ISSN 1045-6333

CIVIL DISCOVERY: ITS EFFECTS AND OPTIMAL SCOPE

Bruce L. Hay

Discussion Paper No. 141

2/94

Harvard Law School
Cambridge, MA 02138

The Program in Law and Economics is supported by a grant from the John M. Olin Foundation.

CIVIL DISCOVERY: ITS EFFECTS AND OPTIMAL SCOPE

Bruce L. Hay

Abstract

The rules of discovery enable a litigant to compel his opponent to disclose, before trial, the evidence in his possession. The positive issue explored in this paper is how such rules affect the outcome of litigation. Two such effects are contrasted: providing parties with information about the (exogenously determined) likely result of a trial, and providing them with evidence that would not otherwise emerge (thus affecting the likely result of trial).

The normative issue explored is the proper scope of discovery, given the costs of demanding and disclosing information. A model is developed in which the scope of discovery affects the accuracy of judicial decisionmaking; judicial accuracy, in turn, is viewed as affecting the incentives of potential defendants to take precaution against harm. A test for discovery that is socially desirable *ex ante* (pre-accident) is contrasted with a test that focuses on the *ex post* (post-accident) costs and benefits to the parties. Finally, attention is given to some second-best issues raised by the possibility that parties may decide to settle without undertaking discovery.

CIVIL DISCOVERY: ITS EFFECTS AND OPTIMAL SCOPE

Bruce L. Hay*

In this paper I use Cooter and Rubinfeld's insightful article on the discovery process[1] as the point of departure for analyzing two basic issues about this central feature of civil litigation. First, what are the effects of enabling litigants to compel adversaries to turn over evidence in their possession before trial? And second, how much discovery should be allowed in a given case?

Section I examines the first, positive issue. My focus is on discovery's informational effects.[2] The discovery process

---

* Assistant Professor of Law, Harvard University. I thank Louis Kaplow, David Shapiro, Steven Shavell, and Jennifer Zacks for comments.

[1] Robert Cooter and Daniel Rubinfeld, An Economic Model of Legal Discovery, 23 J. Legal Stud. ___ (1994).

[2] Discovery has effects other than providing the parties with information; a notable one is imposing costs on one or both parties. As I note in section II, these two effects may collide. Rather than comply with costly discovery requests from the plaintiff (for example), the defendant may buy its way out of them by making an acceptable settlement offer. In such cases, the effect of allowing discovery is not to provide the plaintiff with information but to line her pockets on a basis unrelated to the merit of her claim.

My main purpose in section I is to qualitatively assess discovery's informational effects, which have not been extensively explored. For that purpose, I proceed on the assumption that the parties actually undergo discovery rather than settling without it. I make no attempt to appraise the relative magnitude of discovery's informational effects and the settlement effects just described, though that is an important issue.

provides litigants with information about their case. But what kind of information exactly? In Cooter and Rubinfeld's model, discovery give the parties information about the probable outcome of trial. The defendant (say), but not the plaintiff, knows the contents of certain evidence that will emerge at trial. If the evidence is helpful to the plaintiff's case, the defendant prefers not to disclose it before trial, since to do so will inevitably strengthen the plaintiff's settlement position. (Without disclosure, the plaintiff's estimate of her chances at trial, and hence her settlement demands, are tempered by the possibility that the evidence in question will be against her. With disclosure, the plaintiff will realize her chances are better than she thought, and will raise her settlement demands accordingly.) Discovery rules, properly invoked and enforced, force the defendant to turn over the evidence; the effect, say Cooter and Rubinfeld, is to give the plaintiff information she would not otherwise receive about the likely outcome of trial.

But are discovery rules really needed to give the plaintiff this information? Even if there were no such rules, the defendant might not be able to keep the evidence secret. The plaintiff knows that if the evidence were favorable to the defendant, the defendant would eagerly disclose the evidence. Hence if the defendant conceals the evidence, the plaintiff will conclude the evidence is favorable to her side and raise her settlement demands. Knowing the plaintiff will do this, the defendant will want to disclose the evidence, even if it is

favorable to the plaintiff -- so long as it is less favorable than the plaintiff would conclude in the event of concealment. The threat of adverse inference by the plaintiff thus induces the defendant to disclose evidence he would prefer to keep secret. As a result, one way or the other -- directly if the defendant discloses, or by inference if the defendant conceals -- the plaintiff may learn the strength of the evidence in the defendant's possession, all without the aid of discovery rules.

There are, to be sure, limits on the plaintiff's ability to draw, and act on, inferences of this sort. Perhaps the evidence has surprise value at trial, tempting the defendant to hide even evidence favorable to his position; if so, the fact of concealment tells the plaintiff little about the contents of the evidence. Surprise value aside, perhaps the defendant will conceal evidence (if it is allowed) because he knows that the costs of litigation will discourage the plaintiff from raising her settlement demands at the risk of going to trial. For these reasons and others I will examine, the combination of voluntary disclosure and adverse inference will not eliminate all information asymmetries regarding the likely outcome of trial. Formal discovery rules must play some role in the exchange of such information. But it is unclear how substantial that role is.

Discovery rules have another, no less plausible, informational effect. They compel the disclosure of evidence that would not -- were it not for the fact of pretrial disclosure

itself -- emerge at trial. The defendant may, I have already suggested, have an incentive to disclose even some evidence unfavorable to the defense, if that evidence is expected to come out at trial. But there is no comparable incentive to disclose unfavorable evidence that will otherwise remain unrevealed at trial. Voluntarily turning over the latter evidence would pointlessly (as the defendant sees things) strengthen the plaintiff's case. Discovery rules force the disclosure of this evidence, thereby giving the court access to information it would not otherwise have in deciding the case. The result, presumably, is more accurate decisionmaking by courts.

Section II turns to the problem of identifying the optimal scope of discovery, assuming it enhances adjudicatory accuracy in the manner explored in the previous section. I propose a simple model for assessing the discovery's value at the margin, where the social objective is to minimize the costs of harmful behavior and its prevention and detection. In the model, the social desirability of discovery depends on whether it induces, at acceptable cost, defendant precautions against harm.

I use the model to make two points. One is that Cooter and Rubinfeld's proposed standard of an "abusive" discovery request -- which would compare the request's value to the requesting party to the costs of processing it -- probably does not yield the socially desirable level of discovery. Drawing on Steven Shavell's analysis of the divergence between social and private incentives to sue, I argue that there is no necessary correlation

between discovery's value to the parties and its value to society. The other point concerns the sheer difficulty of designing an optimal judicial standard for regulating the scope of discovery. Quite apart from the challenge of identifying the ideal amount of discovery, which turns on such imponderables as its deterrent value at the margin, there are (I suggest) some second-best problems introduced by the possibility that the parties may settle without undergoing discovery.

A couple of assumptions should be made clear at the outset. The first, positive, assumption is that litigants act exclusively in their rational self-interests. In the present context, that premise has some discomfiting implications; it suggests, for example, that litigants are prepared to conceal, and even lie about, unfavorable information if doing so is to their advantage. That may or may not faithfully describe actual litigant motivations; but since my purpose is to speculate about discovery's disclosure-forcing effects, it is an analytically useful assumption. The second, normative, assumption is that discovery rules should be designed to minimize social costs. That aim may sometimes conflict with the goal of the laws that discovery is used to enforce;[3] how that conflict should be resolved is an issue I do not address here.

---

[3] Indeed, the example I employ in section I involves the antidiscrimination laws, which are usually justified on grounds other than efficiency.

I. Discovery's Informational Effects

I propose distinguishing between two potential effects of discovery rules. One is to inform the parties, in advance of trial, the evidence that will be presented at trial. The other is to give the parties, and thus the court, access to evidence that otherwise -- but for the fact of its pretrial disclosure -- would not be presented at trial. The following hypothetical case illustrates the distinction.

An employee of a large firm brings a sex discrimination suit against her employer after she is denied a promotion.[4] Assume that the central issue in the case is whether the defendant witheld the promotion because of the plaintiff's sex. Assume further that the defendant, but not the plaintiff, has in its possession essentially all of the evidence bearing on the central issue in the case.

Some of the evidence in the defendant's possession will come out at trial even if there is no pretrial disclosure. An example might be the firm's annual written performance evaluations of the plaintiff and of employees who were given promotions. The plaintiff, we may assume, will subpoena these documents at trial even if she has never seen them. Before trial, only the defendant knows what they contain. A possible effect of discovery is to inform the plaintiff of the contents of the performance evaluations, so that she knows what the evidence will

---

[4] This hypothetical is loosely based on Blank v. Sullivan & Cromwell, 16 Fair Empl. Prac. Cas. (BNA) 87 (SDNY 1976).

show at trial.

But other evidence in the defendant's possession may not come out at trial unless there has been some sort of pretrial disclosure. An example might be certain internal memoranda or minutes of meetings regarding the firm's promotion decisions -- evidence that, if brought to the court's attention, would probably affect the likely outcome of the case. Absent pretrial disclosure the plaintiff, we shall assume for purposes of analysis, would not know of the existence of these documents and therefore would not subpoena them at trial. A possible effect of discovery is to give the plaintiff access to such evidence -- to evidence, that is, that would not be presented at trial <u>in the absence of its disclosure before trial</u>. This second effect is not to <u>inform the plaintiff</u> whether certain evidence will emerge at trial, but to <u>determine</u> whether it emerges.

In what follows I speculate on the relative likelihood of these effects given the parties' structural incentives to disclose or withold particular types of information in litigation. I make no claims about which effect predominates in practice, a question that turns on a number of empirical issues. My analysis suggests, however, that Cooter and Rubinfeld are mistaken in giving exclusive focus to the first effect; the second may be the more important one. I discuss the two effects in turn.

A. Disclosure Of Evidence That Will Emerge At Trial

Suppose that there are no rules of discovery; the courts will not compel disclosure of evidence before trial. Suppose also that there is some evidence -- the performance evaluations of other employees, in our hypothetical case -- that will (even though there is no discovery) emerge at trial. The plaintiff, let us say, intends to subpoena the performance evaluations at trial. If the defendant is made aware of this intention, the defendant may well have an incentive to disclose the evaluations' contents voluntarily before trial.

The incentive to disclose derives from the fact that it is costly to go to trial. As between going to trial and settling for the expected value of the judgment, both parties would prefer to settle and save the costs of trial.[5] Since the defendant knows what the evaluations say while the plaintiff does not, the parties' estimates of the expected value of the judgment may diverge. As Cooter and Rubinfeld observe, the defendant has an incentive to correct the plaintiff's "false optimism" about the outcome of trial. If the plaintiff thinks the performance evaluations will be more damning to the defendant than they in fact are, the defendant has an incentive to show the plaintiff her error. In my view, however, Cooter and Rubinfeld give insufficient weight to the possibility that this incentive may -- through a familiar unravelling process -- induce the defendant to

---

[5] Throughout, I assume the parties are risk neutral.

disclose the evaluations even if they are very damning indeed.[6]

Disclosure through unravelling comes about in this way. Assume that the performance evaluations are either damaging in some measure to the defendant or have no effect on the case. In no event do the evaluations actually help the defendant in the sense of negating liability.[7] Suppose that the defendant, by virtue of the evaluations' contents, is one of eleven "types." The types can be labelled with the integers ranging from 0 to 10. Type 0 means the evaluations do nothing to support the plaintiff's claim of discrimination. Type 10 means the evaluations' contents are highly unfavorable to the defendant; they are strong evidence of discrimination.

