THE BARRISTER AND THE BOMB: THE DYNAMICS OF COOPERATION, NUCLEAR DETERRENCE, AND DISCOVERY ABUSE

JOHN K. SETEAR*

I. INTRODUCTION

Pre-trial civil discovery under the Federal Rules of Civil Procedure is something like the MX missile. When first conceived, it seemed like a good way to limit the possibility of a surprise attack by one's opponent at trial, or by the Russians, respectively. As the years pass, however, there are deep but distant rumbles that suggest something is awry. Many have come to believe that the promise of the original idea has been cannily perverted by those who actually give it form, whether their uniforms are of Brooks Brothers gray or Air Force blue. Critical articles proliferate. Multiple study commissions arise to stab with pins a beast that continues calmly to devour time and money.

This Article discusses both strategic nuclear weapons, of which the MX is perhaps the most potent in the U.S. arsenal, and discovery, which may be the most potent weapon in the litigator's arsenal. I examine these two areas through the lens of game theory, a tool commonly employed in defense analysis but less frequently used in legal discussions. The quasi-mathematics of game theory, of course, is a far-from-perfect lens for either viewing. It tends to obscure the common-sensical truths of personality and habit, both of which clearly influence the decisions of the superpowers and of opposing counsel. It filters out the fear and fatigue of a nuclear crisis and of a prolonged litigation. It requires careful attention, but fails to reward even the attentive with simple solutions.

A game-theoretic analysis of discovery abuse is nonetheless one useful perspective on the problem. By its exclusions, game theory can free us to consider the aspects of the problem that do not vary from case to case; the relative rigor of its style encourages clarity (and caution) in one's assumptions. And as we shall see, game-theoretic analysis of discovery abuse can provide some predictive, even some reformist, implications. The air is thin but clean.

Part II of this Article sketches a brief introduction to game theory, with particular emphasis on the "Prisoner's Dilemma." Part III examines discovery abuse as a static, single-decision game between attorneys, and then

* J.D., Yale Law School, 1984. Mr. Setear is currently employed by The RAND Corporation as a defense analyst. The views expressed herein are entirely those of the author, and do not reflect the views of The RAND Corporation or its research sponsors.
reviews the implications for the reform-minded of this static view of discovery abuse. In Part IV, the Article turns to dynamic, multi-move versions of the Prisoner’s Dilemma in superpower interaction and in discovery. The first of three dynamic perspectives considered is the “theory of cooperation.” This theory, of which Robert Axelrod is the foremost exponent, focuses upon the conditions under which cooperation can evolve in multi-move Prisoner’s Dilemma games despite the incentives for uncooperative behavior inherent in the single-move Prisoner’s Dilemma. The second dynamic perspective that the paper examines is what I call the “classical theory of deterrence.” This theory, commonly employed in the analysis of nuclear deterrence, focuses upon the conditions under which two parties will cooperate with one another because each threatens to harm the other if cooperation is not forthcoming. The final dynamic perspective that the paper examines is what I call the “revisionist theory of nuclear escalation.” This theory, of which Paul Bracken is the foremost exponent, focuses upon the conditions under which informational ambiguities might lead to the breakdown of nuclear deterrence under conditions that the classical theory of deterrence would consider sufficient for stability. Part V proceeds to apply these various theories to the issues of discovery abuse and to discuss various reforms suggested by the application of dynamic theories to discovery.

II. Static Game Theory

A. An Introduction to Game Theory

This section provides an introduction to game theory and sets forth some simple games. Those completely familiar with the subject may skip to section C; those who are generally familiar with the subject but might profit from a brief refresher on the Prisoner’s Dilemma may skip to section B.

Classical game theory begins by defining a “game” as a set of interactions of players’ choices in which each and every combination of the players’ choices lead unambiguously to a particular set of payoffs for each of the players. The number of possible choices for each player may be few or many. The number of players may also be few or many, though I will focus here on two-player games. The number of moves before the game ends may also be few or many, though there must be a clear way to specify when the game ends. Game theory is the child of economists, and the individuals hypothetically playing such games are a race of the species homo economus: they are rational enough to know from which strategies they and their opponents may choose, and to know the payoffs for all players if a particular

---

1 See infra text accompanying notes 36-64.
2 See infra text accompanying notes 64-71.
3 See infra text accompanying notes 72-89.
4 See infra text accompanying notes 89-117.
combination of choices occurs. Uncertainty may generally exist only to the extent that a particular forthcoming roll of the dice is unknown; the odds of the game, however, are clearly specified.

A simple example of a "game" in this sense is what might be called "Grandpa's Game." One player, i.e. Grandpa, magically produces a quarter from behind his ear, to the giggled delight of the other player, i.e. Granddaughter. Grandpa puts the quarter behind his back, along with his hands, and chooses whether to place the quarter in his closed right fist or his closed left fist. The other player (Granddaughter) chooses "right" or "left." If her choice matches the fist with the quarter, she receives the quarter as her payoff. If her choice does not match Grandpa's, then Grandpa keeps the quarter and she receives only a good-natured tweak on the cheek.

