Stampede to Judgment: Persuasive Influence and Herding Behavior by Courts

Andrew F. Daughety and Jennifer F. Reinganum, *Vanderbilt University*

We model appeals courts as Bayesian decision makers with private information about a supreme court's interpretation of the law; each court also observes the previous decisions of other appeals courts in similar cases. Such "persuasive influence" can cause "herding" behavior by later appeals courts as decisions progressively rely more on previous decisions and less on a court's private information. We provide an example drawn from a recent United States Supreme Court decision finding unconstitutional a basic provision of a law previously found constitutional by six circuit courts. Herding on the wrong decision may remain uncorrected, since review of harmonious decisions is rare.

1. Introduction

A hierarchy of courts is a nonmarket (and, ideally, nonpolitical) means for generating and aggregating decisions, in particular, decisions about the law. Many scholars have discussed the role of vertical relationships in influencing the formation of decisions: from a court's viewpoint, binding precedent involves some reliance on past decisions of that particular court and those superior to it in the hierarchy. The effect of following binding precedent (*stare decisis*) has been the subject of considerable discussion; we review some of the discussion below. Less well understood is the effect of horizontal relationships on legal decisions: what happens when

Thanks to John Goldberg, Richard Posner, Kathryn Spier, Nicholas Zeppos, and an anonymous referee for helpful conversations and suggestions.

Send correspondence to: Andrew Daughety, Department of Economics and Business Administration, Vanderbilt University, 415 Calhoun Hall, Nashville, TN 37235; Fax: (615) 343-8495; E-mail: andrew.f.daughety@vanderbilt.edu.

decisions in a legal system reflect the influence of nonbinding precedent? This occurs when a court in one jurisdiction relies upon a decision made by a court in another jurisdiction.

Such persuasive influences among courts are the subject of this article. "Herding," also called "informational cascades," occurs among agents when their decisions are decreasingly determined by their own information and increasingly determined by the actions of others; this has been the subject of a growing literature, which we briefly summarize below. In that literature, much of which focuses on such phenomena in markets, the persistence of such correlated actions may be fragile to releases of new public information: if agents "herd" on the "wrong" outcome, new public information may correct the problem. In our case, the passive nature of courts (that is, a case must be brought) means that erroneous decisions made by one level in a system of courts potentially remain uncorrected, even though all courts are trying to make principled judgments. We also show how such herding can amplify the sensitivity of individual court decisions to (possibly incorrect) judgments made by other courts, leading to vastly different sequences of holdings for the same set of cases, depending only on the order in which they arise. Moreover, we find that too much harmony among court decisions may be cause for as much concern (and possibly more) than too much conflict.

In Section 2 we provide a brief discussion of results on herding and a brief overview of models of court decisions. Section 2 also includes an example that illustrates the model we provide and examine in Sections 3 and 4. Our example is the June 25, 1998, decision by the U.S. Supreme Court in *Eastern Enterprises v. Apfel*, wherein a 5–4 decision by the Court found unconstitutional a basic provision of the Coal Industry Retiree Health Benefit Act of 1992. This act had previously survived constitutional challenge in cases brought in six U.S. Circuit Courts of Appeals (in sequence, over the period April 1995 through April 1997, the Second, Seventh, Sixth, Third, Fourth, and First Circuits).

Section 3 provides the relevant elements of a model of adversarial evidence generation, trial and appeal that we have detailed elsewhere (Daughety and Reinganum, 1998a, 1998b). Using this model, Section 4 considers a collection of privately informed Bayesian courts facing both binding and nonbinding (but influential) precedents. The hierarchy we analyze consists of a system of courts of appeals, each superior to a trial court, all inferior to the same supreme court. Section 5 contains a summary of the results and implications.

2. Review of Related Literature and a (Potential) Example of Herding by Courts

Background on the Herding Literature

This paper focuses on decision making by a collection of "horizontal" courts; for example, the U.S. Circuit Courts of Appeals or the state supreme courts.[1] Within such a collection of courts, there is no binding precedent; nevertheless, decisions made by one court may be influential in another court's decision making. We adapt to the judicial context a model developed by Banerjee (1992) and Bikchandani, Hirshleifer, and Welch (1992), in which they demonstrate how an accumulation of public information garnered from the previous decisions of others can overwhelm an individual decision maker's private information.[2] This results in what Bikchandani, Hirshleifer, and Welch call an "informational cascade" and Banerjee calls "herding:" decision makers early in the sequence rely (though, to a decreasing extent) on their private information, while later decision makers simply go along with the herd. They suggest that such an information externality may play a role in the occurrence of (and rapid changes in) mass behavior, including fashion, smoking, drinking, and various market behaviors.[3] As these authors point out, when an agent's de-

1. Lynn (1993) observes that studies of state supreme court opinions find that one-third of case citations refer to out-of-state cases; the most often cited were New York, Massachusetts, and California. In what follows, we consider cases involving issues over which a single supreme court has ultimate jurisdiction. When state supreme courts consider cases involving federal law, the U.S. Supreme Court has final jurisdiction, while state supreme courts have final jurisdiction in issues of state law. Our model addresses cases of the first type but not the second.

2. After completing this paper, we became aware of a working paper by Talley (1996; rev. 1998), in which he suggests herding as an endogenous source of *stare decisis* and convergence in the evolution of the common law.

3. Related models have been used extensively in the finance area to examine the investment behavior of money managers (Scharfstein and Stein, 1990), initial public offerings (Welch, 1992), and buying frenzies and crashes (Bulow and Klemperer, 1994). Choi (1997) incorporates both informational and network externalities in order to examine technology adoption. See also Bikchandani, Hirshleifer, and Welch (1998) for a recent survey.

cision is not a perfect signal of his private information, it is eminently possible for the agents to end up coordinating on the "wrong" outcome.[4]

Bikchandani, Hirshleifer, and Welch emphasize that informational cascades may be fragile because a public release of new information may cause wholesale changes in decisions, as previously decided agents change their minds and the informational cascade begins again, possibly settling on a very different outcome. For instance, if consumers have coordinated on a particular health fad, a government release of information contradicting its putative benefits may disrupt the outcome. In our context, a "public" release of information would correspond to a Supreme Court opinion; since the Supreme Court must wait for cases to be appealed (and even then, it must find a case sufficiently compelling relative to the others awaiting review), it is less able to intervene directly than is the government in the health example above. Consequently, outcomes in which a collection of courts herds on the wrong decision may be persistent.

A Potential Example of Herding: Decisions About the Constitutionality of the Coal Act

In Section 4 below we provide a model of how persuasive influence (that is, the sequential influence of horizontal courts on each other's decisions) can result in herding. In this subsection we briefly discuss an example that has the earmarks of such behavior; we return to this example in Section 5, after presenting the model, in order to provide further discussion.[5]

On June 25, 1998, the U.S. Supreme Court voted 5 to 4 to hold a basic provision of the Coal Industry Retiree Health Benefit Act of 1992 (the Coal Act) unconstitutional. The Coal Act was designed to fund defined lifetime benefits for miners and their dependents by allocating liability to all signatories to any wage agreement from 1950 on (only those agreements from 1978 on contained an explicit promise of defined lifetime

4. Lee (1993) asks what is necessary to ensure that decision makers converge to the "right" outcome; he finds that a discrete action set always permits herding on the wrong outcome. Chamley and Gale (1994) and Gul and Lundholm (1995) examine the potential for informational cascades when agents are allowed to choose when to move, rather than being ordered exogenously. Chamley and Gale maintain discrete action sets and obtain herding; Gul and Lundholm employ a continuum action set, which allows perfect revelation of private information.

5. We thank Nicholas Zeppos for bringing this case to our attention.

benefits).[6] Justices O'Connor, Rehnquist, Scalia, and Thomas found the provision unconstitutional because it constituted an uncompensated taking, while Justice Kennedy found that it violated principles of substantive due process. The particular case involved Eastern Enterprises, which had mined coal from 1946 through 1965, when it left the industry (though it continued to receive revenues from a wholly owned subsidiary from 1966 through 1987). This case was the sixth of a sequence of fairly similar cases, concerning the retroactive reach of the Coal Act, which had been heard by U.S. Courts of Appeals in as many circuits over a two-year period stretching from April 1995 through April 1997;[7] it was one of four of these that were appealed to the Supreme Court, the only one (of four) to be granted certiorari.[8]

Each succeeding appeals court opinion referenced all the previous decisions. Moreover, each opinion became progressively shorter (except in *Eastern Enterprises v. Chater*, which raised an extra issue) and applied progressively similar criteria to reach the same conclusion: while the Coal Act involved retroactive application of liability to the firm(s) in question, due process was not violated and the action did not constitute an uncompensated taking (most referred to it as a tax). A few choice quotes help illustrate the discussion:

6. For a brief review of the events leading up to the passage of the Coal Act, see the plurality opinion, written by Justice O'Connor, in *Eastern Enterprises v. Apfel.*

7. A seventh (and earlier) case, *Barrick Gold Exploration, Inc., et al. v. Hudson*, Sixth Circuit, decided February 24, 1995, considers a somewhat different issue (the lack of credit under the Coal Act for withdrawal liabilities previously paid) from those addressed by the cases we focus on.