These "types" can be thought of as corresponding to the probability that the defendant will be held liable as a result of the evidence being introduced at trial; multiplying that

---

[6] Cooter and Rubinfeld acknowledge the possibility of unravelling but point to impediments, which I examine below. See Cooter and Rubinfeld, supra note 1, at [11-12]; text around note 15 infra. Unravelling in economic transactions outside the litigation context is described in Sanford Grossman, The Informational Role of Warranties and Private Disclosure about Product Quality, 24 J L & Econ 461 (1981), and Paul R. Milgrom, Good News and Bad News: Representation Theorems and Applications, 12 Bell J Econ 380-91 (1981). Steven Shavell, Sharing of Information Prior to Settlement or Litigation, 20 Rand J. Econ. 183 (1989), and Douglas G. Baird, Robert H. Gertner and Randal C. Picker, Strategic Behavior and the Law (preliminary draft, April 1, 1993), both discuss the unravelling model as it applies to settlement negotiations, but give less attention than I do to the plausibility of the model's assumptions in this context.

[7] In other words, they do not constitute evidence that the defendant itself would want to introduce at trial. My purpose in making this assumption is to focus on the case in which voluntary disclosure seems least likely.

probability by the plaintiff's damages yields the expected value of the judgment. Thus, suppose the plaintiff's damages are $100,000. If the defendant is type 0, the expected judgment against it is 0; if it is a type 5, the expected judgment is $50,000; if it is a type 10, the expected judgment is $100,000.[8]

The defendant knows its own type, but the plaintiff does not; the plaintiff knows only the distribution of types, which for simplicity's sake is assumed to be uniform. If the plaintiff has no information about the contents of the performance evaluations, her estimate of the expected judgment will be the mean of all possible defendant types, or $50,000. But then, if the defendant's true expected liability is lower than the mean, it has an incentive to reveal this to the plaintiff. If, for example, the defendant is a type 4, making the expected judgment $40,000, it will want to disclose the evaluations to the plaintiff. If it does so, the plaintiff will revise her estimate downward, and be willing to settle for less than she would have in the absence of disclosure. Similarly, if the defendant is type 0, 1, 2, or 3, it will want to reveal this to the plaintiff.

Thus, if the defendant does not disclose the evaluations, the plaintiff will infer that the defendant must be among the types 5 through 10. Her revised estimate of the expected

---

[8] It is unrealistic to assume that a defendant could have a 0% chance of being held liable, or that it could have a 100% chance of being held liable. The assumption is made purely for clarity's sake. The analysis would be qualitatively unchanged if we instead assumed that the probability of being held liable always fell somewhere between these figures.

judgment will be the mean for types 5 through 10, or $75,000. But then, by the above reasoning, if the defendant is type 5, 6, or 7, it will want to disclose its evidence to the plaintiff. Thus, if the defendant does not disclose, the plaintiff will infer that the defendant is a type 8, 9, or 10, and that the expected judgment is $90,000. But then if the defendant is a type 8, it will want to reveal this fact. If the defendant does not disclose, the plaintiff will infer the defendant is a type 9 or 10; then, if the defendant is a type 9, it will disclose. The upshot is that the absence of formal discovery rules is immaterial. The plaintiff learns as much from voluntary disclosure (and adverse inference from silence) as she would get if she could compel disclosure of the evidence.

This result seems to contradict the common-sense notion that the defendant would want to disclose the evidence only if it were favorable to the defendant's case, and instead to conceal it if it were unfavorable. But there is no contradiction. Favorable and unfavorable are relative terms. The evidence of a type 8 defendant (say) is unfavorable compared to type 3 but favorable compared to type 9. A type 8 would be happy to conceal its evidence if it could thereby be pooled with lower types who have more favorable evidence. But lower types will not willingly be pooled with a type 8; they would rather disclose their evidence. As a result, if the type 8 conceals, it will be pooled only with higher types. Compared to the other types in the pool, the type

8 has relatively favorable evidence.[9] And so the defendant discloses, damning though its evidence may be. Concealment would be even more damning.

To be sure, it is wrong to suppose that there would be complete unravelling in the manner I have described. The costs of litigation -- which help propel the unravelling dynamic by making settlement attractive to the defendant -- also have the effect of limiting the extent of voluntary disclosure. The reason is this: Assume that the defendant has not disclosed, and that the plaintiff must make her final settlement demand before going to trial. She could simply assume that the defendant is a type 10, and make a demand that only a type 10 would accept. But if she does so and is wrong about the defendant's type, the defendant will refuse her demand, forcing the plaintiff to go to trial. Given this risk, the plaintiff's optimal strategy is to scale back her demands -- in effect to assume that the defendant could be any one of several "high types" (say, 8, 9, or 10), and to offer settlement terms attractive enough that all of those types will accept.[10] Anticipating this result, the type 8 (and a fortiori, the type 9) defendant may have no reason to disclose if not compelled by law to do so. There would be at least some

---

[9] Another way of putting this is to say that if the type 8 defendant conceals, the plaintiff will be "falsely optimistic" that it confronts a type 9 or 10. The defendant will want to set the plaintiff right.

[10] For a formal demonstration, see Shavell, supra note 6, at 187-88. The extent to which defendants will be pooled in this fashion depends, among other things, on the size of the stakes and of the parties' litigation costs.

pooling of defendant types in the absence of discovery.

Nonetheless, the incentive structure suggests that there are strong tendencies toward voluntary disclosure of information about trial. Silence is itself costly to the defendant if it signals to the plaintiff that her prospects in court are better than she thought. The costs of silence may often be enough to induce disclosure. If so, discovery rules have no effect here; they simply require what the defendant would do anyway.

But will the plaintiff treat silence as a signal about the strength of her case? Since everything rides on the assumption that she will, I want to examine the assumption with some care. There are three main reasons that it might be false. One is that the defendant might not want to correct false optimism on the plaintiff's part. Another is that the plaintiff might not know of the evidence in the defendant's possession. A third is that the defendant might not be truthful when "disclosing" the evidence in its possession. Let us look at each.

Incentive to Disclose Favorable Evidence. The above analysis is driven by the defendant's presumed incentive to reveal (relatively) favorable information in order to buttress its settlement position. But perhaps the defendant would prefer to conceal favorable evidence, in the hopes of surprising the plaintiff with it at trial (thereby giving her less opportunity to rebut the evidence, adapt her trial strategy to it, and so

forth).[11] Perhaps, in other words, the defendant wants the plaintiff to be falsely optimistic about her chances of winning at trial. This is a potent objection to my analysis. If defendants prefer concealing favorable information to disclosing it, then all bets are off; discovery rules are necessary to induce disclosure.

Still, it is a fair bet that the defendant will frequently like revealing favorable information better than concealing it, surprise value notwithstanding. To pursue the strategy of surprising the plaintiff, the defendant probably must be willing to give up the possibility of settling -- meaning the strategy carries with it the cost of going to trial.[12] The defendant will

---

[11] See, for example, Richard A. Posner, An Economic Approach to Legal Procedure and Judicial Administration, 2 J. Legal. Stud. 399, 422 (1973); Shavell, supra note 6, at 193-94.

[12] Suppose that if the defendant reveals its evidence, the expected judgment will be $20,000, and this will be known to both parties. On the other hand, suppose that if the defendant conceals its evidence, the expected judgment will be $10,000, but this will be known only to the defendant; the plaintiff's estimate of the expected judgment will be the mean of all defendant types, or $50,000. (The plaintiff will adopt this mean estimate because she can infer nothing from the defendant's silence; silence might mean the evidence is favorable and the defendant wants to keep it for a surprise at trial, or that it is unfavorable and the defendant is hoping to settle before the weakness of its case is revealed.) If the defendant discloses, settlement should be feasible, since the parties will agree on the expected value of the judgment. If the defendant does not disclose, settlement will be difficult, since there is a large gap between the parties' estimates.

A formal representation of the matter, on which I will draw at a few points in this paper, is as follows. Let l be the plaintiff's losses (and expected damage award); let c_P and c_D be the parties' respective costs of going to trial; and let π_P and π_D be the parties' respective estimates of the plaintiff's chances of success at trial. The precondition for settlement is (assuming the parties are risk-neutral) that the defendant's expected loss from

only pursue the strategy if the anticipated return to the claim generated by surprise at trial exceeds the defendant's share of the surplus that would be generated by settling.[13] That condition may hold in certain cases where the stakes are large

---

going to trial exceed the plaintiff's expected gain; that is, $\pi_D l+c_D > \pi_P l-c_P$. Rearranging terms, we have as the precondition for settlement

$$c_P+c_D > (\pi_P-\pi_D)l.$$

(Compare Richard A. Posner, Economic Analysis of Law 556 (4th ed. 1992). An additional precondition for settlement, which I put to one side here, may be that the plaintiff's expected recovery at trial exceed her costs of litigation. See generally Lucian Bebchuk, The Credibility and Success of Suits Known to Be Made Solely to Extract a Settlement Offer (unpublished manuscript, November 1991).)

In the situation discussed in the text, the plaintiff's chances of prevailing at trial depend on whether the evidence is disclosed before trial. Let π^* be the plaintiff's chances of prevailing if the evidence is disclosed; its value is known to both parties once the evidence is disclosed. Let $(\pi^*-\Delta\pi)$ be the plaintiff's chances of prevailing if the evidence is not disclosed; its value is known only to the defendant. If the defendant discloses the evidence, the case will settle; since the parties' expectations about trial are identical, the precondition for settlement is always satisfied. If on the other hand the defendant does not disclose the evidence, the case will only settle if

$$c_P+c_D > (\pi_P-\pi^*)l+\Delta\pi l.$$

As the value of $\Delta\pi l$ -- which is in essence the "surprise value" of keeping the evidence secret until trial -- increases, the likelihood of satisfying the precondition for settlement declines.

[13] If the parties split the surplus (i.e. cost savings) from settlement, then upon the defendant's disclosing the evidence the case settles for $\pi^*l+\frac{1}{2}(c_D-c_P)$. If the defendant does not disclose and the case fails to settle, the defendant's anticipated loss from litigation will be $(\pi^*-\Delta\pi)l+c_D$. Comparing these figures, the defendant will only conceal the evidence if

$$\Delta\pi l > \tfrac{1}{2}(c_D+c_P).$$

relative to the costs of going to trial; it is less likely to be met otherwise.[14] How often the condition holds is an empirical question. But it seems safe to speculate that in a broad range of cases, the defendant would prefer the settlement value of disclosure to the surprise value of concealment.

