentiary standards. In our model, J_T 's preferred e^* is insensitive to the possibility of being reversed, guaranteeing that e^* and \hat{e} will generally diverge. This divergence would persist even if J_T suffered a loss from a successful appeal, though the difference between e^* and \hat{e} would shrink. In general the appeals process cannot guarantee $e^* = \hat{e}$, since in such an equilibrium litigants would have no incentive to appeal—they would never expect J_A to reverse J_T 's decision—thus destroying J_T 's incentive to choose any evidence “cutoff” other than his preferred cutoff e^* .

While we have so far focused on the judicial hierarchy, there is another sense in which the legal system is hierarchical, as discussed in the previous section. The system of legal rules is also hierarchical, encompassing higher level rules and lower level rules. The appeals court judge can use a higher level rule to affect the decisions made by the judicial hierarchy.

In order to analyze the choice of a higher level rule, we suppose that, in addition to ruling on the merits of a particular dispute, J_A can also choose the legal regime that governs all disputes of a particular kind. Formally, J_A is able to choose a higher level rule $R \in \{strict, loose\}$. By selecting $R = strict$, J_A announces that the charge against the defendant, once substantiated, will never be excused. This is equivalent to restricting both J_A and J_T to the action set $\{y (= \text{accept the plaintiff's argument})\}$ rather than $\{y, n (= \text{reject the plaintiff's argument})\}$. By choosing $R = loose$, on the other hand, J_A does not restrict either judge's action, and in any particular case the equilibrium that arises will be as described in Proposition 2.

For example, in contractual disputes, the higher court judge can promulgate a rule stating that contracts will never be excused (even if efficiency dictates that a contract be excused once it has been breached), in order to preserve the incentives of people in the entire jurisdiction to honor future contracts. Higher court judges have similar power with respect to antitrust issues. The higher court judge can choose the higher level rule $R \in \{per\ se\ illegal, rule\ of\ reason\}$. If $R = per\ se$, the freedom of the lower court is more restricted than if $R = rule\ of\ reason$.

Why might J_A want to choose $R = strict$? As demonstrated above, J_A 's preferred decision generally differs from J_T 's preferred decision. Moreover, because J_A must rely on the litigants to bring appeals, she cannot be confident that she will be able to reverse all decisions that she believes are wrong; few litigants may appeal the decisions made by J_T , either because the cost of

appeals is high or because of uncertainty about the outcome. Consequently, J_A might be limited in her ability to establish precedents that reflect the interests of her broad jurisdiction.

By choosing $R = \textit{strict}$, J_A may improve—from her point of view—the quality of the decisions made by the judicial hierarchy. In our model, this will tend to be the case when J_A rarely wants to excuse the defendant's offense, and therefore suffers little from giving up the ability to excuse it; when J_T more frequently wishes to throw out the charge against the defendant; and, finally, when the appeals process is costly and uncertain and therefore seldom used. Intuitively, if J_A wishes to choose \hat{e} close to ∞ but seldom has the chance to influence a case because litigants rarely appeal, then J_A may wish to force J_T to choose $e^* = \infty$. J_A does this by establishing the appropriate higher level rule when the opportunity arises. Thus, J_A is able to influence the legal rules chosen in all N jurisdictions without hearing any further cases.

VI. BROADER IMPLICATIONS

The argument of this paper is that higher level courts will provide more protection to rules, and less to particular individuals, than will lower level courts. The pre-New Deal Supreme Court represented the epitome of this argument. This Court frequently overturned lower court decisions because these did not provide sufficient protection for long term rights, primarily rights to free contract and to property (for one discussion, see Horwitz, 1977, Chapter 1). However, the political reaction to the Great Depression changed the long term focus of the Supreme Court.

We would in general expect the political process to place more weight on individuals than do the courts. There are several reasons for this. First, in the political process identifiable

individuals are involved. A tariff, for example, benefits current firms and identifiable workers in a particular industry. FDA and other safety regulation is aimed at protecting identifiable individuals, even though the net effect may be harmful. Sometimes these individuals are in a position to seek their own benefit as shown in the rent seeking literature. Moreover, appeals to identifiable individuals apparently have more force with voters than do appeals to general principles. Voters, like jurors, are not trained in long term abstract analysis, and they are able to see effects on identifiable individuals directly. Politicians also have relatively short time horizons. They must be reelected and must appeal to voters who have short time horizons and are concerned with individuals.⁶

The political process, in attempting to alleviate the shock of the Great Depression, behaved in its normal manner and adopted various short term solutions as part of the New Deal. Since the problem was catastrophic, the solutions adopted were extraordinary.⁷ Among these were the Supreme Court rulings reducing contractual freedom. While these rulings were aimed at protecting statutes that would have previously been overturned, the effect was to institute rules that overturned explicit contracts. As Epstein (1984, at 742) argues: “After all, if the contract clause prevents impairment of contracts with the state, then it also prevents impairment of contracts between private parties” (and conversely). It was a natural step from allowing statutes overturning contracts to directly overturning contracts by the courts. Lower level courts place more weight on individuals than does the Supreme Court. Thus, once freed from the constraints

⁶ It is for these reasons that Buchanan and others have argued for Constitutional constraints on the political system. Constitutional constraints will force the system to act as if it has a long time horizon and to pay more attention to rules (see, for example, Buchanan, 1975).