8. The following listing provides a title of the case, the circuit involved, the date of the appeals court's decision followed parenthetically by the disposition of any subsequent appeal. The six appeals courts cases, in sequence, are: (1) *in re Chateaugay*, Second Circuit, decided April 17, 1995 (cert. denied October 10, 1995); (2) *Davon, Inc., et al. v. Shalala*, Seventh Circuit, decided January 25, 1996 (cert. denied in *Templeton Coal, et al.*, October 7, 1996); (3) *Blue Diamond Coal Co. v. Shalala*, Sixth Circuit, decided March 21, 1996 (rehearing en banc by Sixth Circuit denied June 28, 1996; cert. denied January 6, 1997); (4) *Lindsey Coal Mining Co. v. Chater*, Third Circuit, decided July 26, 1996; (5) *Holland v. Keenan Trucking Co.*, Fourth Circuit, decided December 12, 1996; (6) *Eastern Enterprises v. Chater*, First Circuit, decided April 7, 1997 (cert. granted October 20, 1997; *Eastern Enterprises v. Apfel* decided June 28, 1998). It is also worth observing that the Court's membership (Justices Breyer, Ginsgurg, Kennedy, O'Connor, Rehnquist, Scalia, Souter, Stevens, and Thomas) remained constant over the entire period of these cases.

"We conclude that the Coal Act's retroactive financing provision challenged by plaintiffs easily passes the due process rationality test. Our conclusions are consistent with every other court that has addressed the issue" (Commenting on substantive due process issue, followed by references to *Chateaugay, Barrick Gold,* district court decisions in *Blue Diamond, Lindsey Coal and Unity Real Estate Co. v. Hudson,* to which we return later). *Davon, Inc., et al.,* at 40.

"We are thus in agreement with all but one federal court to have decided the issue." (Commenting on takings issue, followed by references to *Chateaugay, Barrick Gold,* district court decisions in *Blue Diamond* and *Lindsey Coal*). "But see *Unity Real Estate Co. v. Hudson*... (finding unconstitutional taking)." *Davon, Inc., et al.,* footnote 12 at 50.

"Three circuit courts, including the Sixth Circuit, and several district courts have considered challenges to the constitutionality of various provisions of the Coal Act... In *Davon, Inc. v. Shalala,*..., a case nearly identical to the instant one." *Blue Diamond v. Shalala,* at 8. *Davon* is used at a number of later points to buttress the argument.

"The reporters are full of court of appeals' decisions concluding, like the district court, that the Coal Act is rational economic legislation that comports with the substantive requirements of the Due Process Clause" (followed by references to *Davon, Blue Diamond, Barrick Gold* and *Chateaugay*). "We agree with the views expressed so well in these cases." *Lindsey Coal v. Chater,* at 694.[9]

With respect to the takings issue (discussed in a total of two paragraphs at 695):

"We agree. As with the Due Process challenge, every court of appeals to consider a 'takings' challenge to the Coal Act has rejected it" (followed by references to *Davon, Blue Diamond, Barrick Gold* and *Chateaugay*). "We endorse the reasoning of these cases."

"In upholding the constitutionality of the Coal Act, we join the unanimous opinion of the circuits which have considered this issue" (followed by references to *Lindsey, Blue Diamond, Davon* and *Chateaugay* appeals courts' decisions). *Holland v. Keenan,* at 16.

"The constitutional arguments are retreads which have taken their lumps from circuit courts of appeals in five other circuits" (followed by references to *Holland, Lindsey, Blue Diamond, Davon* and *Chateaugay* cases). "Although these decisions

9. It is worth noting that *Davon* had cited the very district court decision on the Lindsey Coal Mining Company (see quote above) that the Third Circuit Appeals Court was reviewing on appeal. A similar citing result occurs in *Blue Diamond,* which cites *Davon,* which had cited the district court Blue Diamond decision.

are not binding on us, we find them convincing." *Eastern Enterprises v. Chater*, at 2.

The plurality decision in the Supreme Court reviewed a number of its previous decisions (many or all of which were cited in various of the appeals courts decisions), stating "Our decisions, however, have left open the possibility that legislation might be unconstitutional if it imposes severe retroactive liability on a limited class of parties that could not have anticipated the liability, and the extent of that liability is substantially disproportionate to the parties' experience. We believe that the Coal Act's allocation scheme, as applied to Eastern, presents such a case." (O'Connor at 54, 55).

It is worth noting that in *Unity Real Estate v. Hudson* (December 7, 1994), U.S. Magistrate Judge Keith A. Pesto (Western District of Pennsylvania, which is in the Third Circuit) issued a preliminary injunction restraining enforcement of the Coal Act; his argument emphasized previous Supreme Court holdings suggesting that the retroactive effect of legislation should be limited, a central point of the eventual plurality decision written by O'Connor. Pesto also quotes a similar reservation held by the district court in the Blue Diamond case (however, that district court found against Blue Diamond). Pesto's decision was later (June 7, 1995) supported by District Court Judge D. Brooks Smith. After the Third Circuit's July 26, 1996, decision in *Lindsey Coal*, Smith reversed his earlier decision on March 14, 1997. *Unity* was noted in the appeals courts' decisions in *Davon* (Seventh Circuit) and *Blue Diamond* (Sixth Circuit) but then disappeared from sight until the First Circuit noted its holding and its reversal, stating that it was "not persuasively reasoned and its precedential force has been undermined severely by the Third Circuit's summary rejection of a Takings Clause challenge" (citing *Lindsey*). *Eastern Enterprises v. Chater*, footnote 8 at 35.

From our perspective, the story is as follows. The Coal Commission, Congress, and the appeals courts found that the coal companies had promised defined lifetime benefits, either explicitly or—in the case of those who had signed only pre-1978 agreements—implicitly, and the purpose of the Coal Act was to require the companies to make good on this promise. In contrast, the majority of Supreme Court justices found that there was no such implicit promise, so the Coal Act involved severe retroactive liability that could not have been anticipated. Thus, while the decisions in the cases *In re Chateaugay and Holland v. Keenan* (involving signatories to 1978 or later agreements) *might* have been upheld by the Supreme Court, the *Davon, Blue Diamond,* and *Lindsey* cases were factually similar to *Eastern Enterprises* in dealing with companies that had exited the industry before any explicit promises of defined lifetime benefits were made. This suggests that the appeals court decisions in *Davon, Blue Diamond, Lindsey,* and *Eastern Enterprises might* reflect overdependence on the public signals being created by the sequence of appeals courts' decisions.

Background on Models of Court Decision-Making

There is a vast literature on decision making by courts, which we only briefly summarize here. Decision making by trial courts, whose objective has been assumed to be accuracy, has been modeled in (at least) four ways.[10] First, courts have been modeled as (exogenously specified) processors of evidence and/or expenditure. That is, the probability of the plaintiff's prevailing is a given function of the evidence or, more usually, the expenditure of the parties (Brauetigam, Owen and Panzar, 1984; Danzon, 1983; Hause, 1989; Katz, 1987, 1988; Landes, 1993; Plott, 1987). Second, the court has been modeled as a fully Bayesian processor of evidence and/or expenditure. That is, the court infers liability or damages from either the evidence (or lack thereof) that the parties present (Daughety and Reinganum, 1995; Sanchirico, 1997a; Shin, 1994, 1998; Sobel, 1985) or their expenditures (Rubinfeld and Sappington, 1987; Sanchirico, 1997b, 1997c). Third, courts have been modeled as "naive" to varying extents. For instance, Milgrom and Roberts (1986) examine both fully Bayesian courts and courts that are naive but computationally able: the court takes all the evidence provided at face value, but, if provided with all of the relevant evidence, it can compute the socially optimal outcome. This strand of the literature (see also Lipman and Seppi, 1995; Seidmann and Winter,

10. There is also an enormous (primarily statistical) literature in political science, law, and psychology on modeling decision making by jurors and/or judges that does not necessarily presume that the objective is accuracy (for some recent contributions, see George and Epstein, 1992; Hastie, 1993; Segal and Spaeth, 1993). A recent paper that shows that appeals courts have incentives to follow legal doctrine is Cross and Tiller (1998). Other papers examine the objectives of judges (Ashenfelter, Eisenberg, and Schwab, 1995; Cohen, 1992; Posner, 1993) and issues of judicial independence (Cooter and Ginsburg, 1996; Ferejohn and Weingast, 1992; Spiller and Tiller, 1996).

1997) focuses on characterizing sufficient conditions for all relevant evidence to be provided. In a related vein, Froeb and Kobayashi (1993, 1996; see also Farmer and Pecorino, 1998) model the court as "naive (and potentially biased) but Bayesian"; essentially, the court uses Bayesian updating with respect to the evidence presented but not with respect to the strategic incentives of the parties. Finally, Daughety and Reinganum (1998a, 1998b) assume that trial court decision making is constrained by rules of evidence and procedure to satisfy several desirable properties, or axioms, with respect to the treatment and aggregation of evidence. These two papers provide the basis for our analysis and are discussed in greater detail in Section 3.

Models of appeals court (or Supreme Court) decision making largely focus on voting issues within a single court, since these decisions are typically made by panels of judges (Easterbrook, 1982; Gely and Spiller, 1990; Kornhauser, 1992a, 1992b; Spiller, 1992; Toma, 1996). In this context, several issues naturally arise, including the issues of "path dependence" and precedent.[11] For instance, Easterbrook (1982) argues that cycling under majority rule can lead such courts to decide a sequence of cases inconsistently; moreover, the decision in the instant case will depend on the "path" or order in which previous cases were considered within the same court. Kornhauser (1992a) argues that requiring a court to respect its internal precedents (i.e., its own previous decisions either with respect to rules or outcomes) will prevent cycling, but may still involve path dependence. We suppress the voting aspect of appeals court decision making in what follows, in order to focus on the informational externalities between courts which operate on the same level. We find that a different notion of path dependence operates (which we refer to as "interjurisdictional path dependence"), in which the outcome in the instant case may depend crucially on the (essentially random) order in which related cases were considered by courts in other jurisdictions.