Parties' Knowledge. A second objection, this one raised by Cooter and Rubinfeld, is that the plaintiff may not know what evidence the defendant has in its possession.[15] Unravelling won't occur unless the defendant expects the plaintiff to draw an adverse inference from the defendant's decision to conceal information; to draw such an inference, the plaintiff needs to realize that the defendant is concealing information. Perhaps, say Cooter and Rubinfeld, the plaintiff will be unaware that the defendant has certain evidence that will come out at trial. If so, she will infer nothing from its concealment, and the defendant may accordingly choose not to reveal it.

This argument is hard to square with the premise that the evidence will come out at trial. Recall that we are focusing on evidence that will, even in the absence of discovery, be brought out at trial. By hypothesis, the evidence is damaging (or at best neutral) to the defendant's position at trial. If the

---

[14] The defendant will only conceal the evidence if (rearranging terms in the inequality in the previous footnote) $2\Delta\pi > (c_D+c_P)/l$, which is less likely as the larger the parties' costs are relative to l.

[15] See Cooter and Rubinfeld, supra note 1, at [10].

evidence is damaging to the defendant, who will introduce it at trial? Presumably the plaintiff, since the defendant would just as soon keep the evidence out. Yet to introduce the evidence at trial, the plaintiff must know (more precisely, believe) that the defendant has it. And if she has this belief at the time of trial, she can normally be expected to have it before trial as well. The assumption that the evidence will come out at trial thus probably implies that the plaintiff is aware of its existence in advance of trial.

It is tempting to reply that some evidence will not come to the plaintiff's attention until trial, as a result of the revelation of other evidence. Perhaps, for example, the plaintiff only learns of the existence of the performance evaluations from the testimony of one of the defendant's managers. But this possibility just pushes the analysis one step back. If the plaintiff expects the defendant's manager to testify at trial, then the defendant has an incentive to reveal the contents of the manager's testimony; otherwise the plaintiff will draw the most adverse possible inference from the defendant's silence. And if the manager's testimony is disclosed before trial, the plaintiff will learn of the performance evaluations at that time. And once she learns of them, the defendant will be induced to reveal their contents. In short, it makes no difference that the plaintiff's awareness of evidence *X* depends on the revelation of evidence *Y*; if unravelling leads the defendant to disclose *Y*, it will also lead to the disclosure of

X.[16]

A somewhat stronger variant of this second objection concerns the defendant's knowledge. Perhaps (it might be argued) the defendant will not realize the plaintiff is aware of the performance evaluations and that she intends to subpoena them at trial. If that is the case, the defendant will not expect the plaintiff to draw adverse inferences from the concealment of the evidence; and so the defendant may decide to conceal it.

This argument overlooks the plaintiff's incentive to reveal her intentions to the defendant. If she believes the evidence exists and plans to introduce it at trial, it is likely in her interest to inform the defendant of this before trial. To see why, suppose the defendant is (wrongly) convinced that the plaintiff will not subpoena the performance evaluations. Then the defendant will in effect consider itself a "type 0" as that term is used above: the evaluations can do it no harm at trial, because they will not be introduced. The plaintiff, on the other hand, will expect (absent further information) the defendant to be the average of all possible types, which is type 5. The disparity in the parties' estimates of the expected judgment will be $50,000. Nonetheless, no voluntary disclosure will occur, because the defendant erroneously thinks the plaintiff is unaware

---

[16] The plaintiff's awareness of the existence of evidence *Y* may itself depend on the revelation of other evidence *Z*. If unravelling would induce the defendant to disclose *Z*, it would lead to the disclosure of *Y* and hence to the disclosure of *X*. In principle evidence *X* might be many times removed from the evidence of which the plaintiff is aware, yet nonetheless come to the plaintiff's attention before trial.

of the evaluations' existence and won't introduce them at trial.

This state of affairs is unacceptable for the plaintiff. She would be better off informing defendant of her intention to subpoena the evidence at trial. If she doesn't inform the defendant of her intention, the case will be difficult to settle, given the divergence in the parties' expectations about trial.[17] Moreover, if the case does settle, the plaintiff will have to accept an amount reflecting the defendant's belief that it is a type zero. In effect, the settlement will treat the defendant as if it were a low type -- even if the defendant is in fact a high type. If instead the plaintiff does inform the defendant of her intention to subpoena the evidence, then the unravelling process is set in motion; the defendant discloses the evidence. The case becomes easier to settle, since the parties' expectations about trial are brought into line. In addition, the plaintiff recovers more in settlement, since the settlement amount is now determined by the defendant's true type.[18] It is thus unlikely that the

---

[17] A related reason the case may fail to settle is that the defendant may believe that the case is unprofitable for the plaintiff to take to trial, and therefore that the plaintiff will drop the case in the event no settlement is reached. Whether the defendant has this belief depends on whether it thinks the plaintiff has evidence (other than the performance evaluations) to support her claim.

[18] The plaintiff's incentive to reveal her intentions can be shown as follows, using the notation introduced in note 12 above.

1. Plaintiff does not disclose. Suppose first that the plaintiff does not reveal her intentions. The plaintiff's estimate of her chances of winning (π_P) is .5, since the defendant has not disclosed its type and the types are assumed to be uniformly distributed. The defendant's estimate (π_D), assuming there is no

defendant would remain ignorant of the plaintiff's intention to introduce the evidence; the plaintiff has a strong incentive to reveal her intention.

There may, however, be counterincentives. One is of a sort we have seen already: the plaintiff might do better at trial if she conceals her intentions beforehand and then surprises the defendant with an unexpected subpoena.

Another possibility is that concealing her intentions may

---

other evidence available to the plaintiff bearing on the defendant's liability is 0. The plaintiff's minimum settlement demand is $.5l-c_P$; the defendant's maximum offer is c_D. (This puts aside the fact that the defendant might well offer nothing because it believes the plaintiff lacks a credible threat to go to trial. My conclusion would merely be strengthened if we took that into account.)

The parties will settle if $c_D > .5l-c_P$. Assume that if the case settles, the settlement amount is (on average) at the midpoint of the settlement range. Then the plaintiff's expected settlement recovery is $\frac{1}{2}(.5l-c_P+c_D)$, or $.25l+\frac{1}{2}(c_D-c_P)$. If the parties fail to settle, the plaintiff's expected gain from trial is $.5l-c_P$.

2. Plaintiff discloses. Now suppose the plaintiff reveals her intentions. This will induce the defendant to disclose the evidence. The parties then both know the plaintiff's true chances of success (π^*) and the case settles for $\pi^*l + \frac{1}{2}(c_D-c_P)$. But as we have seen, before revealing her intentions, the plaintiff estimates the value of π^* to be .5. Thus, the amount the plaintiff expects to recover in settlement if she discloses her intentions is $.5l+\frac{1}{2}(c_D-c_P)$.

That amount exceeds the amount she expects to get (whether through trial or settlement) from concealing her intentions. That is, $.5l+\frac{1}{2}(c_D-c_P)$ is greater than $.5l-c_P$ and is greater than $.25l+\frac{1}{2}(c_D-c_P)$, for all positive l, c_P and c_D. Thus, in expected value terms, the plaintiff is always better off revealing her intentions.

One complicating factor here is that the plaintiff may sometimes not be able to avoid revealing her intentions. The amount she demands in settlement may signal to the defendant that she is aware of the evidence and intends to introduce it. Unravelling may then occur. But this complication does not undermine the basic point; given the choice between concealing and revealing her intentions, she is better off revealing.

induce her opponent to divulge information that it would not otherwise reveal. Suppose, for example, that the defendant has some damaging document *W* in its possession, and that the defendant is unsure whether the plaintiff is aware of its existence. If the plaintiff announces her intention to introduce documents *X*, *Y*, and *Z* (but no others), the defendant will be tempted to conclude that she is unaware of *W* and that she will draw no adverse inferences if the document is not disclosed. If, on the other hand, the plaintiff says nothing of her intentions, the defendant will be uncertain whether she will draw adverse inferences from the document's concealment; and this uncertainty may induce disclosure, even if the plaintiff is in fact unaware of the document. For this reason the plaintiff may find it to her advantage to keep her intentions to herself. But the strategy may backfire; the defendant may conclude it doesn't need to disclose <u>any</u> evidence. The plaintiff may then want to disclose her intentions after all.

To the extent the plaintiff does find it in her interest to reveal her intentions, a system of voluntary disclosure would in the end closely resemble the process of formal discovery. The plaintiff's lawyer would say to the defendant's lawyer: "I believe your client has in its possession some performance evaluations. I intend to subpoena these at trial, so you might as well show them to me now." If the defendant believes that the

plaintiff will in fact subpoena the documents,[19] and that the court will allow their introduction into evidence, it will turn them over. Disclosure of the documents may, depending on their contents, prompt another round of inquiries by the plaintiff. In this way, all of the evidence that will come out at trial would in principle be disclosed in advance.

Truthful Disclosure. But all of this assumes that the defendant is truthful when it "discloses" its evidence to the plaintiff. A third objection to my analysis is that the defendant might lie. A high type might masquerade as a low type by doctoring the performance evaluations; by falsely announcing that no such evaluations exist; or by turning over only the least damaging evaluations, and falsely claiming that there are no others. If the defendant can get away with such tactics, then self-interested defendants, regardless of their true type, will claim to be low types. Assuming the plaintiff cannot tell whether the defendant is lying, the defendant's "voluntary disclosure" will give her no information.

This is a problem. Unravelling models usually specify that the information recipient has some way of verifying the information she receives -- if not at the time of receipt, then after the fact.[20] (The car buyer can verify the seller's claims

---

[19] The credibility of the plaintiff's threat depends on the subpoena's costs to the plaintiff and/or her attorney. These include lawyer time and (perhaps) lost opportunities to pursue other lines of inquiry at trial. In some instances these costs might render the plaintiff's subpoena threat incredible.

[20] See, for example, Grossman, supra note 6, at 461.

about gas mileage by test driving it before the purchase; alternatively, she can test drive it after purchase and rescind the sale if the seller's claims are false.) Verification is not so easy in the present context. Before trial, the defendant turns over some documents to the plaintiff,[21] and says that these are (all of) the very documents the plaintiff will get if she subpoenas them at trial. But the whole point of this pretrial exercise is to avoid trial by settling. If the exercise succeeds, the plaintiff will never know whether the defendant was telling the truth about the evidence that would emerge at trial. The plaintiff has nothing against which to test the defendant's purported disclosures. Realizing this, the defendant will be tempted to make fraudulent disclosures. Realizing that, the plaintiff will not credit the information she receives.