⁷The scope of legal change needed to implement the New Deal was enormous (see, e.g., Ely, 1992, Chapter 7).

of *Lochner*,⁸ the lower courts were able to indulge their preferences for solutions that implied benefits for particular individuals at the expense of rules. This has caused the weakening of contract noted by many observers (e.g., Barnett, 2003, Horwitz, 1977, Epstein, 1984, and Rubin, 1993).

Of course, this outcome was not instantaneous. Lower courts could not immediately adopt principles favoring individuals at the expense of rules. For one thing, the subsidiary rules were not known or even knowable because under a regime of free contract, many events that would later be interpreted as violations were not so considered. If a contract protected a manufacturer against all claims arising from all product related injuries, for example, then there would be no need to determine if liability was based on negligence or was strict, and no need to distinguish between design and manufacturing defects. The law would not even recognize the existence of these differences; they are legally hidden distinctions.⁹ Moreover, to change the law requires bringing cases that establish new precedents. This is also a time consuming process. A case must first be heard at the level of the trial court. Before it becomes a precedent, however, it must be appealed through the higher level courts and affirmed. Each step takes time. Agents conducting the appeals process to change the law in the way discussed here include plaintiffs, and also attorneys (Rubin and Bailey, 1994).

However, as argued in this paper, lower level courts provide less protection to rules, and more to individuals, than do higher level courts. At any given time, lower courts would prefer more individual oriented rulings than they are allowed by the higher courts. Thus, if a higher court changes rulings to allow more attention to individuals (as did the Supreme Court after the

⁸*Lochner v. New York*, 198 U.S. 45 (1905).

Great Depression), then lower courts will gladly adopt these rulings. On the other hand, if the higher court has moved in the other direction, announcing more emphasis on rules (perhaps the situation that now obtains in the courts, with many Reagan-Bush judges in the Supreme Court), then we would expect the lower level courts to resist moving to this new level. Thus, for example, it should take longer to reverse the movement away from freedom of contract than it took to implement the movement in the first place.

VII. CONCLUSION

The basic theory here has empirical implications regarding differences between trial court and appellate decisions at a legal equilibrium. The implications are that appellate courts will be more rule oriented and less individual oriented than trial courts, and that in a multilevel system this tendency will continue as we move up the hierarchy. With respect to tort, this means that higher level courts will be more pro-defendant than trial courts. With respect to contract, higher level courts will be more likely to enforce explicit terms than will lower level courts. In property rights cases, higher level courts should provide more protection to property rights than do lower level courts. However, testing these implications will require correcting for selection effects with respect to those cases that are actually litigated and appealed (see Priest and Klein, 1984, and Waldfogel, 1995).

⁹This is related to, but distinct from, the notion, developed by Hadfield (1992), that legal evolution will be biased because courts will only see certain disputes under any given legal rule.

ADDENDIX

Proof of Lemma 1. Fix the evidence e and suppose the plaintiff appeals with probability α_n if J_T chooses $d = n$. Suppose that the appeals court judge J_A chooses $d = y$ if a litigant chooses to appeal. Then J_T 's expected utility from $d = n$ is:

$$(1 - \alpha_n)(u_L(e, n) + u_p(e, n)) + \alpha_n(u_L(e, y) + u_p(e, y)).$$

His expected utility from $d = y$ is:

$$u_L(e, y) + u_p(e, y).$$

Therefore, J_T chooses $d = n$ if and only if

$$u_L(e, y) + u_p(e, y) - u_L(e, n) - u_p(e, n) \geq 0.$$

By monotonicity of J_T 's preferences, this is satisfied if $e \geq e^*$, where e^* is given in the Lemma.

Similar arguments establish the result when J_A chooses $d = n$. QED

Proof of Lemma 2. Follows from monotonicity of J_A 's preferences. QED

Proof of Proposition 1. Follows from Lemmas 1 and 2. QED

Proof of Lemma 3. $\frac{\int_{e^*}^{\hat{e}} g(e) f(s|e) de}{\int_{e^*}^{\infty} g(e) f(s|e) de}$ is decreasing in s if

$$\frac{\int_{e^*}^{\hat{e}} g(z) f_s(s|z) dz}{\int_{e^*}^{\hat{e}} g(z) f(s|z) dz} < \frac{\int_{e^*}^{\infty} g(z) f_s(s|z) dz}{\int_{e^*}^{\infty} g(z) f(s|z) dz}.$$

To establish this, it is sufficient to show that $\frac{\int_{e^*}^x g(z) f_s(s|z) dz}{\int_{e^*}^x g(z) f(s|z) dz}$ is increasing in x , i.e. that

$$\frac{f_s(s|x)}{f(s|x)} > \frac{\int_{e^*}^x g(z) f_s(s|z) dz}{\int_{e^*}^x g(z) f(s|z) dz} = \frac{\int_{e^*}^x g(z) f(s|z) \frac{f_s(s|z)}{f(s|z)} dz}{\int_{e^*}^x g(z) f(s|z) dz}.$$