There are relatively few combined models of trial and appeal. Shavell (1995) discusses the optimal allocation of resources to a trial and an ap-

11. Miceli and Cosgel (1994), O'Hara (1993), and Rasmusen (1994) use a complete-information repeated games analysis to characterize circumstances under which judges will follow the previous decisions of others (even if it were not required of them) in order to have their decisions followed by subsequent judges. Thus, these papers (along with the Talley paper mentioned earlier) provide equilibrium models of *stare decisis*.

peals court, where the appeals court is charged with correcting errors made by the trial court. Daughety and Reinganum (1998a) model the trial and appeals courts as performing different functions, in a setting of incomplete information about a supreme court's interpretation of law; since we employ parts of this model in the current analysis, we delay a discussion of this paper to the next section. Spitzer and Talley (1998) also use an incomplete information model to examine trial and appeal. Their model is similar in some respects to the model we develop in Section 4, but there are several important differences as well. First, trial and appeal pertain to the same issue in their model, whereas we distinguish between the trial court's focus on evidence and the appeals court's focus on issues of law. This is a major reason why we include all three levels of the judicial hierarchy. Second, there is no difference in ideology between the trial and appeals courts in our model, whereas this is an important element of their model. Finally, there are several differences between the models in terms of "who knows what when." They show that the pattern of "audit" and reversal is sensitive to the motive for review: if review is occurring to correct error, then it will be two-sided, whereas review to correct ideological bias will be one-sided.

3. Relevant Elements of a Model of Trial Courts

As indicated in the introduction, our hierarchy consists of a tree structure involving three levels: trial courts, which are grouped under appeals courts, which in turn are under a supreme court. We employ the standard dichotomy in U.S. judicial systems: trial courts determine issues of fact, based upon the evidence presented (and constrained in their decisions by rules of evidence and procedure), while appeals courts consider issues of law (Posner, 1992, p. 584). Our focus here will be on a civil suit and the issue will be liability, so the outcome of interest concerns a defendant being found liable or not liable.[12] Thus, while a trial court may weigh the evidence about an incident and determine whether a defendant is liable (and if so, the size of a damages award to make to a plaintiff), an appeals court would take the facts as determined but would affirm or reverse

12. See Kornhauser (1992a) for a discussion of why many legal issues can be viewed as primarily binary in nature.

a case based on arguments about the application of the law by the trial court.

We view a generic trial process as being comprised of three elements: evidence production, liability assessment and liability determination; we consider each of these in turn. First, there is the generation and presentation of evidence by the two parties (the plaintiff, P, and the defendant, D).[13] In Daughety and Reinganum (1998b) we presented a model of adversarial evidence generation based on strategic sequential search by each party over a space of possible *credible* evidence. In that paper both parties faced a trial over liability, followed (if D was found liable) by a trial to determine the damages award. Only evidence that was credible (such as estimates of liability or damages by qualified experts) could be presented in court; noncredible evidence (which might be self-serving opinions) was not considered as part of the case. We make this same assumption here. Furthermore, in this paper we only consider trials weighing evidence about the liability of D. Thus, the award to P (should D be found liable) is assumed to be common knowledge.

The second element of a trial process is an assessment function for the trial court (T), whose arguments are the evidence presented by P and D individually and which yields an aggregate assessment by T. Finally, the third element is an exogenously specified evidentiary standard with which to compare the assessment so as to render a judgment. The outcome of the trial (in this paper) involves a binary assessment as to D's liability.

Formally, let x and y be the assessments of D's liability proffered at trial by P and D, respectively, with x and y both in the interval $[0, 1]$. Following Daughety and Reinganum (1998b), we assume that P has developed her case, represented by the assessment x, by sequentially sampling $[0, 1]$ and selecting the best (that is, most advantageous) credible evidence to present at trial, recognizing that D was doing the same for his case, represented by the assessment y. P and D develop these cases aware of (1) T's assessment process; (2) the award that D would pay should he be found liable; and (3) the individual costs to P and D of incrementally constructing their cases.

The second element, T's assessment process, is represented by the function $\ell(x, y)$, which is assumed to be twice continuously differentiable; $\ell(x, y)$ is a number between 0 and 1 (inclusive) which is the trial court's

13. We abstract from any issues of monitoring of, and incentives for, performance by attorneys for P and D. For a discussion of such questions, see Miller (1987).

assessment. Following Daughety and Reinganum (1998a), we model the trial court's assessment process axiomatically: we pose a set of "desirable properties" that $\ell(x, y)$ should possess and use them to derive a specific functional form for $\ell(x, y)$. In Daughety and Reinganum (1998a) (and here, in Section 4) we model an appeals court as Bayesian. The reason behind the use of these different approaches was addressed at some length in Daughety and Reinganum (1998a); we take a few moments to summarize this reasoning now.

First, as indicated earlier, trial courts are different from appeals courts in that, in general, the job of the former is to define what the facts of a matter are (and to use those facts to come to a judgment), whereas the job of the latter is primarily to evaluate issues of law arising from actions taken by the trial court. This division of labor is reflected in the resources devoted to fact-finding versus legal research at each level. Trial courts do not receive significant resources for the purpose of extensive legal research, but devote considerable time and expense to hearing and documenting evidence. On the other hand, appeals courts take the "facts" as determined by the trial court and focus time, attention and resources (such as law clerks and other support personnel) on research and evaluation of the legal issues.

Second, there is a distinction between how the system treats the interpretation of the evidence and how it treats the interpretation of the law. Thus, one outcome of the division of labor is that trial courts have considerably less discretion than appeals courts. For example, rules of evidence and procedure exist to provide restrictions on how assessments of evidence are to be made.

Third, as shown in Daughety and Reinganum (1998b), there is an important technical reason one hesitates to employ a fully Bayesian model to analyze evidence submitted at a trial when the parties have private information. This reason is that each litigant selects the evidence they will present; hence, the trial court and the opposing party are unaware of unpresented evidence; that is, there is not common knowledge of all the facts, even between the litigants. The trial court is also unable to observe the litigants' optimal stopping rules for evidence production; it observes only the results of their employment. This means that a Bayesian court would rely heavily on its prior beliefs rather than the evidence presented. The law goes out of its way to severely limit such reliance. Rules of evidence and of civil procedure limit what a jury or judge may consider, so as not to appeal to prejudices, or even to rationally formed correlations. For example, much of the detail of pretrial bargaining is inadmissible at trial, even though it might be informative. In jury trials, most states discourage or prohibit judges from giving "inference instructions" wherein they draw the jury's attention to specific evidence and comment on what conclusions might be drawn. Errors in applications of rules of evidence or procedure may lead to appeals and reversals.

All of this suggests that one should view trial court decision processes as being subject to a set of desirable properties. The job of the rules of evidence and procedure is to force strategic behavior out of the court's decision process itself, to the greatest degree possible, and confine it to the adversarial behavior of the litigants. This is because the trial court itself is not an adversary: it is a neutral referee.

As in Daughety and Reinganum (1998a), in the sequel we enforce the following "desirable properties" (axioms) for ℓ to possess for all values of (x, y) in $[0,1] \times [0,1]$:

(i) $\ell_x > 0, \ell_y > 0$; (monotonicity)
(ii) $\ell_{xy} \geq 0$; (complementarity)
(iii) $\max\{x, y\} \geq \ell(x, y) \geq \min\{x, y\}$; (interiority)
(iv) a) $\ell(x, y) = \ell(y, x)$, (symmetry)
b) $\ell(\lambda x, \lambda y) = \lambda \ell(x, y)\ \forall \lambda, 0 < \lambda \leq 1$; (homogeneity)
(v) $\forall u, v, w, z,\ \ell(\ell(u, w), \ell(v, z)) = \ell(\ell(u, z), \ell(v, w))$. (divisibility)

The first property, monotonicity, requires that ℓ be increasing in each of its terms. Since x and y are assessments of D's liability, an increase in either party's assessment should lead to an increase in the court's assessment. The second property reflects an assumption of complementarity in the court's assessment of x and y. The reason for this assumption is that this reflects plausible properties of the best response functions in the trial stage when the two litigants are strategically searching for, and providing, the evidence x and y. In Daughety and Reinganum (1998b), we show that, under this assumption, the defendant's best response to increased aggressiveness in evidence gathering by the plaintiff is for the defendant to become increasingly aggressive. This would seem reasonable, since it is the defendant whose wealth is at risk, as being found liable leads to his paying damages. Reversing property (ii) leads to an implausible best response in which the defendant becomes more timid in the face of increasing aggressiveness on the part of the plaintiff.

Property (iii), interiority, requires that the assessment be primarily driven by the evidence. Property (iv) emphasizes unbiasedness via an absolute and a proportional property. The absolute property (iva) is symmetry (or anonymity): inasmuch as the evidence is hypothesized to be credible, it should not matter in the court's assessment who provided the evidence; the assessment based on the two evidence submissions should be the same as would be made if the two submissions were switched. The second property (ivb) is a form of linear homogeneity. This property requires that proportional scaling alone of the evidence should not influence the assessment disproportionately towards one party or the other.

Finally, the last property reflects a notion of divisibility of the evidence. Imagine that the case that the plaintiff provides consists of a list of items, the totality of which yields the assessment x; similarly, the case for the defendant consists of a list of items, the totality of which yields the assessment y. Now assume that each of these lists could be (arbitrarily) divided into two parts so that, say, for the plaintiff the totality of the evidence on the first part of the list yielded assessment u and that on the second part of the list yielded assessment v (this does not presume that $x = u + v$, nor does this presume that $x > u$ and v). Similarly, let the sublist assessments for the defendant be denoted w and z. The fifth property asserts that the court's decision process should be able to come to the same conclusion by comparison of the sublists, followed by comparisons of the assessments based on the sublists, independently of how the sublists are compared. Thus, note that on the left u is compared with w and v with z, whereas on the right, w and z have been switched. Essentially, this axiom removes any role for psychology or "style" from the court's assessment procedure: the court's assessment is not influenced by how the evidence is presented. Alternatively, we assume that all parties to a dispute have access to equally qualified litigators, and that any such effects wash out of the court's assessment.