But none of this undermines my argument. The temptation to lie does not merely hinder voluntary exchanges of information between the parties. It also hinders compulsory exchanges. Even if the law requires the defendant to satisfy the plaintiff's requests for evidence in its possession, the defendant will be tempted to falsify or deny the existence of damaging evidence. It is thus unclear that discovery rules contribute anything to the exchange of information about trial. The same problems of verification that impede the voluntary disclosure of reliable information about trial will impede its compulsory disclosure.

Yet discovery rules do enable courts to punish false or

---

[21] Or, alternatively, says there are no documents.

incomplete disclosures with sanctions such as attorney fee awards, contempt citations, and default judgments.[22] It might be argued that this arsenal of sanctions is what discovery rules contribute to the exchange of trial information: the threat of penalties discourages parties and their attorneys from making fraudulent disclosures, thereby (the argument would run) making the exchange of information more reliable than it would be in a system of purely voluntary disclosure. This argument has some surface appeal. The difficulty, however, is there are severe sanctions for fraudulent disclosure quite apart from those provided by discovery rules. A litigant or lawyer who doctors evidence before disclosing it or falsely denies its existence would expose himself to possible civil or criminal liability; a lawyer would also expose himself to disbarment or other professional discipline. The threat of these penalties would presumably discourage fraudulent behavior in a system of purely voluntary disclosure. (To bring these deterrent devices into play, the plaintiff could ask the defendant and its lawyer to certify that its disclosure was complete and truthful. If the defendant or the lawyer were unwilling to do so, she would draw and adverse inference about their veracity. The certification requirements in discovery rules can be duplicated by private

---

[22] See, for example, FRCP 37; Cal Civ Proc Code § 2023.

agreement.)[23] Perhaps discovery rules add to the deterrence by piling on further sanctions; perhaps in some cases the additional deterrence means the difference between fraudulent and truthful disclosure.[24] But it is unclear how often discovery rules make a difference here.

B. Disclosure of Evidence That Would Not Otherwise Emerge At Trial

To this point we have focused on the pretrial disclosure of evidence that would -- independently of discovery -- come out at trial. I have argued that discovery rules may often have comparatively little impact on the pretrial disclosure of such

---

[23] This analysis is, I think, a sufficient response to the argument made in Shavell, supra note 6, at 191, that discovery rules are sometimes needed to induce disclosure when some informed parties are unable to credibly establish their type. In our example, assume that sometimes the defendant is unable to disclose the performance evaluations (perhaps because it is still searching for them). The plaintiff has no direct way of knowing when this is the case. Accordingly, a high type that is able to disclose its evaluations may be tempted to claim (falsely) that it is unable to disclose. With discovery rules, the defendant won't be able to tell such a lie; if it has the evidence at its disposal, it must turn it over. Therefore -- the argument concludes -- when a defendant claims to be unable to reveal its evaluations, the plaintiff will know the defendant is telling the truth. The difficulty with this argument is its assumption that discovery rules are necessary to lend credibility to the defendant's claims regarding disclosure. For reasons I have given, the assumption is doubtful.

[24] It might be that discovery rule sanctions are more effective deterrent devices than professional discipline or criminal penalties because they can involve private enforcement. The defrauded party (the plaintiff in our example) has an incentive to seek sanctions such as attorney fee awards or default judgments. On the other hand, such sanctions may be small potatoes for the defendant's attorney; the threat of disbarment may be the greater deterrent. The comparative effectiveness of different sanctions is an empirical question.

evidence; given that the evidence will come out at trial, it may well be disclosed whether or not the law so requires. But let us now assume that there is some evidence in the defendant's possession that would not emerge at trial unless it had previously been disclosed to the plaintiff. Earlier I suggested that certain memoranda or minutes of meetings could be examples.

Why might such evidence fail to come out at trial without pretrial disclosure? If the plaintiff were aware of the existence of particular memoranda or minutes, she could subpoena them.[25] But suppose that she doesn't know what documents exist; she simply believes that the defendant's files contain evidence to support her claim. Two constraints might prevent her from unearthing this evidence at trial. One is admissibility. The evidence the plaintiff is looking for is, by hypothesis, admissible. But it may be that the only way to locate that evidence is to examine or sift through other, inadmissible, items. Given admissibility restrictions, such sifting may not be possible at trial. The other reason is limited judicial resources. Admissibility problems aside, a court can't give the plaintiff the weeks or months (even years) of trial time it might take to go through the defendant's files in search of usable evidence.

---

[25] The plaintiff could simply try to subpoena at trial "all memoranda and minutes" (or better, "all documents") pertaining to the subject matter of the litigation. The assumption in the argument that follows, however, is that this will not (always) be able to do so. If the assumption seems unrealistic in our hypothetical discrimination case, the reader might substitute a complex antitrust or securities case.

Discovery rules free the plaintiff of both of these constraints. Admissibility restrictions don't apply to discovery requests; the plaintiff can examine even inadmissible materials if doing so may lead her to admissible evidence.[26] And since pretrial discovery takes place outside of court, the plaintiff can in principle go through reams and reams of the defendant's files without burning up judicial resources. Judges are forced to supervise discovery, sometimes intensively. But this is not the same as having the entire process of exchanging and examining documents take place at trial. Plaintiffs routinely can demand the production of far more documents in pretrial discovery than they could ever hope to subpoena at trial. Allowing discovery thus increases the likelihood that our hypothetical memoranda will come to the plaintiff's attention and be introduced at trial.

But is compulsory disclosure really necessary to bring the evidence to the plaintiff's attention? It might be asked whether a system of voluntary disclosure would not achieve the same result. Suppose once again that there are no rules of discovery. Absent pretrial disclosure, the plaintiff would not know of, and would not introduce at trial, the defendant's memoranda. Would the logic of unravelling induce the defendant to disclose the memoranda even without legal compulsion?

It would not. If the defendant realizes the plaintiff is unaware of the documents, it has no reason to call her attention

---

[26] See, for example, FRCP 26(b)(1).

to them. In part A above, what led the defendant to disclose the performance evaluations was the fact that the plaintiff's belief that they had a positive expected value at trial -- premise on her intention to introduce them as evidence -- influenced her settlement demands. Since performance evaluations had an expected value of $50,000, the plaintiff's initial settlement demand was set accordingly; this gave low types (those with evidence less damaging than average) an incentive to disclose, and set the unravelling process in motion. Yet none of this happens unless the plaintiff expects the evidence to be introduced at trial. If she is unaware of the memoranda, they are in effect assigned a value of 0 in her settlement demands. The defendant has no incentive to correct her beliefs; turning over the evidence would simply give the plaintiff ammunition for trial (hence for settlement) she would otherwise never get. Even if the memoranda are only slightly damaging evidence, the defendant is better off saying nothing.[27]

To be sure, the defendant may not know whether the plaintiff is aware of the evidence. An equilibrium may exist in which the plaintiff does not reveal her beliefs and in which the defendant, uncertain whether the plaintiff knows of their existence, turns over the memoranda.[28] But as I suggested earlier, there are

---

[27] Here I am maintaining the assumption, introduced in part A above, that the evidence in question is either damaging to the defendant or has no effect on the case. If the evidence were exonerating, the defendant would eagerly turn it over.

[28] See text preceding note 19 supra.

strong incentives for the plaintiff to announce in advance the evidence she intends to subpoena at trial, in order to ensure that the defendant discloses it before trial. If the memoranda are not on the list she sends over, the defendant will be inclined to conclude she is unaware of them. And in any event, there are probably some types of evidence the defendant cannot be bluffed into turning over. If finding a given document would require the plaintiff to ferret through stacks of (perhaps inadmissible) defendant files, the defendant might safely conclude that the document won't come out at trial -- and that the plaintiff doesn't expect it to. There would be no point in the defendant voluntarily turning over such a document.

Discovery rules, then, give the plaintiff access to evidence that would otherwise not come to light. Here I do not simply have in mind the "smoking gun" buried in the defendant's files, but also other less dramatic material. The evidence uncovered through discovery may not break open the litigation. But it must often enable the plaintiff to build a case that she could not have built from simply subpoenaing documents at trial. As a result, the trier of fact is confronted with evidence it never would have seen in the absence of discovery.

C. Consequences For Adjudication And Settlement

I do not know the relative magnitude of the two effects of discovery I have identified -- informing parties of the evidence that will (exogenously) emerge at trial versus bringing out

evidence that would not otherwise emerge. Since both effects are possible, both should be kept in mind when asking what impact discovery has on the outcome of litigation. In what follows I consider briefly the potential consequences of discovery on both the quality of trials and settlements and the relative likelihood of trials and settlements.[29]

Accuracy of Outcome. If discovery rules bring out evidence that would otherwise remain concealed at trial, it presumably adds to the accuracy of judicial decisions. Accuracy of decision depends on the quantity of information available to the court, which depends (in an adversary system) on the quantity available to the parties. By expanding the pool of evidence available to the plaintiff in our example, discovery expands the pool available to to the court. With discovery, the court gets to see both the performance evaluations and the internal memoranda; without discovery, it only gets to see (at most) the evaluations. Sometimes the evidence of wrongdoing (when it has occurred) will be in the memoranda, not the evaluations. With discovery, the court catches some violations it would otherwise miss. The prospect of greater accuracy in trial outcomes, moreover, can lead to greater accuracy in settlements.[30] Once the memoranda

---

[29] Cooter and Rubinfeld also explore the impact of discovery on the filing of claims and the costs of litigation, subjects which in the interest of brevity I will not address.

[30] It does so only on the assumption that discovery in fact occurs. The anticipated costs of discovery may induce the parties to settle before discovery. It is possible that the terms of such a settlement will be worse -- will less accurately reflect the merits of the case -- than the settlement the parties would reach

are disclosed, the plaintiff can demand in settlement what they are worth at trial. If a trial with discovery is more accurate than a trial without it, then a settlement reflecting the anticipated outcome of a trial with discovery is likely more accurate than a settlement reflecting the anticipated outcome of a trial without it.

This connection between discovery rules and accuracy of outcome is different from (though consistent with) the one identified by Cooter and Rubinfeld. In their view, discovery improves the quality of trials and settlements by telling the plaintiff what evidence will later come out at trial. Trials improve because the parties are better prepared when they appear in court.[31] Settlements improve[32] because both parties will be informed about the evidence, meaning that the settlement will reflect the "merits" (the probable outcome of trial).[33] My analysis suggests, however, that there is more to the story. Discovery adds to the accuracy of outcomes not only by eliminating surprises about the evidence but also by expanding the pool of evidence available to the parties.

---

if there were no discovery rules. I look at this problem below in section IIC.

[31] See Cooter and Rubinfeld, supra note 1, at [12]. Cooter and Rubinfeld rightly note that party preparation doesn't always advance accurate decisionmaking; sometimes it may be better to surprise a party at trial, so that he lacks the chance to dilute or obfuscate the evidence against him. See id. at [15].

[32] Assuming settlement follows discovery.

[33] See Cooter and Rubinfeld, supra note 1, at 14.