Since $\frac{f(s|e')}{f(s|e'')}$ is increasing in s if $e' > e''$, by the MLRP, we have $\frac{f_s(s|e')}{f(s|e')} > \frac{f_s(s|e'')}{f(s|e'')}$, so that

$$\frac{\int_{e^*}^x g(z) f(s|z) \frac{f_s(s|z)}{f(s|z)} dz}{\int_{e^*}^x g(z) f(s|z) dz} < \frac{f_s(s|x) \int_{e^*}^x g(z) f(s|z) dz}{f(s|x) \int_{e^*}^x g(z) f(s|z) dz} = \frac{f_s(s|x)}{f(s|x)},$$

which establishes the first part of the Lemma. Analogous arguments establish the second part of the Lemma. QED

Proof of Lemma 4. Follows from Lemma 3, which implies that the litigants' payoffs are monotone in s . QED

REFERENCES

- Barnett, Randy E., 2003. *Restoring the Lost Constitution: The Presumption of Liberty*, Princeton University Press, Princeton.
- Buchanan, James M., 1975. *The Limits of Liberty: Between Anarchy and Leviathan*, University of Chicago Press

- Cameron, Charles M. and Kornhauser, Lewis A., 2004. "Decision Rules in a Judicial Hierarchy," University of Michigan Law and Economics Working Paper.
- Cohen, Mark A., 1992. "The Motives of Judges: Empirical Evidence from Antitrust Sentencing," 12 *International Review of Law and Economics* 13.
- _____, 1991. "Explaining Judicial Behavior or What's 'Unconstitutional' About the Sentencing Commission," 7 *Journal of Law, Economics & Organization* 183-199.
- Daughety, Andrew F. and Reinganum, Jennifer F., 2000. "Appealing Judgments," 31 *Rand Journal of Economics* 502-525.
- Ely, James W. Jr., 1992. *The Guardian of Every Other Right: A Constitutional History of Property Rights*, Oxford University Press, New York.
- Epstein, Richard A., 1984. "Toward a Revitalization of the Contract Clause," 51 *University of Chicago Law Review* 703.
- _____, 1993. "Holdouts, Externalities and the Single Owner: One More Salute to Ronald Coase," 36 *Journal of Law and Economics* 553.
- Hadfield, Gillian K., 1992. "Biases in the Evolution of Legal Rules," 80 *Georgetown Law Journal* 583.
- Higgins, Richard S. and Rubin, Paul H., 1980. "Judicial Discretion," 9 *Journal of Legal Studies* 129-38.
- Horwitz, Morton J., 1977. *Transformation of American Law*, Harvard University Press, Cambridge, Massachusetts.
- Kornhauser, Lewis A., 1992a. "Modeling Collegial Courts I: Path Dependence," 12 *International Review of Law and Economics* 169-185.

- _____, 1992b. "Modeling Collegial Courts II: Legal Doctrine," 8 *Journal of Law, Economics and Organization* 441-470.
- Kobayashi, Bruce H. and Lott, John R. Jr., 1994. "Judicial Reputation and the Efficiency of the Common Law," mimeo, George Mason University.
- Levy, Gilat, 2004. "Careerist Judges," *Rand Journal of Economics*, forthcoming
- Miceli, Thomas J. and Cosgel, Metin M., 1994. "Reputation and judicial decision-making," 23 *Journal of Economic Behavior and Organization* 31-51.
- O'Hara, Erin, 1993. "Social Constraint or Implicit Collusion?: Toward a Game Theoretic Analysis of Stare Decisis," 24 *Seton Hall Law Review* 736-778.
- Posner, Richard A., 1992. *Economic Analysis of Law*, Fourth Edition, Little, Brown, Boston.
- _____, 1993. "What Do Judges Maximize? (The Same Thing Everybody Else Does)," 3 *Supreme Court Economic Review* 1-41.
- Priest, George and Klein, Benjamin, 1984. "The Selection of Disputes for Litigation," 13 *Journal of Legal Studies* 1.
- Rasmusen, Eric, 1994. "Judicial Legitimacy as a Repeated Game," 10 *Journal of Law, Economics & Organization* 63-83.
- Rubin, Paul H., 1993. *Tort Reform by Contract*, American Enterprise Institute, Washington.
- _____ and Bailey, Martin J., 1994. "The Role of Lawyers in Changing the Law," 23 *Journal of Legal Studies* 807.
- Shavell, Steven, 1995. "The Appeals Process as a Means of Error Correction," 24 *Journal of Legal Studies* 379-426.

_____, 2004. "The Appeals Process and Adjudicator Incentives," NBER Working paper 10754.

Spitzer, Matt and Talley, Eric, 2000. "Judicial Auditing," 29 *Journal of Legal Studies* 649-683.

Waldfogel, Joel, 1995. "The Selection Hypothesis and the Relationship between Trial and Plaintiff Victory," 103 *Journal of Political Economy* 229-260.

Whitman, Douglas G., 2000. "Evolution of the Common Law and the Emergence of Compromise," 29 *Journal of Legal Studies* 753-781.