In Daughety and Reinganum (1998a), we show that these axioms combine to provide a unique family of continuous functions (for a proof, see Daughety and Reinganum [1998a]), indexed by the parameter q.

Theorem 1. The family of functions, indexed by the parameter q, given by: (a) $\ell(x, y; q) = \{(x^q + y^q)/2\}^{1/q}$, for q in $(-\infty, 1]$ such that $q \neq 0$; (b) $\ell(x, y; 0) = (xy)^{.5}$, is the unique family of continuous functions satisfying (i)–(v) above.

In Daughety and Reinganum (1998a), we show that $\ell(x, y; q)$ is strictly increasing in q when $x \neq y$. Thus, for a given evidence pair (x, y) (with $x \neq y$), higher values of q are associated with higher assessments, by T, of D's probability of being liable. At one end of the spectrum (as $q \to -\infty$), T's assessment is $\min\{x, y\}$. At the other end of the spectrum ($q = 1$), T's assessment is $(x + y)/2$. Intuitively, the closer the separately-provided evidence is (the closer x is to y), the less difference it makes what value of q is used, while the greater the disparity between x and y, the greater the sensitivity of T's assessment to the value of q that is used.[14]

We interpret q as a parameter reflecting basic legal properties of the assessment process. For example, notions such as duty, cause, foreseeability and proportionality may be either broadly or narrowly construed. Narrow interpretations (of, for example, to whom a defendant owes a duty) are more likely to lead to his being found not liable. Thus, a narrow interpretation corresponds to a lower value of q, whereas a broader interpretation corresponds to a higher value.

A basic question is, What value of q should be used? Assume that there is a supreme court (that is, a court superior to the trial and appeals courts) and that there is a preferred value of q, denoted q_S. This value is unknown to the litigants and the trial court, but we assume these agents have a common prior probability density over q_S. At the trial court level, we assume that neither the court nor the litigants have any private information about q_S. This is a plausible assumption for the trial court, since

14. The function $\ell(x, y; q)$ lends itself to two interpretations. First, it is a *quasi-arithmetic mean*, sometimes called *the mean of order* q. For example, when $q = 1$, we have the arithmetic mean of x and y, $q = 0$ yields the geometric mean, $q = -1$ the harmonic mean and $q \to -\infty$ yields the minimum of x and y. Alternatively, $\ell(x, y; q)$ is a *CES production function*, which is symmetric in the inputs, with the parameter multiplying the inputs equal to 1/2. In this case $q = 1$ corresponds to a linear production function, $q = 0$ to the Cobb-Douglas production function and $q \to -\infty$ yields the Leontief production function. Finally, note that eliminating axiom (ii) means that q ranges from negative to positive infinity. In particular, as q approaches positive infinity, $\ell(x, y; q)$ approaches $\max\{x, y\}$, an implication inconsistent with the notion that the burden is on the plaintiff to prove her case. Thus, in what follows, we maintain axiom (ii).

they do not receive resources to support legal research; their best guess about q_S is its unconditional mean. If we assume that the trial court is motivated by accuracy, then it will use this expectation, denoted q_T, in its liability assessment and determination.[15] Thus, in choosing what evidence to present, both the plaintiff and the defendant make use of the known family of assessment functions $\ell(x, y; q)$, the value q_T and the evidentiary standard to gather their evidence x and y, respectively, which is then presented at trial.

This brings us to the third element of the trial process: the determination of whether D will be held liable. Let γ be an exogenously specified evidentiary standard, with γ in $[0, 1]$. Three levels of γ are frequently employed: (1) "preponderance of the evidence" appears to translate to (approximately) 0.51 and applies in most civil cases; (2) "clear and convincing" is murkier and appears to lie somewhere in the interval $[0.6, 0.8]$ and is used by the majority of states in punitive damages cases; and (3) "beyond a reasonable doubt" involves a very high level of γ (say, 0.99) and is used in criminal cases and by Colorado in punitive damages cases.

The outcome of the trial is determined by the evidence, the assessment function and the evidentiary standard to answer the question: is the defendant liable? The liability determination function, $L(\ell(x, y; q); \gamma)$, is:

$$L(\ell(x, y; q); \gamma) = \begin{cases} 1 & \text{if } \ell(x, y; q) \geq \gamma \\ 0 & \text{otherwise,} \end{cases}$$

where $L = 1$ means that D is found liable, and $L = 0$ means D is found not liable.

The outcome of the trial may generate an appeal; we focus on appeals about the value of q alone (that is, the appeal is about legal issues, and not about the evidence (x, y) or the evidentiary standard γ). Given the assessment function and the evidentiary standard, notice that if (x, y) is such that D will always be found not liable, independent of the value of q used (that is, $L(\ell(x, y; q), \gamma) = 0$ for all q in $(-\infty, 1]$ or, equivalently, $(x + y)/2 < \gamma$), then we will assume that no appeals will ever arise in this case. In such cases P's evidence is too weak relative to D's. Alternatively,

15. While one could imagine the litigants observing private signals about q_s as a consequence of their own legal research, in Daughety and Reinganum (1998a) we show that this private information will have no impact, in equilibrium, on a trial court's decision in the absence of a private signal of its own.

given the assessment function and the evidentiary standard, if (x, y) is such that D will always be found liable, independent of the value of q used (that is, $L(\ell(x, y; q), \gamma) = 1$ for all q in $(-\infty, 1)$ or, equivalently, $\min\{x, y\} \geq \gamma$), then we will assume that no appeals will ever arise in this case, too. In these cases, D's evidence is too weak relative to P's.

Between these two alternatives lies a region of (x, y) pairs such that the value of q influences the outcome of the trial (given the evidentiary standard). In this case, for a given (x, y, γ) there is a specific value of q (which we denote as $q_{\min}$) such that $L(\ell(x, y; q), \gamma) = 0$ for all $q < q_{\min}$ and $L(\ell(x, y; q), \gamma) = 1$ for all $q \geq q_{\min}$. This value of q will play a pivotal role in the analysis presented in the next section.

To summarize, T compares q_T to $q_{\min}$ to determine liability; the record of the trial is the tuple (x, y, γ, L), describing the evidence provided by each party, the evidentiary standard employed and the trial's outcome. To simplify matters, we only consider appeals that are due to cases wherein $L = 1$ at trial. In Daughety and Reinganum (1998a), we developed a detailed model of D's choice to appeal a loss and an individual appeals court's decision as to whether to affirm or reverse T. There, both the potential appellant and the appeals court receive private signals about how the supreme court would decide the case; the appeals court uses its private signal and the appellant's decision to appeal to determine (what it considers to be) the "right" parameter value q, and to affirm or reverse the trial court's decision on this basis. In this article, we build on the previous one by considering a sequence of appeals courts evaluating the decisions of a sequence of trial courts in a sequence of substantially similar cases. So as to avoid unnecessary complexity, in the next section we assume that it is a dominant strategy for the defendant to appeal his loss at trial to the relevant appeals court.

4. A Model of the Impact of Persuasive Influence

We assume there is a collection of n courts that operate at the same level (e.g., federal circuit courts of appeals). Thus, there is no binding precedent among these courts. Superior to these n courts is a supreme court, whose decisions will be binding on the collection of courts below. All n courts are assumed to be rational and nonstrategic; that is, they are fully Bayesian and use all available information, *and* they do not attempt to influence the decisions to be made by other courts (appeals or supreme). We assume that a particular type of case is considered in sequence by this collection of courts. Specifically, we assume that the case records are identical in each instance of this type of case, where the case record is given by $(x, y, \gamma, 1)$.

At some point, this type of case may also be considered by the supreme court. Suppose that the supreme court, if it were to consider this case, would use the parameter q_S to determine its assessment of the evidence and, ultimately, the outcome of the case ($L_S = 0$ or $L_S = 1$). We assume that the appeals courts do not know the true value of q_S, but each observes a private signal, denoted q_i for Court i, which is correlated with q_S (for example, this may be viewed as the outcome of legal research performed by Court i). In addition, we assume that these private signals are independent conditional on q_S, and that the courts share a common prior density over q_S, denoted $h_S(q_S)$. This means that the conditional density function for q_i given q_S is the same for all i; denote it by $g(q_i|q_S)$. Finally, this means that the joint density function for the random variables $(q_S, q_1, \ldots, q_n)$ can be written as $f(q_S, q_1, \ldots, q_n) = h_S(q_S)g(q_1|q_S)\ldots g(q_n|q_S)$.[16]

We further assume (Assumption 1; see the Appendix for a formal statement) that the function $g(q_i|q_S)$ satisfies the monotone likelihood ratio property. This property captures the intuitive notion that a higher private signal to Court i is more likely to be associated with a higher value of the supreme court's preferred value q_S. This property implies (Milgrom and Weber, 1982, p. 1099) that the random variables $(q_S, q_1, \ldots, q_n)$, and all subsets of these random variables, are *affiliated.* As a consequence, each court that considers its respective case can use its private signal (and any other publicly observable information) to update its beliefs about q_S. In particular, Milgrom and Weber show (1982, Theorem 5, p. 1100) that the *expected value* of q_S, *given* that q_1 is in $[a_1, b_1], \ldots,$ and q_n is in $[a_n, b_n]$, which we denote by $E_n(q_S|q_1 \in [a_1, b_1], \ldots, q_n \in [a_n, b_n])$, is a nondecreasing function of $a_1, b_1, a_2, b_2, \ldots, a_n$ and b_n; note that in our application it is also continuous. Since affiliation is a rather weak relationship (for instance, independent random variables are affiliated), we will make a further strengthening assumption (Assumption 2; see the Appendix for a for-

16. All densities are assumed to be continuous with positive support on $(-\infty, 1]$.

mal statement): this function is strictly increasing in $a_1, b_1, a_2, b_2, \ldots, a_n$ and b_n.