Settlement vs. Trial. I adopt the conventional premise that the parties settle if and only if the defendant's (subjective) expected loss from trial exceeds the plaintiff's expected gain.[34] In the present context, the probability of that occurrence depends in large measure on the parties' beliefs about the evidence in the defendant's possession. I suggest that discovery's impact, if any, on the likelihood of settlement depends on what the defendant (as the informed party) knows about the plaintiff's beliefs -- an issue Cooter and Rubinfeld do not consider explicitly.

1. Assume first that the defendant knows the plaintiff's beliefs at all times, enabling the defendant to predict the plaintiff's settlement demands following disclosure or concealment of the evidence. Granted this assumption, discovery rules either have no effect on the likelihood of settlement, or reduce it. Perhaps surprisingly, discovery rules never make settlement more likely.[35] On this point I part company with Cooter and Rubinfeld, who argue that discovery rules can make settlement more likely by reducing party disagreement over the probable outcome of trial.[36] Let us examine two possible cases.

---

[34] See Cooter and Rubinfeld, supra note 1, at [6] (adopting the same assumption). The parties' gains and losses are interpreted here to be net of litigation costs.

[35] Discovery may enhance the likelihood of settlement by raising the costs of litigation, but I put this to one side. My focus is on discovery's informational effects.

[36] More precisely, proposition 1 of Cooter and Rubinfeld's paper states that discovery makes settlement more likely when it makes the recipient (the plaintiff, in our example) more

(a) In the simplest case, the parties interpret the evidence identically; that is, upon seeing it, they each form the same appraisal of its expected value at trial. In this case, discovery has no impact on the likelihood of settlement; the dispute will settle regardless of whether the defendant is required to disclose the evidence. This can be seen as follows.[37] Let s be the amount the plaintiff will demand to settle the case if the defendant has disclosed the evidence before trial.[38] Let $\hat{s}$ be the amount the plaintiff will demand if the defendant has instead concealed the evidence before trial.

By hypothesis, if the defendant discloses the evidence, the parties will settle because they agree on the expected result of trial.[39] This implies that the defendant must be willing to pay s to settle the case, which implies that s is less than the

---

pessimistic about trial. See Cooter and Rubinfeld, supra note 1, at [8]. But since the authors assume that there will be a settlement if the parties agree (or, a fortiori, if they are mutually pessimistic) about the outcome of trial, proposition 1 must refer to situations in which, before discovery, the parties are mutually optimistic about trial. In that situation, making the plaintiff more pessimistic is equivalent to bringing the parties closer to agreement about the probable outcome of trial.

[37] I am indebted here to Steven Shavell, who proves the point formally in Shavell, supra note 6, at 186. My analysis extends his argument to cases in which the parties do not necessarily interpret the evidence identically.

[38] More precisely, s is the minimum amount the plaintiff will accept to settle the case.

[39] This follows from the assumption that the parties settle whenever the defendant's expected loss (net of costs) from trial exceeds the plaintiff's expected gain. Since litigation is costly, this condition is satisfied if the parties agree on the expected outcome of trial -- as they will if they both see all the evidence and interpret it identically.

defendant's expected loss from going to trial. But then the case will definitely settle. If $\hat{s} > s$, the defendant will disclose and agree to the plaintiff's demand of s. If $\hat{s} < s$, the defendant will conceal and agree to the plaintiff's demand of $\hat{s}$.

Thus, granted the assumption that the parties interpret the evidence identically once they see it, discovery has no effect on the likelihood of settlement. If disclosure is required, the parties will settle for s. If disclosure is not required, the parties will settle for either s or $\hat{s}$, whichever is smaller. Discovery rules potentially affect the terms of settlement, but not its likelihood.

(*b*) In the next case, the parties don't necessarily evaluate the evidence identically once they see it. If the plaintiff sees the evidence, she may view it as more favorable to her case than does the defendant. As a result, settlement is no longer guaranteed following disclosure. The plaintiff, having seen the evidence, may demand more in settlement than the defendant expects to lose from going to trial -- meaning there will be a trial if disclosure occurs.

The upshot is that discovery may sometimes prevent settlement.[40] Suppose that s, but not $\hat{s}$, exceeds the defendant's (subjective) expected loss from trial. In effect, concealment is a prequisite to settlement. If disclosure is not required, the

---

[40] This analysis assumes that the defendant knows at all times the values of s, the amount the plaintiff will demand. The assumption may be unrealistic, given the premise that the parties interpret the evidence differently.

defendant will conceal the evidence and the case will settle for $\hat{s}$.[41] If disclosure is required, the defendant will disclose the evidence, the plaintiff will unsuccessfully demand s, and the case will go to trial. Requiring disclosure prevents the parties from settling.

On the other hand, requiring disclosure never helps bring about settlement (under the given assumptions). Suppose that $\hat{s}$, but not s, exceeds the defendant's expected loss from trial. Then disclosure is a prerequisite to settlement. But discovery isn't needed to induce disclosure. Whether or not disclosure is required, the defendant will disclose and the case will settle for s. Thus, when the parties interpret the evidence differently, discovery either reduces the likelihood of settlement or leaves it unaffected.[42]

2. So far we have assumed the defendant knew the plaintiff's beliefs and so could predict the value of s and $\hat{s}$ before deciding whether to disclose the evidence. When this is not the case, discovery may sometimes make settlement more likely. Suppose, for example, that the defendant wrongly thinks that s, but not $\hat{s}$, exceeds the defendant's expected loss from

---

[41] This follows from the preceding analysis, which showed the defendant pursuing the path (disclose or conceal) leading to the lowest possible settlement demand, which in this case is $\hat{s}$.

[42] When both s and $\hat{s}$ exceed the defendant's expected loss from trial, the case fails to settle regardless of whether disclosure occurs. When neither s nor $\hat{s}$ exceeds the defendant's expected loss from trial, the case settles regardless of whether disclosure occurs. In both of these instances -- as well as the instance discussed in the text -- requiring disclosure has no effect on whether the case settles.

trial, when in reality it is the other way around. If disclosure is not required, the defendant will conceal, thinking this will lead to the lower settlement demand; but the result will be a trial. If disclosure is required, upon disclosing the defendant (unexpectedly) will receive an acceptable demand, and the case will settle.[43]

II. The Optimal Scope of Discovery

Discovery is costly, and many of its costs are externalized by the requesting party.[44] Plaintiff discovery requests may force the defendant[45] to screen thousands of documents for privileged material or spend weeks in depositions; the evidentiary payoff to the plaintiff may be small in relation to the defendant's costs, but that is no concern to her. On top of that, there is a temptation for the plaintiff to use discovery strategically, as a means of imposing costs on the defendant in the hopes of extracting a settlement unrelated to the merit of her claim.[46] The potential for waste raises a fundamental issue:

---

[43] It is also possible, of course, that for other values of s and $\hat{s}$, discovery will reduce the likelihood of settlement or have no effect on it.

[44] The requesting party may of course spend a great deal of time and money processing the information that is turned over.

[45] I will continue to focus on discovery requests made by the plaintiff. Most of the following analysis also applies to requests made by the defendant.

[46] The defendant may respond with (threats of) burdensome discovery requests of its own, but that is merely a further instance of the problem. There is no reason to expect the parties' equilibrium strategies will result in the right amount of

Given its costs, what is the right amount of discovery in a case?

The problem is brought into sharp focus by Cooter and Rubinfeld's effort to formulate a definition of discovery abuse.[47] They propose a standard that focuses on discovery's net return to the requesting party. Under their standard, a discovery request by the plaintiff is acceptable if the costs it imposes on the parties are less than the anticipated return it brings to the plaintiff's claim at trial. Conversely, a request is "abusive" and should be stricken (or discouraged) if its costs to the parties exceed its anticipated return at trial.[48]

From a social welfare perspective, the trouble with this approach -- which, according to Cooter and Rubinfeld, reflects "current law and practice"[49] -- is that it rests on a purely ex post perspective on the litigation process. The rules of

---

discovery. The strategic use of discovery is explored in Frank H. Easterbrook, Discovery as Abuse, 69 B.U. L. Rev. 635 (1989).

[47] See Cooter and Rubinfeld, supra note 1, at [20].

[48] See Cooter and Rubinfeld, supra note 1, at [21].

[49] Cooter and Rubinfeld treat their proposed standard as an interpretation of the Federal Rules' prohibition of "unduly burdensome or expensive" discovery. See id. at [21] & n. 16 (quoting FRCP 26(b)(1)). They do not seek to defend their standard on social welfare grounds; rather, their purpose is to "clarif[y]" existing law. See id. at [22].

The purpose of my remarks in this section is to show that the standard they put forth does not necessarily advance social welfare. I take no position on whether their standard correctly mirrors existing law; if it does, my remarks can be taken as a criticism of that law. (Although my research has not uncovered a judicial opinion explicitly defining abusive or "unduly burdensome and expensive" discovery, my guess is that many district judges and magistrates would endorse something like Cooter and Rubinfeld's test.)

litigation don't just shift around the losses occasioned by past behavior (violations of the law, tortious conduct, and so forth).[50] They influence the likelihood that wrongful (or injurious) behavior, and litigation, will occur in the future. Any discovery standard concerned with minimizing social costs must take these ex ante effects into account.

A. Discovery's Social Value at the Margin

If discovery increases the accuracy of the legal process, it may thereby induce greater compliance with the law (or, equivalently, greater precautions against causing harm). Other things equal, increased accuracy means an increased likelihood that wrongdoers are held liable for their actions.[51] Ex ante, an increased likelihood of liability means greater compliance incentives. What follows is a model that compares the value of increased compliance (if it occurs) and the costs generated by

---

[50] If that were all litigation did, it is unclear how it -- or its constituent parts, including discovery -- could ever improve social welfare, if the latter term is defined in terms of wealth maximization. Why is giving the plaintiff a transfer payment with expected value $\$X$ worth charging the parties (up to) $\$X$ in discovery costs? The transfer itself does not produce wealth; assuming equal marginal utilities of wealth, the plaintiff's gain is offset by the defendant's loss.

[51] See generally Louis Kaplow, The Value of Accuracy in Adjudication, 23 J. Legal Stud. ___, ___ (1994). The deterrent effect of discovery has been noted by others (see, for example, Easterbrook, supra note 46, at 637 n.12), but it has not to my knowledge been explicitly modeled.

discovery.[52]

In the model, a potential injurer chooses between taking a specified precaution or taking no precaution (equivalently, complying with the law or violating it). Taking the precaution reduces the probability of injuring the plaintiff. If the plaintiff is injured by the defendant, she brings suit;[53] if she can prove she was injured by the defendant, she recovers her losses, and otherwise recovers nothing. (The defendant is strictly liable for any injury it is found to have caused. In this respect the model differs from the example given in part I above.)[54] The model's notation is as follows. Let

x	=	cost to the defendant of taking precaution
p	=	probability of plaintiff injury if the defendant takes no precaution
Δp	=	reduction in probability of plaintiff injury if the defendant takes precaution)
l	=	plaintiff's losses if she is injured

---

[52] This model draws heavily, in both structure and notation, on the one developed in Steven Shavell, The Social versus the Private Incentive to Bring Suit in a Costly Legal System, 11 J. Legal Stud. 333, 334-35 (1982). The purpose of the ensuing discussion is in essence to apply to the discovery context Shavell's insights concerning the social desirability of filing suit.