The order in which the courts consider their respective cases depends on the behavior of litigants in each jurisdiction, since a case must be brought before it can be considered. Without loss of generality, we label the first court to consider its case Court 1, the second to consider its case Court 2, and so on. Thus, the information available to Court i when it makes its decision is the prior density $h_S(q_S)$, the action taken by the trial court (which is uninformative about q_S, and which we therefore suppress for notational simplicity), the actions taken by Courts $1, 2, \ldots, i-1$, and its private signal q_i. Let the decision taken by Court i be denoted L_i, where $L_i = 1$ means Court i affirms while $L_i = 0$ means Court i reverses.

Throughout we assume that Court i is striving for accuracy in the sense of using that value of q which minimizes the mean-squared distance from q_S, conditional on all information available to i at the time of decision.[17] Equivalently, Court i calculates its expectation of how the supreme court would assess the case (its best estimate of q_S) and uses this estimate in its own decision process. We denote Court i's best estimate of q_S as $B_i(q_i; L_1, \ldots, L_{i-1})$. This is consistent with assuming that Court i believes the supreme court has the ultimate truth, and that Court i is also pursuing this ultimate truth, or with the more prosaic assumption that Court i wants to minimize the embarrassment of a supreme court opinion that takes a substantially different interpretation than Court i.

In the analysis that follows we will be deriving court-specific "cutoffs" (indicated by overbars) which partition the space for the court's private signal (that is, $(-\infty, 1]$) into two regions, one for affirmance and the other for reversal. We focus on cutoffs that are *interior* (finite, but strictly less than 1).[18]

First, consider Court 1's decision. Its best estimate of q_S is given by $B_1(q_1) \equiv E_1(q_S|q_1)$ (see the Appendix for a precise description of Court 1's beliefs and expectations). Thus, Court 1 will decide[19] $L_1 = 0$

17. This objective implies that the court perceives an equal loss from over- and under-estimating q_S.

18. Cutoffs could become arbitrarily small (approach $-\infty$) or become 1, in which case the court in question, and all following it, will always make the same decision regardless of their private signals.

19. Recall the definition of $q_{\min}$ provided at the end of the previous section.

if $B_1(q_1) < q_{\min}$ and $L_1 = 1$ if $B_1(q_1) \geq q_{\min}$. Court 1's decision represents an affirmance of the trial court if $L_1 = 1$ and a reversal of the trial court if $L_1 = 0$. The function $B_1(q_1)$ is continuous and strictly increasing in q_1 (by Assumption 2). Thus, there exists a unique value of q_1 in $(-\infty, 1)$, denoted $\bar{q}_1$, such that $B_1(\bar{q}_1) = q_{\min}$. Thus, if Court 1 observes $q_1 < \bar{q}_1$, it chooses $L_1 = 0$ (reverses), while if Court 1 observes $q_1 \geq \bar{q}_1$, it chooses $L_1 = 1$ (affirms). Furthermore, observe that $\bar{q}_1$ is computable by all appeals courts.

Now consider a subsequent instance of (essentially) the same case which arises in another jurisdiction (e.g., another circuit). Court 2, which is to decide this case, is not bound by Court 1's decision, but *it may find it influential.* This is because Court 1's action reveals something about Court 1's private signal (which is affiliated with Court 2's signal), namely, whether q_1 was greater than or less than $\bar{q}_1$. Court 2 finds this information of interest because it helps Court 2 to further refine its beliefs regarding q_S.[20] Upon observing L_1 and its own private signal q_2, Court 2 forms its posterior beliefs about q_S. Note that, since none of the courts are strategic (in the sense that none of them are trying to influence the decisions of the courts that follow them), common knowledge of the assessment function $\ell(x, y; q)$ and the role of q, as well as the observed decision by Court 1, means that there is no need to disentangle information and strategy. Thus, observing that (say) $L_1 = 0$ simply means that $q_1 < \bar{q}_1$. Court 2's best estimate of q_S is given by $B_2(q_2; L_1) \equiv E_2(q_S | q_2, L_1)$ (again, see the Appendix for a precise description of Court 2's beliefs and expectations).

The function $B_2(q_2; L_1)$ is continuous in q_2 and (by Assumption 2) is strictly increasing in both q_2 and L_1. Thus, there exists a unique value of q_2 in $(-\infty, 1)$, denoted $\bar{q}_2(L_1)$, such that $B_2(\bar{q}_2(L_1); L_1) = q_{\min}$. Therefore, if Court 2 observes $q_2 < \bar{q}_2(L_1)$, it chooses $L_2 = 0$, while if Court 2 observes $q_2 \geq \bar{q}_2(L_1)$, it chooses $L_2 = 1$. Note, however, that now the cutoff value of q_2 *depends on the previous decision*; this is because, while it was assumed that the trial court's decision was not informative regarding q_S, we specifically assume that Court 1's decision is based on its private signal regarding q_S, which is therefore informative to Court 2. In particular, since both $\bar{q}_2(1)$ and $\bar{q}_2(0)$ are assumed to be interior, they are defined

20. We currently assume that opinions, in and of themselves, add no further information. We address this issue in the next section.

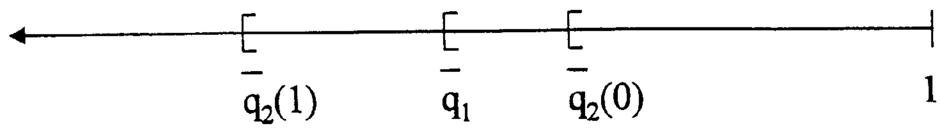

Figure 1. Intervals of affirmance/reversal.

implicitly by $B_2(\bar{q}_2(L_1); L_1) = q_{\min}$. Then since $B_2(q_2; 1) > B_2(q_2; 0)$, it follows that $\bar{q}_2(1) < \bar{q}_2(0)$. This means that Court 2 is more willing to choose $L_2 = 1$ (i.e., it chooses $L_2 = 1$ for a greater range of q_2 values) after observing $L_1 = 1$ than after observing $L_1 = 0$. Alternatively put, Court 2 is more likely to affirm its corresponding trial court decision if Court 1 has affirmed its trial court's decision. Similarly, Court 2 is more willing to choose $L_2 = 0$ (i.e., it chooses $L_2 = 0$ for a greater range of q_2 values) after observing $L_1 = 0$ than after observing $L_1 = 1$. Alternatively put, Court 2 is more likely to reverse its corresponding trial court decision if Court 1 has reversed its trial court's decision. In this sense, Court 1's decision represents *persuasive influence*; it influences, but does not dictate, Court 2's decision.

It is further shown (in the Appendix) that $\bar{q}_2(1) < \bar{q}_1 < \bar{q}_2(0)$. Thus, Court 2 is more likely to make the same decision as Court 1 when it observes Court 1's decision than when it acts independently on the basis of its own private signal alone (in which case Court 2 would also use the value $\bar{q}_1$ to partition the interval $(-\infty, 1]$ into regions wherein $L_2 = 0$ versus $L_2 = 1$). These relationships are illustrated in Figure 1.

In the Figure, the right-facing brackets ([) indicate that any signal received in the interval from the bracket up to (and including) 1 results in an affirmance for the court in question, with the cutoffs for Court 2 indicating their history-dependence on the outcome at Court 1.

Next, consider Court 3's decision. Upon observing the history of decisions (L_1, L_2) and its private signal q_3, reasoning as we did above for Court 2 implies that Court 3's best estimate of q_S is given by $B_3(q_3; L_1, L_2) \equiv E_3(q_S | q_3, L_1, L_2)$ (again, see the Appendix for a precise description of Court 3's beliefs and expectations). The expectation is continuous in q_3 and a strictly increasing function of q_3, L_1 and L_2 by Assumption 2. It is clear that $B_3(q_3; 1, 1) > B_3(q_3; 1, 0) > B_3(q_3; 0, 0)$ and $B_3(q_3; 1, 1) > B_3(q_3; 0, 1) > B_3(q_3; 0, 0)$. It is shown in the Appendix that $B_3(q_3; 0, 1) > B_3(q_3; 1, 0)$: when the appellate decision history is mixed, it is the most recent decision that is most influential.

That is, a mixed appellate decision history of one reversal (in Court 1) followed by one affirmance (in Court 2) results in a higher estimate of q_S than a mixed appellate decision history of one affirmance (in Court 1) followed by one reversal (in Court 2). This makes sense because the presence of influence makes it less likely that Court 2 will disagree with Court 1; thus when Court 2 is observed to disagree with Court 1, this public signal is more informative than if there were no influence.

Thus, Court 3's best estimate of q_S conditional on the history (L_1, L_2) can be ordered as follows: $B_3(q_3; 1, 1) > B_3(q_3; 0, 1) > B_3(q_3; 1, 0) > B_3(q_3; 0, 0)$. The interior cutoffs, denoted $\bar{q}_3(L_1, L_2)$, are defined implicitly by $B_3(\bar{q}_3(L_1, L_2); L_1, L_2) = q_{\min}$. It can be shown (see the Appendix) that $\bar{q}_3(1, 1) < \bar{q}_3(0, 1) < \bar{q}_3(1, 0) < \bar{q}_3(0, 0)$. Thus, Court 3 is most likely to affirm its respective case (i.e., set $L_3 = 1$) if both previous appellate courts have also affirmed, and most likely to reverse its respective case (i.e., set $L_3 = 0$) if both previous appellate courts have also reversed. In cases of mixed appellate decision histories, Court 3 is more likely to affirm if Court 1 reversed but Court 2 affirmed than if Court 1 affirmed but Court 2 reversed.

In addition, it is shown in the Appendix that $\bar{q}_3(1, 1) < \bar{q}_2(1)$ and $\bar{q}_2(0) < \bar{q}_3(0, 0)$. That is, Court 3 is more willing to affirm after observing a history of two previous affirmances than Court 2 would be after observing a history of one previous affirmance. Conversely, Court 3 is more willing to reverse after observing a history of two previous reversals than Court 2 would be after observing a history of one previous reversal. Thus, courts progressively begin to rely upon the previous decisions of other courts, which is the essence of herding.[21]

It is easy to show (by duplicating the proof for $n = 3$ in the Appendix) that this latter result is quite general (assuming that all these expressions are interior): $\bar{q}_n(1, \ldots, 1) < \bar{q}_{n-1}(1, \ldots, 1) < \cdots < \bar{q}_1 < \cdots < \bar{q}_{n-1}(0, \ldots, 0) < \bar{q}_n(0, \ldots, 0)$. This is illustrated in Figure 2.

21. Technically, the herding literature classifies herding as occurring when agents fully ignore their private signals; here, the private signals must be (progressively) more extreme in order to have any influence on a decision. Bikchandani, Hirshleifer, and Welch (1992) show that as the number of decision-makers grows without bound, the probability that herding will eventually occur goes to one. In our setting the number of decision-makers (courts) is finite, so this extreme form of herding is less likely (though not impossible).

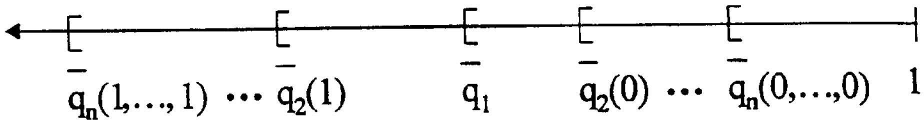

Figure 2. Uniform histories and affirmance/reversal.

Again, the brackets indicate that signals in the interval to the right of (and including) the bracket result in affirmance. Alternatively, for those cutoffs indicating a history of reversals to that point, a signal strictly to the left of the bracket results in reversal for the indicated court.

5. Implications

The analysis above has several interesting implications. First, previous decisions (which were assumed to be rational and nonstrategic) are influential because the random variables $(q_S, q_1, \ldots, q_n)$ are affiliated, and rational behavior by Court i means that this relationship should be exploited by incorporating the public information contained in the decisions by Courts 1 through $i-1$. Second, sequences of outcomes such as $(1, \ldots, 1)$ and $(0, \ldots, 0)$ are more likely (and mixed sequences are less likely) when courts are influenced by previous decisions than if they were to decide each case independently. Thus, herding increases the likelihood of agreement and decreases the likelihood of conflict among the appeals courts.[22] Since there is a "correct" decision (either that implied by q_S in a first-best scenario or that implied by the best estimate of q_S given the realized signals $(q_1, \ldots, q_n)$ in a second-best analysis), this means that the "correct" sequence (all ones or all zeros, as appropriate) is more likely to occur—but so is the incorrect sequence (all zeros or all ones, as appropriate)—when courts are influenced by previous decisions than if they were to decide each case independently (recall that the likelihood

22. Richard Posner and Nicholas Zeppos have pointed out to us the following additional reasons (which are outside our model) for harmony among appeals courts: (1) most circuits require a panel to consult with the full court if a decision would create an intercircuit conflict; (2) inasmuch as the U.S. Supreme Court may be called upon to resolve intercircuit conflicts, such conflicts should be avoided when reasonable; and (3) harmony contributes to certainty. Posner (1990, p. 458) also argues that the existence of a harmonious decision arising from a diverse (as compared with a homogeneous) set of courts is more informative. This issue is beyond the current model, as here all courts have the same conditional likelihood function g and objective.

of mixed sequences will fall). Third, later decisions that contradict earlier ones are more informative—because this pattern of outcomes is less likely to occur—when courts are influenced by previous decisions than if they were to decide each case independently. Fourth, a form of path dependence, which we will dub *interjurisdictional path dependence*, arises, since the sequence of outcomes may be critically dependent on the order in which the private signals arrive. For instance, if courts that address the issue early in the sequence receive low private signals (and those that address the issue later receive higher private signals), this may result in a sequence such as $(0, \ldots, 0)$ since the early signals cause decisions which drive the minimum signal needed for affirmance (the cutoffs) closer to one; on the other hand, if the high signals are received early in the sequence (and the lower signals are received later), this could well result in a sequence such as $(1, \ldots, 1)$. This occurs in both cases because later decisions become more heavily-reliant on public information (i.e., the sequence of previous decisions) relative to private information.

Fifth, the observation that a collection of courts agrees on an outcome cannot be taken as indicating that this outcome is the correct one. Moreover, such agreement may discourage further appeal, since potential appellants are also trying to estimate what the supreme court would do, based on the sequence of outcomes from Court 1, Court 2, ..., as well as private signals of their own. If further appeal is completely discouraged, then the error never gets corrected; if it is simply delayed until an appellant receives an extremely strong (and contradictory to the history of decisions) private signal, then the correction is correspondingly delayed. This occurs because the supreme court must wait for an appeal, rather than reaching down and designating cases for review, and an appellant would need a sufficiently extreme signal (and a sufficiently low cost of appeal) to make the appeal worthwhile. It does suggest, however, that an appeal to the supreme court after a sequence of outcomes such as $(1, \ldots, 1)$ or $(0, \ldots, 0)$ should not be denied cert lightly, as appeals court harmony may well correspond to herding on the "wrong" outcome.[23]

23. Stern, Gressman, and Shapiro (1986) indicate that cert is often (though not always) granted when there is conflict among the appeals courts. While the Supreme Court may also grant cert as a check against extensive harmony, we've not seen any discussion on this point.

In our discussion of the Circuit Court cases before *Eastern Enterprises v. Apfel* we noted that three earlier cases (*Davon, Blue Diamond* and *Lindsey Coal*) appear to share a critical attribute of *Eastern Enterprises* (extensive retroactivity in light of no explicit promise). The first two of these were denied cert and the third apparently did not pursue it. Our model does not address cert decisions by a supreme court.[24] We feel that this string of cases is consistent with the predictions of our model: each opinion on the same set of issues became progressively briefer and endorsed the arguments made in previous opinions. Interestingly, the Pesto/Smith opinions in *Unity*, which cited (and seemed to be responding to) statements made by members of the Supreme Court in earlier cases (statements subsequently mentioned in O'Connor's opinion in *Eastern*) were available to the Third Circuit (in fact, Pesto and Smith are in the Third Circuit) when they considered *Lindsey Coal*, but were not cited (though the *Davon* and *Blue Diamond* cases, which were cited in *Lindsey Coal*, cited *Unity* as arguing that the Coal Act effected an uncompensated taking). The information in previous Supreme Court opinions (about q_S) seems to have been disregarded, or at least overwhelmed.

Finally, we have assumed that the public signal is the outcome of affirmance or reversal. While opinions might provide more information, this is likely only to improve the interval estimates for Court i of the previous Courts' signals, since opinions are not likely to be sufficiently precise as to generate perfect revelation. Moreover, as Lynn (1993, p. 21) observes, "A common response of appellate courts to the growing burden of opinion writing, because of caseload pressures, has been to increase the number of cases decided without opinion or to issue sketchy per curiam opinions which reveal little." This practice may contribute to a greater likelihood of herding on outcomes, some of which should be corrected, but may never reach the Supreme Court.

Appendix

Assumption 1. (Monotone Likelihood Ratio Property) For all $q_i' > q_i$ and $q_S' > q_S$, $[g(q_i|q_S)/g(q_i|q_S')] \geq [g(q_i'|q_S)/g(q_i'|q_S')]$, $i = 1, 2, \ldots, n$.

24. There was no dissent written in these cert denials, so the reasons for denial are unknown; see Epstein and Knight (1998) for a discussion of the strategic considerations that come into play in the Supreme Court's cert decision.

Assumption 2. $E_n(q_S|q_1 \in [a_1, b_1], \ldots, q_n \in [a_n, b_n])$ is known to be a non-decreasing function of $\{a_i, b_i\}_{i=1}^n$, for all n; we hereby assume that it is a strictly increasing function of $\{a_i, b_i\}_{i=1}^n$, for all n.

In what follows, wherever there is an integral, its domain of integration is $(-\infty, 1]$.

Court 1's Beliefs and Expectations:

Court 1's posterior beliefs about q_S given q_1 are:

$$\mu_1(q_S|q_1) = [h_S(q_S)g(q_1|q_S)]/[\int h_S(r)g(q_1|r)\,dr].$$

Then $E_1(q_s|q_1) = \int q_S\mu_1(q_s|q_1)\,dq_s$.

Court 2's Beliefs and Expectations:

Court 2's posterior beliefs about q_s given q_2 and L_1 are given by:

$$\mu_2(q_S|q_2, 1) = \{h_S(q_S)[1 - G(\bar{q}_1|q_S)]g(q_2|q_S)\} \div \{\int h_S(r)[1 - G(\bar{q}_1|r)]g(q_2|r)\,dr\}$$

and

$$\mu_2(q_S|q_2, 0) = \{h_S(q_S)[G(\bar{q}_1|q_S)]g(q_2|q_S)\} \div \{\int h_S(r)[G(\bar{q}_1|r)]g(q_2|r)\,dr\},$$

where $G(q_i|q_S)$ is the cumulative distribution function of $g(q_i|q_S)$. Then $E_2(q_S|q_2, L_1) = \int q_S\mu_2(q_S|q_2, L_1)\,dq_S$. Equivalently, we can write $E_2(q_S|q_2, 1) = E_2(q_S|q_2, q_1 \in [\bar{q}_1, 1])$ and $E_2(q_S|q_2, 0) = E_2(q_S|q_2, q_1 \in (-\infty, \bar{q}_1))$.