[53] We assume for the moment that the plaintiff sues if and only if she has been injured by the defendant -- that is, if and only if she has a valid claim against the defendant. The assumption is relaxed in part IIC below.

[54] The analysis would be largely unaffected if we instead assumed that the defendant would not be liable for the injuries it had caused unless it had violated the law (or the standard of care).

The plaintiff's chances of proving that the defendant caused her injuries depend on the amount of discovery allowed by the court. There are two possible levels of discovery, called low and high. High discovery is the more costly to the parties, and gives the plaintiff the better chance of proving her case against the defendant.[55] It is assumed that the plaintiff takes as much discovery as she is allowed by the court,[56] and that the case does not settle before discovery.[57] It is also assumed that the amount of discovery the court allows is fixed in advance of the defendant's decision whether to take precaution, and is known at all times to both parties. To denote the remaining variables, let

π = plaintiff's probability of winning given low discovery

$\Delta\pi$ = plaintiff's additional probability of winning

---

[55] For simplicity of analysis, it is assumed that the defendant is never erroneously held liable. The court may fail to impose liability on a defendant who caused the plaintiff harm, but it never imposes liability on a defendant who did not cause harm. Relaxing this assumption would not change the essentials of the analysis.

[56] In other words, if high discovery is allowed, the plaintiff takes it (implying that that $\Delta c_P < \Delta\pi l$); otherwise she takes low discovery. She never takes less discovery than the amount allowed.

[57] The assumption that the parties never settle before discovery is relaxed below in part IIC. Regarding the parties' behavior <u>following</u> discovery, it makes no difference whether we assume that the parties settle or go to trial -- provided that, if the parties do settle after discovery, they settle (on average) for the expected amount of the judgment. (If we assume they settle, C_P and C_D can be though of as involving settlement costs rather than trial costs.)

from allowing high discovery[58]

c_P = plaintiff's litigation costs given low discovery

Δc_P = increase in plaintiff's litigation costs from allowing high discovery

c_D = defendant's litigation costs given low discovery

Δc_D = increase in to defendant's litigation costs from allowing high discovery

The task of the social planner (court, rulemaking body, or legislature) is to decide which level of discovery the court should allow. The planner is assumed to have the objective of minimizing social costs -- that is to say, minimizing the sum of the defendant's costs of taking the precaution, the plaintiff's losses from being injured, and the parties' litigation costs.[59] Choosing the level of discovery that minimizes these costs involves resolving two issues: first, whether high discovery induces the defendant to take the precaution; second, if so, whether the gains from the precaution outweigh the costs of inducing them.

Deterrent effect of additional discovery. The first precondition for allowing high discovery is that it encourage precaution on the part of the defendant. For if high discovery had no effect on the defendant's behavior, then allowing it would

---

[58] That is, the plaintiff's probability of winning under high discovery is $\pi + \Delta\pi$.

[59] Here I follow Shavell, supra note 52, at 335.

plainly drive up total social costs: there would be no fewer injuries, and hence no fewer lawsuits, but the cost of each lawsuit would go up.[60] Thus high discovery cannot be desirable unless it tends to induce defendant precautions.

The formal expression of this precondition is as follows. Suppose the court allows only low discovery. If the defendant does not take precautions, its costs are $p(\pi l+c_D)$. If the defendant does take precautions, its anticipated costs are $x+(p-\Delta p)(\pi l+c_D)$. Assuming risk neutrality, the defendant will take precautions if

$$x+(p-\Delta p)(\pi l+c_D) < p(\pi l+c_D),$$

that is, if

$$x < \Delta p(\pi l+c_D).$$

The right hand term represents the benefits -- in reduced exposure to liability and legal expenses -- the defendant derives from taking precautions. If these exceed the cost of precautions, then low discovery is sufficient to induce the defendant to take precautions; there is then no point in allowing high discovery.

If, however, the right hand term above is less than the cost

---

[60] Recall our provisional assumption that all injured plaintiffs sue.

of precautions,[61] then low discovery will not induce defendant precautions. The question then is whether high discovery will do so. Suppose that the court allows high discovery. If the defendant takes no precautions, its anticipated costs are now $p[(\pi+\Delta\pi)l+(c_D+\Delta c_D)]$. If the defendant takes precautions, its anticipated costs are $x+(p-\Delta p)[(\pi+\Delta\pi)l+(c_D+\Delta c_D)]$. The defendant will take precautions if

$$x+(p-\Delta p)[(\pi+\Delta\pi)l+(c_D+\Delta c_D)] < p[(\pi+\Delta\pi)l+(c_D+\Delta c_D)],$$

that is, if

$$x < \Delta p[(\pi+\Delta\pi)l+(c_D+\Delta c_D)],$$

where the right hand term again represents the reduction in liability exposure and legal expenses achieved by taking precautions. If these exceed the cost of precautions, then high discovery successfully induces defendant precautions. But if the right hand term is less than the cost of precautions, then even high discovery is insufficient to induce precautions; and again there is no point in using high discovery.

Thus, high discovery cannot be worthwhile unless it is both necessary and sufficient to induce the defendant to take precautions. That condition is met if and only if the following inequality is satisfied:

---

[61] I ignore the case in which the expected costs of taking precautions equal the expected costs of taking no precautions.

$$\Delta p(\pi l+c_D) < x < \Delta p[(\pi+\Delta\pi)l+(c_D+\Delta c_D)]. \quad (1)$$

In practical terms, satisfaction of this condition means that high discovery, but not low discovery, leads the defendant to take precautions.

Social Value of Deterrence. Even if high discovery is necessary and sufficient to induce defendant precautions, it does not follow that high discovery is preferable to low. The deterrence achieved by employing high discovery may not be worth its costs.[62] Resolving this issue requires a comparison of the total social costs of high discovery and low discovery. In making this comparison I assume that inequality (1) is satisfied; we already know that if it is not, low discovery is preferable.

Suppose the court allows only low discovery. By inequality (1), the defendant will take no precautions. The plaintiff will be injured with probability p, and with that probability will sue and cause both sides to incur legal expenses. Total social costs per defendant under low discovery are thus

$$p(l+c_P+c_D).$$

Now suppose the court allows high discovery. By inequality (1),

---

[62] My accounting of the social benefits of discovery excludes the precedential value of judicial decisions. The scope of discovery probably has relatively little effect on precedent formation in any event. Rules of law (such as "defendants are liable when event E occurs") are often articulated on motions to dismiss that precede discovery.

the defendant will choose to take precautions. The plaintiff will be injured with probability $p-\Delta p$, and with that probability will sue. Total social costs under high discovery are thus

$$x+(p-\Delta p)\,[1+(c_P+\Delta c_P)+(c_D+\Delta c_D)]\,.$$

The social planner's task is to find the smaller of these two sums. The social costs of high discovery are lower than those of low discovery if and only if

$$p(1+c_P+c_D) > x+(p-\Delta p)\,[1+(c_P+\Delta c_P)+(c_D+\Delta c_D)]\,.$$

Rearranging terms, we get a second precondition for high discovery to be desirable:

$$\Delta p(1+c_P+c_D) > x+(p-\Delta p)\,(\Delta c_P+\Delta c_D)\,. \qquad (2)$$

The left hand side of inequality (2) represents the social benefits of high discovery, namely, the sum of plaintiff losses and legal costs that are <u>avoided</u> as a result of inducing defendant precautions. The right hand side represents the social costs of high discovery, namely, the sum of precaution costs and legal costs that it <u>generates</u>. If costs saved are greater than the costs generated, high discovery is worthwhile.

In sum, high discovery is warranted if and only if inequalities (1) and (2) are satisfied. This analysis gives a

general criterion for deciding whether allowing a given discovery request is or is not socially desirable. A request seeks a marginal increase in the level of discovery. "Low" discovery in effect refers to the amount of discovery occurs if the request is disallowed; "high" discovery is the amount that occurs if the request is allowed. The request is socially desirable if the two preconditions set forth above; otherwise not.

B. *Privately vs. Socially Desirable Discovery*

As Steven Shavell has shown, there is no necessary connection between private and social benefits of allowing a lawsuit to be brought.[63] So too with allowing (additional) discovery in the lawsuits that are brought. The private benefits of discovery lie in the increase it brings to the plaintiff's expected recovery at trial.[64] The social benefits lie in the reduction discovery brings to the sum of plaintiff losses, costs of precaution, and costs of running the legal system. The value of the plaintiff's private benefits (net of cost) is not a

---

[63] See Shavell, supra note 52, at 336. Refinements and objections to Shavell's argument appear in Peter Menell, A Note on Private Versus social Incentives to Sue in a Costly Legal System, 12 J. Legal Stud. 41 (1983); Louis Kaplow, Private Versus Social Costs in Bringing Suit, 15 J. Legal Stud. 371 (1986); and Susan Rose-Ackerman and Mark Geistfeld, The Divergence Between Social and Private Incentives to Sue: A Comment on Shavell, Menell and Kaplow, 16 J. Legal Stud. 483 (1987).

[64] The private benefit of discovery is defined here as its value to the requesting party.

reliable indicator of the net social benefits.[65]

Cooter and Rubinfeld's criterion of allowable discovery focuses entirely on the private costs and benefits of discovery.[66] Their standard would allow high (i.e. additional) discovery if, in our notation,

$$\Delta\pi l > \Delta c_P + \Delta c_D; \qquad (3)$$

otherwise it would allow only low (no additional) discovery. This test is very different from inequalities (1) and (2); unlike them, it says nothing about the deterrent value of discovery in relation to its expected costs. For that reason, it does not tell us whether the additional discovery is socially desirable. Sometimes inequality (3) will prohibit discovery that should, from a social standpoint, be allowed; other times it will allow discovery that should be prohibited.

The point is worth emphasizing. It is quite possible that allowing a given increment of discovery costs the parties to litigation more than it contributes to the plaintiff's claim, but nonetheless to reduce overall social costs. For if the prospect

---

[65] A similar discussion of diverging private and social incentives to acquire and present information to the court is found in Kaplow, supra note 51, at ___.

[66] To avoid any misunderstanding, I should emphasize that Cooter and Rubinfeld make no claim that their standard maximizes social welfare. Their standard is intended as a restatement of existing law, which in their view "focuses on private interests of the parties in civil disputes." Cooter and Rubinfeld, supra note 1, at [22]. My aim is to show that if existing law does indeed incorporate their standard, it probably fails to advance social welfare.

of such discovery induces defendant precautions, it lowers the number of injuries and thus the number of lawsuits. Though the discovery may be exceedingly expensive when litigation occurs, these costs may be more than justified by the benefits of having fewer injuries and less litigation in the first place.

Conversely, it is possible that allowing additional discovery costs the parties less than it contributes to the plaintiff's claim, but nonetheless raises overall social costs. The additional discovery may fail to induce defendant precautions, in which case allowing it is plainly wasteful from a social standpoint. Or it may induce defendant precautions, but at a cost (measured by the sum of precaution costs and added legal costs) exceeding the benefits (measured by the sum of plaintiff losses and legal costs avoided). Inexpensive though the discovery may be in relation to its return to the plaintiff's claim, there is no social value in allowing it.

The divergence between the private and social values of discovery can be illustrated with two numerical examples.

Case 1. Begin with a case in which additional discovery costs more than it contributes to the plaintiff's expected recovery. Assume that

Plaintiff's loss if injured (l)	=	\$100,000
Plaintiff's additional chance of winning from high discovery ($\Delta\pi$)	=	20%
Increase in each party's costs from allowing additional	=	\$15,000

discovery (Δc_P and Δc_D)

The return to the plaintiff's claim, $20,000, is less than the additional discovery's cost to the parties, $30,000. Thus, if inequality (3) is the test for allowable discovery, the additional discovery in this case should be prohibited.

But the additional discovery may nonetheless be socially desirable. Assume that

Cost of defendant precautions (x)	=	$12,000
Probability of injury if defendant takes no precautions (p)	=	40%
Reduction in probability of injury if defendant takes precautions (Δp)	=	20%
Plaintiff's chance of winning under low discovery (π)	=	30%
Each party's legal costs given low discovery (c_P and c_D)	=	$20,000

Given these parameters, high discovery is preferable to low. Inequality (1), the first precondition for allowing high discovery, is satisfied: under low discovery, the benefits to the defendant of taking precautions are $10,000, which is less than the costs of precautions; under high discovery, the benefits to the defendant of taking precautions are $17,000, which is greater than the costs of precautions.[67] Thus, high discovery is necessary and sufficient to induce defendant precautions.

---

[67] Per inequality (1), under low discovery, the benefits to the defendant of taking precautions are .2[.3(100,000)+20,000] = 10,000; under high discovery, the benefits are .2[(.3+.2)(100,000)+(20,000+15,000)] = 17,000.

Inequality (2), the second precondition, is also satisfied: the social benefits of the defendant's precautions are $28,000, while the social costs of bringing them about are only $18,000.[68]

Case 2. Now take a case in which additional discovery is privately desirable. Assume that

$$l = \$100,000$$

$$\Delta\pi = 20\%$$

$$\Delta c_P = \Delta c_D = \$5,000$$

Inequality (3) is now satisfied. The return to the plaintiff's claim, $20,000, exceeds the additional discovery's cost, $10,000. Thus the additional discovery passes Cooter and Rubinfeld's test.

But it may not be socially desirable:

(a) The additional discovery may not be necessary to induce defendant precautions. Given the above (case 2) parameters, assume that

$$x = \$12,000$$

$$p = 40\%$$

$$\Delta p = 20\%$$

$$\pi = 50\%$$

$$c_P = c_D = \$20,000$$

---

[68] These figures are derived from the expressions in inequality (2). The social benefits of the defendant's precautions are $\Delta p(l+c_P+c_D)$, that is, .2(100,000+20,000+20,000)=28,000. The social costs of bringing about the precautions are $x+(p-\Delta p)(\Delta c_P+\Delta c_D)$, that is, 12,000+(.4-.2)(15,000+15,000)=18,000.

Inequality (1) is not satisfied: under low discovery, the defendant's benefits from taking precautions are $14,000, which exceeds the costs of taking precautions. Low discovery thus suffices to induce defendant precautions, making high discovery unnecessary.

(*b*) The additional discovery may not be sufficient to induce defendant precautions. Assume that everything is the same as in (*a*) above, except that

$$x = \$18,000$$

$$\pi = 40\%$$

Inequality (1) is again not satisfied, but for a different reason: under high discovery, the defendant's benefits from taking precautions, $17,000, are less than the costs of taking precautions. High discovery, like low discovery, fails to induce precautions; so nothing is gained by allowing high discovery.

(*c*) The additional discovery may induce precautions, but at excessive cost. Assume that everything is the same as in (*a*) above, except that

$$p = 70\%$$

$$\Delta p = 12\%$$

$$\pi = 70\%$$

Inequality (1) is satisfied: under low discovery, the precautions' benefits to the defendant are $10,800, which is less than the costs of precautions; under high discovery, the benefits

are $13,800, which is greater than the costs of precautions. But inequality (2) is not satisfied: the social benefits of the defendant's precautions are $16,800, while the social costs of bringing them about are only $17,800. Inducing precautions with high discovery is not worth its costs.

A lesson of these examples is that the information needed to evaluate inequality (3) is not enough to decide whether additional discovery is socially desirable. Given certain values of l, $\Delta\pi$, Δc_P, and Δc_D, additional discovery may or may not reduce overall social costs, depending on the other information (x, p, Δp, π, c_P, c_D) needed to evaluate inequalities (1) and (2). Nor are the two sets of variables obviously correlated. For example, the plaintiff's losses (l) and additional chance of winning from additional discovery ($\Delta\pi$) do not imply anything about the costs of precautions (x) or precautions' reduction in the chances of injury (Δp). There is then no reason to suppose that using inequality (3) as a test for allowable requests will yield the right level of discovery.

I hasten to add, however, that it is hardly clear what test <u>will</u> bring about the right level of discovery. Taking inequalities (1) and (2) as our criterion of desirable additional discovery,[69] we face the problem of translating these

---

[69] Inequalities (1) and (2) leave out of consideration a number of matters that would figure into a more realistic model, such as the costs (apart from legal expenses) of operating the legal system; the possibility of plaintiffs' bringing nonmeritorious suits or declining (because of discovery costs) to bring meritorious suits; and the role of discovery requests made by

inequalities into judicially manageable standards. It is a substantial problem. A court would face daunting informational obstacles in deciding whether additional discovery satisfies the inequalities; its answer might involve little more than wild guesses about the deterrent value of additional discovery.[70] Finding the approach to discovery regulation that would yield the least error is no small task.[71]

---

defendants. Even in a more realistic model, however, the argument concerning the divergence between social and private benefits of discovery would hold. Compare Shavell, supra note 52, at 339.

[70] It is hard enough even to determine whether inequality (3) is satisfied. Compare Easterbrook, supra note 46, at 638-39 (noting informational barriers to determining what additional discovery contributes to the plaintiff's expected recovery (in our terms, the value of $(\Delta\pi l)$)). Deciding whether inequalities (1) and (2) involves evaluating a host of variables in addition to $\Delta\pi$ and l.

[71] It is conceivable, for example, that applying a rough categorical rule (such as "15 interrogatories for all cases involving less than $100,000") would come closer to satisfying inequalities (1) and (2) than would using a more precise, complex rule (such as "allow additional discovery if [you think] it satisfies inequalities (1) and (2)"). The former, though error-prone, may be no more so than the latter, and is cheaper to administer. On the problem of optimal precision and complexity in rules, see Louis Kaplow, Rules Versus Standards: An Economic Analysis, 42 Duke L.J. 557 (1992); Colin S. Diver, The Optimal Precision of Administrative Rules, 93 Yale L.J. 65 (1983).

Similar problems are encountered if, in lieu of imposing direct limits on discovery, courts use incentive devices such as fee shifting rules or damage multipliers. (An argument for fee shifting is made in Easterbrook, supra note 46, at 645-47.) Many of the social costs and benefits of discovery (as measured by inequality (2)) are externalized by the plaintiff in making the discovery request. A device that focuses on a single variable, such as a rule requiring the loser to pay the other side's costs, will not align her incentives with society's. (Compare Shavell, supra note 52, at 337.) More complex devices might be designed, yielding formulae that take account of other relevant variables. But this reintroduces the difficulty of accurately estimating the value of the variables in inequalities (1) and (2), which the court

C. *The Settlement Issue*

The model developed above ignores an important side of the optimal discovery problem. I have assumed to this point that parties never settle before discovery occurs. The analysis changes considerably if we instead allow for the possibility of settlement without discovery. A standard permitting an otherwise-ideal level of discovery might be undesirable <u>if it leads the parties to settle rather than undertake discovery</u>. In the remaining space I will merely sketch this problem, leaving for another time a full analysis.

The essential point is simple. The value of discovery is that it helps courts distinguish defendants who have committed some wrong ("injurers") from defendants who have not ("innocent defendants"). Such "sorting" is what gives the law its deterrent force; if the legal system treats innocent defendants and wrongdoers the same, it does not discourage wrongdoing.[72] The court's discovery standard may lead both innocent defendants and injurers to settle -- on similar terms -- <u>before</u> discovery. Ex ante, the effect is to weaken deterrence.

Seeing the problem requires relaxing two of our earlier assumptions. First, drop the assumption that there is only one

---

would have to do in deciding how much to award (or penalize) a given plaintiff.

[72] "Same treatment" here refers to both the probability and severity of sanctions. From a deterrence standpoint, it may be acceptable to (mistakenly) sanction innocents as severely as wrongdoers (we cannot hope to do otherwise!), provided that the wrongdoer is more likely to suffer the sanctions.

possible set of precautions the defendant can take. Assume instead that the defendant can take no precautions, a few precautions, or a lot of precautions. (The defendant is still assumed strictly liable for the injuries it causes.)

Second, drop the assumption that the plaintiff sues if and only she had been injured by, and was thus entitled to recover from, the defendant. Let us now assume, more realistically, that the plaintiff sometimes sues an innocent defendant because, before the evidence is disclosed, the defendant's innocence is private information to the defendant. In particular, the plaintiff sues anytime she suffers a given kind of injury. Let r denote the frequency ($0 < r < 1$) with which injuries such as the plaintiff's are caused by the defendant's (actionable) conduct, rather than by something else. Thus, upon suffering her injury, the plaintiff knows only that with probability r the defendant's conduct caused her injury, and that with probability $1-r$ the defendant is innocent.[73]

Our objective is to identify the conditions under which the case will settle before discovery. Let us assume, arbitrarily, that in settlement negotiations preceding discovery, the defendant always makes the final offer.