Claim 1. $\bar{q}_2(1) < \bar{q}_1 < \bar{q}_2(0)$.

Proof: First, we show that $\bar{q}_2(1) < \bar{q}_1$. Recall that $\bar{q}_2(1)$ is defined by $B_2(\bar{q}_2(1); 1) = q_{\min}$, and $\bar{q}_1$ is defined by $B_1(\bar{q}_1) = q_{\min}$.

$$\begin{aligned} B_2(q_2; 1) &= E_2(q_S|q_2, 1) = E_2(q_S|q_2, q_1 \in [\bar{q}_1, 1]) \\ &> E_2(q_S|q_2, q_1 \in (-\infty, 1]) \qquad \text{(by Assumption 2)} \\ &= E_1(q_S|q_2) \qquad \text{(knowing that } q_1 \in (-\infty, 1] \text{ adds no information)} \\ &= B_1(q_2). \end{aligned}$$

Evaluating both sides at $\bar{q}_1$ yields: $B_2(\bar{q}_1; 1) > B_1(\bar{q}_1) = q_{\min} = B_2(\bar{q}_2(1); 1)$; since $B_2(q_2; 1)$ is increasing in q_2, it follows that $\bar{q}_2(1) < \bar{q}_1$. Next, we show that $\bar{q}_1 < \bar{q}_2(0)$.

$$\begin{aligned} B_2(q_2; 0) &= E_2(q_S|q_2, 0) = E_2(q_S|q_2, q_1 \in (-\infty, \bar{q}_1)) \\ &< E_2(q_S|q_2, q_1 \in (-\infty, 1]) \qquad \text{(by Assumption 2)} \end{aligned}$$

$$= E_1(q_S|q_2) \qquad \text{(knowing that } q_1 \in (-\infty, 1] \text{ adds no information)}$$

$$= B_1(q_2).$$

Evaluating both sides at $\bar{q}_1$ yields: $B_2(\bar{q}_1;0) < B_1(\bar{q}_1) = q_{\min} = B_2(\bar{q}_2(0);0)$; since $B_2(q_2;0)$ is increasing in q_2, it follows that $\bar{q}_1 < \bar{q}_2(0)$. QED

Court 3's Beliefs and Expectations:

Court 3's posterior beliefs about q_S given q_3, L_1 and L_2 are given by:

$$\mu_3(q_S|q_3, 1, 1) = \{h_S(q_S)[1 - G(\bar{q}_1|q_S)][1 - G(\bar{q}_2(1)|q_S)]g(q_3|q_S)\} \div \{\textstyle\int h_S(r)[1 - G(\bar{q}_1|r)][1 - G(\bar{q}_2(1)|r)]g(q_3|r)\,dr\},$$

$$\mu_3(q_S|q_3, 1, 0) = \{h_S(q_S)[1 - G(\bar{q}_1|q_S)][G(\bar{q}_2(1)|q_S)]g(q_3|q_S)\} \div \{\textstyle\int h_S(r)[1 - G(\bar{q}_1|r)][G(\bar{q}_2(1)|r)]g(q_3|r)\,dr\},$$

$$\mu_3(q_S|q_3, 0, 1) = \{h_S(q_S)[G(\bar{q}_1|q_S)][1 - G(\bar{q}_2(0)|q_S)]g(q_3|q_S)\} \div \{\textstyle\int h_S(r)[G(\bar{q}_1|r)][1 - G(\bar{q}_2(0)|r)]g(q_3|r)\,dr\},$$

and

$$\mu_3(q_S|q_3, 0, 0) = \{h_S(q_S)[G(\bar{q}_1|q_S)][G(\bar{q}_2(0)|q_S)]g(q_3|q_S)\} \div \{\textstyle\int h_S(r)[G(\bar{q}_1|r)][G(\bar{q}_2(0)|r)]g(q_3|r)\,dr\}.$$

Then $E_3(q_S|q_3, L_1, L_2) = \int q_S\mu_3(q_S|q_3, L_1, L_2)\,dq_S$. Equivalently, we can write $E_3(q_S|q_3, L_1, L_2)$ for the four possible histories as follows:

$$E_3(q_S|q_3, 1, 1) = E_3(q_S|q_3, q_1 \in [\bar{q}_1, 1], q_2 \in [\bar{q}_2(1), 1])$$

$$E_3(q_S|q_3, 1, 0) = E_3(q_S|q_3, q_1 \in [\bar{q}_1, 1], q_2 \in (-\infty, \bar{q}_2(1)))$$

$$E_3(q_S|q_3, 0, 1) = E_3(q_S|q_3, q_1 \in (-\infty, \bar{q}_1), q_2 \in [\bar{q}_2(0), 1])$$

$$E_3(q_S|q_3, 0, 0) = E_3(q_S|q_3, q_1 \in (-\infty, \bar{q}_1), q_2 \in (-\infty, \bar{q}_2(0))).$$

Claim 2. $B_3(q_3;0,1) > B_3(q_3;1,0)$.

Proof:

$$B_3(q_3;1,0) = E_3(q_S|q_3, q_1 \in [\bar{q}_1, 1], q_2 \in (-\infty, \bar{q}_2(1)))$$

$$= E_3(q_S|q_3, q_1 \in (-\infty, \bar{q}_2(1)), q_2 \in [\bar{q}_1, 1]) \qquad (q_1 \text{ and } q_2 \text{ are i.i.d. given } q_S)$$

$$< E_3(q_S|q_3, q_1 \in (-\infty, \bar{q}_1), q_2 \in [\bar{q}_2(0), 1]) \qquad (\text{Assumption 2 and } \bar{q}_2(0) > \bar{q}_1 > \bar{q}_2(1))$$

$$= B_3(q_3;0,1). \quad \text{QED}$$

Claim 3. $\bar{q}_3(1,1) < \bar{q}_2(1)$ and $\bar{q}_2(0) < \bar{q}_3(0,0)$.

Proof: First we show that $\bar{q}_3(1,1) < \bar{q}_2(1)$.

$$B_3(q_3;1,1) = E_3(q_S|q_3, q_1 \in [\bar{q}_1, 1], q_2 \in [\bar{q}_2(1), 1])$$

$$> E_3(q_S|q_3, q_1 \in [\bar{q}_1, 1], q_2 \in (-\infty, 1])$$

(by Assumption 2)

$$= E_2(q_S|q_3, q_1 \in [\bar{q}_1, 1])$$

(knowing that $q_2 \in (-\infty, 1)$ adds no information)

$$= B_2(q_3;1).$$

Evaluating both sides at $\bar{q}_2(1)$ yields: $B_3(\bar{q}_2(1);1,1) > B_2(\bar{q}_2(q);1) = q_{\min} = B_3(\bar{q}_3(1,1);1,1)$; since $B_3(\bar{q}_3;1,1)$ is increasing in q_3, it follows that $\bar{q}_3(1,1) < \bar{q}_2(1)$. Next we show that $\bar{q}_2(0) < \bar{q}_3(0,0)$.

$$B_3(q_3;0,0) = E_3(q_S|q_3, q_1 \in (-\infty, \bar{q}_1), q_2 \in (-\infty, \bar{q}_2(0)))$$

$$< E_3(q_S|q_3, q_1 \in (-\infty, \bar{q}_1), q_2 \in (-\infty, 1])$$

(by Assumption 2)

$$= E_2(q_S|q_3, q_1 \in (-\infty, \bar{q}_1))$$

(knowing that $q_2 \in (-\infty, 1]$ adds no information)

$$= B_2(q_3;0).$$

Evaluating both sides at $\bar{q}_2(0)$ yields: $B_3(\bar{q}_2(0);0,0) < B_2(\bar{q}_2(0);0) = q_{\min} = B_3(\bar{q}_3(0,0);0,0)$; since $B_3(q_3;0,0)$ is increasing in q_3, it follows that $\bar{q}_2(0) < \bar{q}_3(0,0)$. QED

References

Ashenfelter, Orley, Theodore Eisenberg, and Stewart J. Schwab. 1995. "Politics and the Judiciary: The Influence of Judicial Background on Case Outcomes," 24 *Journal of Legal Studies* 257–81.

Banerjee, Abhijit V. 1992. "A Simple Model of Herd Behavior," 107 *Quarterly Journal of Economics* 797–817.

Bikchandani, Sushil, David Hirshleifer, and Ivo Welch. 1992. "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades," 100 *Journal of Political Economy* 992–1026.

———. 1998. "Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades," 12 *Journal of Economic Perspectives* 151–70.

Braeutigam, Ronald, Bruce M. Owen, and John Panzar. 1984. "An Economic Analysis of Alternative Fee-Shifting Systems," 47 *Law and Contemporary Problems* 173–85.

Bulow, Jeremy, and Paul Klemperer. 1994. "Rational Frenzies and Crashes," 102 *Journal of Political Economy* 1–23.

Chamley, Christophe, and Douglas Gale. 1994. "Information Revelation and Strategic Delay in a Model of Investment," 62 *Econometrica* 1065–86.

Choi, Jay Pil. 1997. "Herd Behavior, the 'Penguin Effect,' and the Suppression of Informational Diffusion: An Analysis of Informational Externalities and Payoff Interdependency," 28 *Rand Journal of Economics* 401–25.

Cohen, Mark A. 1992. "The Motives of Judges: Empirical Evidence from Antitrust Sentencing," 12 *International Review of Law and Economics* 13–30.

Cooter, Robert D., and Tom Ginsburg. 1996. "Comparative Judicial Discretion: An Empirical Test of Economic Models," 16 *International Review of Law and Economics* 295–314.

Cross, Frank B., and Emerson H. Tiller. 1998. "Judicial Partisanship and Obedience to Legal Doctrine: Whistleblowing on the Federal Courts of Appeals," 107 *Yale Law Journal* 2155–76.

Danzon, Patricia Munch. 1983. "Contingent Fees for Personal Injury Litigation," 14 *Bell Journal of Economics* 213–24.

Daughety, Andrew F., and Jennifer F. Reinganum. 1995. "Keeping Society in the Dark: On the Admissibility of Pretrial Negotiations as Evidence in Court," 26 *Rand Journal of Economics* 203–21.

———. 1998a. "Appealing Judgments," Working Paper No. 98-W06, Department of Economics, Vanderbilt University.

———. 1998b. "On the Economics of Trials: Adversarial Process, Evidence and Equilibrium Bias," Working Paper No. 98-W02, Department of Economics, Vanderbilt University.

Easterbrook, Frank H. 1982. "Ways of Criticizing the Court," 95 *Harvard Law Review* 802–32.

Epstein, Lee, and Jack Knight. 1998. *The Choices Justices Make*. Washington, DC: Congressional Quarterly.

Farmer, Amy, and Paul Pecorino. 1998. "Does Jury Bias Matter?" Working Paper, Department of Economics, University of Tennessee.

Ferejohn, John A., and Barry R. Weingast. 1992. "A Positive Theory of Statutory Interpretation," 12 *International Review of Law and Economics* 263–79.

Froeb, Luke M., and Bruce H. Kobayashi. 1993. "Competition in the Production of Costly Information: An Economic Analysis of Adversarial versus Court-Appointed Presentation of Expert Testimony," Working Paper, Owen Graduate School of Management, Vanderbilt University.

——— 1996. "Naive, Biased, yet Bayesian: Can Juries Interpret Selectively Produced Evidence?" 12 *Journal of Law, Economics, and Organization* 257–76.

Gely, Rafael, and Pablo T. Spiller. 1990. "A Rational Choice Theory of Supreme Court Statutory Decisions with Applications to the *State Farm* and *Grove City* Cases," 6 *Journal of Law, Economics, and Organization* 263–300.

George, Tracey E., and Lee Epstein. 1992. "On the Nature of Supreme Court Decision Making," 86 *American Political Science Review* 323–37.

Gul, Faruk, and Russell Lundholm. 1995. "Endogenous Timing and the Clustering of Agents' Decisions," 103 *Journal of Political Economy* 1039–66.

Hastie, Reid, ed. 1993. *Inside the Juror*. Cambridge: Cambridge University Press.

Hause, John C. 1989. "Indemnity, Settlement, and Litigation, or, I'll Be Suing You," 18 *Journal of Legal Studies* 157–79.

Katz, Avery. 1987. "Measuring the Demand for Litigation: Is the English Rule Really Cheaper?" 3 *Journal of Law, Economics, and Organization* 143–76.

———. 1988. "Judicial Decisionmaking and Litigation Expenditure," 8 *International Review of Law and Economics* 127–43.

Kornhauser, Lewis A. 1992a. "Modeling Collegial Courts I: Path Dependence," 12 *International Review of Law and Economics* 169–85.

———. 1992b. "Modeling Collegial Courts II: Legal Doctrine," 8 *Journal of Law, Economics, and Organization* 441–70.

Landes, William M. 1993. "Sequential versus Unitary Trials: An Economic Analysis," 22 *Journal of Legal Studies* 99–134.

Lee, In Ho. 1993. "On the Convergence of Informational Cascades," 61 *Journal of Economic Theory* 395–411.

Lipman, Barton L., and Duane J. Seppi. 1995. "Robust Inference in Communication Games with Partial Provability," 66 *Journal of Economic Theory* 370–405.

Lynn, Richardson R. 1993. *Appellate Litigation*. 2nd ed. San Francisco: Austin & Winfield.

Miceli, Thomas J., and Metin M. Cosgel. 1994. "Reputation and Judicial Decision-Making," 23 *Journal of Economic Behavior and Organization* 31–51.

Milgrom, Paul R., and Robert J. Weber. 1982. "A Theory of Auctions and Competitive Bidding," 50 *Econometrica* 1089–1122.

Milgrom, Paul, and John Roberts. 1986. "Relying on the Information of Interested Parties," 17 *Rand Journal of Economics* 18–32.

Miller, Geoffrey P. 1987. "Some Agency Problems in Settlement," 16 *Journal of Legal Studies* 189–215.

O'Hara, Erin. 1993. "Social Constraint or Implicit Collusion?: Toward a Game Theoretic Analysis of Stare Decisis," 24 *Seton Hall Law Review* 736–778.

Plott, Charles R. 1987. "Legal Fees: A Comparison of the American and English Rules," 3 *Journal of Law, Economics, and Organization* 185–92.

Posner, Richard A. 1990. *The Problems of Jurisprudence*. Cambridge, MA: Harvard University Press.

———. 1992. *Economic Analysis of Law*, 4th ed. Boston: Little, Brown.

———. 1993. "What Do Judges and Justices Maximize? (The Same Thing Everybody Else Does)," 3 *Supreme Court Economic Review* 1–42.

Rasmusen, Eric. 1994. "Judicial Legitimacy as a Repeated Game," 10 *Journal of Law, Economics, and Organization* 63–83.

Rubinfeld, Daniel L., and David E. M. Sappington. 1987. "Efficient Awards and Standards of Proof in Judicial Proceedings," 18 *Rand Journal of Economics* 308–15.

Sanchirico, Chris William. 1997a. "The Burden of Proof in Civil Litigation: A Simple Model of Mechanism Design," 16 *International Review of Law and Economics*, 431–47.

———. 1997b. "Evidence Production, Adversarial Process and the Private Instigation of Suits," Working Paper, Department of Economics, Columbia University.

———. 1997c. "An Integrated Model of Evidence Production," Working Paper, Department of Economics, Columbia University.

Scharfstein, David S., and Jeremy C. Stein. 1990. "Herd Behavior and Investment," 80 *American Economic Review* 465–79.

Segal, Jeffrey A., and Harold J. Spaeth. 1993. *The Supreme Court and the Attitudinal Model*. Cambridge: Cambridge University Press.

Seidmann, Daniel J., and Eyal Winter. 1997. "Strategic Information Transmission with Verifiable Messages," 55 *Econometrica* 163–169.

Shavell, Steven. 1995. "The Appeals Process as a Means of Error Correction," 25 *Journal of Legal Studies* 379–426.

Shin, Hyun Song. 1994. "The Burden of Proof in a Game of Persuasion," 64 *Journal of Economic Theory* 253–264.

———. 1998. "Adversarial and Inquisitorial Procedures in Arbitration," 29 *Rand Journal of Economics* 378–405.

Sobel, Joel. 1985. "Disclosure of Evidence and Resolution of Disputes: Who Should Bear the Burden of Proof?" in Alvin E. Roth, ed., *Game-Theoretic Models of Bargaining*. Cambridge: Cambridge University Press.

Spiller, Pablo T. 1992. "Commentary," 12 *International Review of Law and Economics* 186–90.

Spiller, Pablo T., and Emerson H. Tiller. 1996. "Invitations to Override: Congressional Reversals of Supreme Court Decisions," 16 *International Review of Law and Economics* 503–21.

Spitzer, Matt, and Eric Talley. 1998. "Judicial Auditing," Working Paper No. 98-22, University of Southern California Law School.

Stern, Robert L., Eugene Gressman, and Stephen M. Shapiro. 1986. *Supreme Court Practice*, 6th ed. Washington DC: Bureau of National Affairs.

Talley, Eric. 1996 (rev. 1998). "On Common Law Learning, Rational Emulation, and the Evolution of Legal Rules," Mimeo, University of Southern California Law School.

Toma, Eugenia F. 1996. "A Contractual Model of the Voting Behavior of the Supreme Court: The Role of the Chief Justice," 16 *International Review of Law and Economics* 433–48.

Welch, Ivo. 1992. "Sequential Sales, Learning and Cascades," 47 *The Journal of Finance* 695–732.

Case References

Barrick Gold Exploration, Inc. v. Hudson, 47 F.3d 832 (6th Cir.), February 24, 1995.

In re: Blue Diamond Coal Co. v. Shalala, 79 F.3d 516 (6th Cir.), March 21, 1996.

In re: Chateaugay Corp. v. Shalala, 53 F.3d 478 (2nd Cir.), April 17, 1995.

Davon, Inc. v. Shalala, 75 F.3d 1114 (7th Cir.), January 25, 1996.

Eastern Enterprises v. Apfel, 118 S. Ct. 2131, June 25, 1998.

Eastern Enterprises v. Chater, 110 F.3d 150 (1st Cir.), April 7, 1997.

Holland v. Keenan Trucking Co., 102 F.3d 736 (4th Cir.), December 17, 1996.

Lindsey Coal Mining Co. v. Chater, 90 F.3d 688 (3rd Cir.), July 26, 1996.

Unity Real Estate Co. v. Hudson, 889 F. Supp. 818 (W.D.Pa), December 7, 1994, January 11, 1995, June 7, 1995.

Unity Real Estate Co. v. Hudson, 977 F. Supp. 717 (W.D.Pa), March 14, 1997.