Low Discovery. Suppose that the court allows only low discovery. Given that she has been injured, plaintiff's expected

---

[73] I ignore the possibility that the plaintiff may sue the wrong defendant, i.e., that she sues defendant *A* when it was in fact defendant *B* that injured her.

gain (before discovery) from litigation[74] is $r\pi l - c_P$. The defendant's expected loss from litigation is $\pi l + c_D$ if the defendant is an injurer, and c_D if the defendant is innocent. Consider the amount an innocent defendant will offer to settle the case before discovery. It will offer the plaintiff an amount equal to the plaintiff's net expected gain from litigation, provided that figure is less than the defendant's costs of litigation; for otherwise the risk-neutral defendant would prefer to continue with the litigation.[75] The plaintiff will accept the offer if it is at least as great as her net expected gain from litigation. Thus, the case will settle before discovery if and only if[76]

$$c_D > r\pi l - c_P; \qquad (4)$$

if it does settle, the settlement amount will be $r\pi l - c_P$. If inequality (4) is not satisfied, discovery occurs and the

---

[74] I am still assuming that all innocent defendants that go through with discovery are exonerated. Thus, at the time of filing the plaintiff's chances of recovering her losses are (probability that the defendant is a wrongdoer) × (probability that the plaintiff can prove it), or $r\pi$.

[75] I am assuming here that there are only two "types" of defendant, namely, injurers and innocents. If there were different types among the injurers -- for example, those with very damaging evidence and those with less damaging evidence -- then there might be some unravelling, since the latter would not want to be pooled with the former. I put this complication to one side, since my purpose is to show how injurers and innocents may be pooled in settlement.

[76] I ignore the case in which the defendants costs of litigation are equal to the plaintiff's expected gain from trial.

defendant is exonerated.

Now consider the amount an injurer will offer. If inequality (4) is satisfied, then the defendant can offer the same amount as would an innocent -- namely, $r\pi l - c_P$ -- and have the offer accepted. Since an innocent would offer that amount, the injurer does not reveal anything about itself by making the offer. But suppose inequality (4) is not satisfied. The defendant must offer at least the plaintiff's expected gain from trial in order to settle the case; but if it makes such an offer, it reveals itself to be an injurer. Upon receiving any offer exceeding c_D, the plaintiff, realizing she faces an injurer, will refuse to settle for less than $\pi l - c_P$. Since that amount is still less than the defendant's expected loss from litigation,[77] the defendant will offer that amount.

Thus, if inequality (4) is satisfied, there is a pooling equilibrium in which both innocents and injurers pay the same amount[78] to settle the case before discovery. If inequality (4) is not satisfied, there is a separating equilibrium in which innocents are exonerated after discovery, while injurers pay their expected liability minus plaintiff's costs[79] to settle the case before discovery. In the separating equilibrium, injurers

---

[77] $\pi l + c_D > \pi l - c_P$ for all positive c_D and c_P.

[78] That is, $r\pi l - c_P$.

[79] That is, $\pi l - c_P$.

wind up incurring greater costs than innocents.[80]

High Discovery. Suppose instead that the court allows only high discovery. The plaintiff's expected gain (before discovery) from litigation is $r(\pi+\Delta\pi)l-(c_P+\Delta c_P)$. The defendant's expected loss from litigation is $(\pi+\Delta\pi)l+(c_D+\Delta c_D)$ if it is an injurer, and $c_D+\Delta c_D$ if innocent. By the reasoning above, if

$$c_D+\Delta c_D > r(\pi+\Delta\pi)l-(c_P+\Delta c_P), \qquad (5)$$

then there will be a pooling equilibrium in which both innocents and injurers pay the same amount[81] to settle the case before discovery. If inequality (5) is not satisfied, there is a separating equilibrium in which innocents are exonerated after discovery while injurers pay their expected liability minus plaintiff's costs[82] to settle the case before discovery. Again, in the separating equilibrium, injurers incur greater costs than innocents.

Now, the cases in which we are primarily interested are those in which one, but not the other, inequality is satisfied. Assume first that inequality (5), but not inequality (4), is satisfied. In this case, a rule allowing low discovery leads to

---

[80] Innocents pay c_D, the costs of litigation; wrongdoers pay $\pi l-c_P$. Since (by hypothesis) $c_D < r\pi l-c_P$, it follows that $c_D < \pi l-c_P$.

[81] That is, $r(\pi+\Delta\pi)l-(c_P+\Delta c_P)$.

[82] That is, $(\pi+\Delta\pi)l-(c_P+\Delta c_P))$.

outcomes in which injurers pay more than innocents, while a rule allowing high discovery leads to outcomes in which the two pay the same amount. It may then be (though it is not necessarily true)[83] that low discovery has a greater deterrent effect than high discovery. If so, allowing low discovery may -- depending on the net social value of deterrence -- be preferable to allowing high discovery.

Moreover, this may be true even though high discovery is itself socially preferable to low. That is, as between having the parties undertake low discovery or having them undertake high discovery, the social planner might well prefer the latter. Perhaps, for example, the prospect of undergoing high discovery will induce the defendant to take a lot of (socially desirable) precautions, while the prospect of undergoing low discovery will only induce a few precautions.[84] Yet these deterrent effects depend on the defendant's expecting high discovery to actually occur. If the defendant instead expects the case to settle without discovery, then the deterrent benefits of a rule allowing high discovery may evaporate. If inequality (5) (but not (4)) is satisfied, the social planner's choice is not between high discovery and low discovery; it is, in effect, between low discovery (if that is what is allowed) and no discovery (if high

---

[83] See note 72 supra.

[84] Recall the assumption that plaintiffs use all the discovery they are allotted. It is conceivable that, if high discovery were allowed, the plaintiff would nonetheless seek only low discovery -- and that, ex ante, defendants would expect plaintiffs to do this.

discovery is allowed). The optimal choice may be to allow "too little" discovery, since the ideal amount is not attainable.

A similar result obtains if we instead assume that inequality (4), but not (5), is satisfied. It might be that the social planner considers low discovery preferable to high, because the prospect of low discovery is all that is needed to induce a desirable level of defendant precautions. But low discovery may not be a real option, If low discovery is all that is allowed, the parties will settle without discovery; since defendants expect that to happen, the rule's deterrent effects may disappear. The social planner's choice may then in effect boil down to having high discovery or no discovery. The optimal choice may be to allow "too much."

The central problem in both cases is that, once again, what is good for the litigants is not necessarily good for the world. Society might in some instances be better off if plaintiffs turned down pre-discovery settlement offers equal to their expected gain from litigation; their doing so may produce separating equilibria with desirable deterrence properties. But when they make settlement decisions, plaintiffs do not worry about deterrence. It is too late, after all, to prevent their own injuries; and their individual settlement decisions are in any event unlikely to affect potential defendants' future behavior. So plaintiffs may individually decide to settle before discovery even when their aggregate welfare might be improved, in ex ante terms, if they could agree in advance not to.

At the margin, then, allowing discovery may sometimes impede the pursuit of accuracy in the legal system. A rule allowing the ideal amount of discovery may backfire if it leads the parties to settle without undertaking discovery.[85] Notice that this has nothing to do with parties using the threat of unreasonable or excessive discovery requests to extract a settlement from opponents. The problem identified here may occur even if no party ever demands, or threatens to demand, more than (what would otherwise be) the socially desirable amount of discovery.

III. Conclusion

This paper has emphasized the relation between discovery and accurate outcomes in litigation. On the positive side, I have distinguished two potential informational effects of discovery rules: informing the parties about the probable outcome of trial, and <u>determining</u> the probable outcome of trial by expanding the pool of evidence (information) available to the court. Both effects seem plausible, though I have suggested that the former effect may be less pronounced than is commonly thought.[86] Both seem to promote accuracy of outcome. The latter improves the

---

[85] A complete model of the problem might allow for the possibility of partial disclosure by the defendant before settlement. Perhaps the defendant could signal its innocence with somthing less than the full amount of disclosure required by discovery rules.

[86] Giving parties notice of the evidence that will emerge at trial has long been thought a -- perhaps the -- central effect of discovery. See, for example, Fleming James, Geoffrey C. Hazard, and John Leubsdorf, Civil Procedure 234-35 (4th ed. 1992), and the sources cited there.

quality of trials; to the extent settlements reflect the probable result of trial, it improves the quality of settlements as well. As for the former effect, it helps bring settlements into line with with the probable outcome of trial, as Cooter and Rubinfeld show. The magnitude of these effects, in relation to each other and to other potential effects -- such as raising the cost of litigation[87] -- remain open questions.

On the normative side, I have highlighted some of the intricacies of determining the appropriate scope of discovery in dispute resolution. If the social objective in litigation is to achieve optimal deterrence, it is fairly clear that a standard focusing on the private costs and benefits to the litigants won't yield the right amount of discovery. But it is scarcely clear how to design a standard (or pricing system)[88] that will. Part of the problem lies simply in identifying, in a given case, the level of discovery that optimizes deterrence. Even if that can be done, second-best issues are raised by the possibility that the parties will contract out of discovery by settling. Finding the best approach to discovery control -- a hardy perennial on the agenda of litigation reform -- will not be easy.

---

[87] See note 2 supra.

[88] See note 71 supra.

REFERENCES

Baird, Douglas C., Robert H. Gertner, and Randal C. Picker, *Strategic Behavior and the Law*, Unpublished Manuscript, University of Chicago Law School, April 1, 1993.

Cooter, Robert and Daniel Rubinfeld, "An Economic Model of Legal Discovery," *Journal of Legal Studies*, 1994, *23* (forthcoming).

Diver, Colin S., "The Optimal Precision of Administrative Rules," *Yale Law Journal*, 1983, *93*, 65-104.

Easterbrook, Frank H., "Discovery as Abuse," *Boston University Law Review*, 1989, *69*, 635-55.

Grossman, Sanford, "The Informational Role of Warranties and Private Disclosure about Product Quality," *Journal of Law and Economics*, 1981, *24*, 461-84.

James, Fleming, Geoffrey C. Hazard, and John Leubsdorf, *Civil Procedure* (4th ed. 1992).

Kaplow, Louis, "Private Versus Social Costs in Bringing Suit," *Journal of Legal Studies*, 1986, *15*, 371-86.

_____, "Rules Versus Standard: An Economic Analysis," *Duke Law Journal*, 1992, *42*, 557-623.

_____, "The Value of Accuracy in Adjudication," *Journal of Legal Studies*, 1994, *23* (forthcoming).

Menell, Peter, "A Note on Private Versus Social Incentives to Sue in a Costly Legal System," *Journal of Legal Studies*, 1983, *12*, 41-54.

Milgrom, Paul R., "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics*, 1981, *12*, 380-91.

Posner, Richard A., "An Economic Approach to Legal Procedure and Judicial Administration," *Journal of Legal Studies*, 1973, *2*, 399-455.

Rose-Ackerman, Susan, and Mark Geistfeld, "The Divergence Between Social and Private Incentives to Sue: A Comment on Shavell, Menell and Kaplow," *Journal of Legal Studies*, 1987, *16*, 483-91.

Shavell, Steven, "Sharing of Information Prior to Settlement or Litigation," *Rand Journal of Economics*, 1989, *20*, 183-95.

_____, "The Social versus the Private Incentive to Bring Suit in a Costly Legal System," *Journal of Legal Studies*, 1982, *11*, 333-40.