The boardgame "Monopoly" is a fairly complicated example of a game. Each player may decide whether or not to buy an unsold property upon which she lands, whether to pay to get out of jail or roll if misfortune places her behind bars, whether and when to buy houses, and so forth. Given several players and the unpredictability of the dice, the number of possible outcomes of the game is huge. Nonetheless, each one is unambiguously (if probabilistically) specified, and one could theoretically subject "Monopoly" to game-theoretic analysis. Most people, however, prefer to use it as a catalyst for familial strife.

Before proceeding further, let us examine the game theorist's equivalent to a Gilbert's or Legalines legal case summary: the "payoff matrix." The payoff matrix sets forth, in tabular form, the strategies and payoffs for a game without bothering with any intermediate steps, such as the players' actual reasoning. In a two-player game, the strategies are listed in the rows for one player and in the columns for the other player. The intersection of two strategies produces a cell in the payoff matrix in which the first of the pairing of numbers in the cell is the payoff for the player whose strategies are listed in the rows, and the second number in the cell is the payoff for the player whose strategies are listed in the columns.

Because a facility with such payoff matrices is as important to the game theorist as Marbury v. Madison \(^5\) is to the constitutional scholar, I, like Chief Justice Marshall, shall belabor the point. Figure 1 shows the payoff matrix for "Grandpa's Game." Figure 2 shows the payoff matrix for the slightly more interesting game of "Rocks, Paper, Scissors." Denizens of summer camp will remember that each of two players in the game simultaneously declares a strategy from among "rocks," "paper," and "scissors." A choice of rocks "breaks" a choice of scissors and triumphs in that pairing of strategies, but is "covered" by a choice of paper and loses in the paper-rocks pairing. Continuing the analogy to the everyday world, scissors cut paper and wins. Although the payoff can vary with context, it is shown in Figure 2 with a value of "1" for the winner and "-1" for the loser.

---

\(^{5}\) 5 U.S. (1 Cranch) 137 (1803) (Marshall, C.J.).
The goal of game theory is to illuminate which choices the rational player will make. The analysis is not much fun in the two games discussed above, since one player always wins exactly what the other loses (known technically as a "zero-sum" game) and since no particular strategy has any significant characteristics to distinguish it from the other available strategies. Hence, no opportunities exist for using cooperation to increase the gains of the players, and the choice of strategies can be made randomly with no harm to the player's own prospects of gain. I will turn now to two games—tantalizingly entitled "Chicken" and the "Prisoner's Dilemma"—in which the payoffs are not zero-sum and in which the choice of strategy can engage the attention of those with more than a playground education. They are also the two games of most interest in the game-theoretic analysis of superpower crisis and, as we shall see, discovery abuse.

**B. The Games of Crisis**

Having introduced the reader to game theory, I now, by way both of review and of preview, examine interactions between the superpowers in grave crisis through a game-theoretic lens involving the games of Chicken and the Prisoner's Dilemma.
1. **Chicken**

The first of these games of crisis is generally called "Chicken," supposedly named after the car game in which two youthful drivers and their respective companions race towards one another on a head-on collision course until either one car veers off (with the concomitant loss to its driver's social status) or the cars collide (with the concomitant loss of life). Perhaps because I grew up in the wrong generation, I have never heard of such a confrontation actually occurring. I therefore turn instead to a similar, but to my mind more plausible, situation to illustrate "Chicken." I call it "Faculty Meeting."
In this game, two professors—let me call them “Derek” and “Duncan”—attend a faculty meeting. An issue arises, and with it a difference of opinion between the two men that is suspended by the need for a break in the meeting. After the break, each player can choose from the strategies “hold forth” and “concede by silence.” If one pedagogue chooses to hold forth and the other chooses to concede by silence, then the holder forth sees his status rise amongst his fellow pedagogues, while the status of his opponent falls. If both concede by silence, both lose a bit of status among their backers. If both hold forth, both incur the mighty wrath of all their peers for delaying the meeting into the dinner hour. A payoff matrix is shown in Figure 3.

The important lesson here is that mutual stubbornness can lead to disaster—if both hold forth, their payoff is negative and of great magnitude. Yet both will be tempted towards such an outcome by the fact that one who outlasts the other gains at the other’s expense, and by the fact that mutual concession leads to a negative (though small) outcome for both.
A crisis between the superpowers is often considered to present an example of Chicken. Imagine yourself in Nikita Khrushchev's position after you have, for reasons known to you but that will forever be debated by historians, ordered the emplacement of offensive nuclear missiles in Cuba. The American President Kennedy, scion of a ruthless capitalist family, both demands that you withdraw the missiles and sets up a naval blockade around the island to prevent further relevant shipments to Cuba. In the vernacular of the time, cold-war tensions are high.

If you persist in your course and Kennedy backs down, your determination will be viewed as brilliant leadership and Kennedy's stock will fall significantly. If you are forced to back down without visible concessions from the United States, you are probably not long for the Kremlin in light of the spottiness of your nation's agriculture and economy—not to mention that your shoe-pounding at the United Nations has reflected poorly upon your nation's space-age sophistication. If you and Kennedy both concede through some political quid pro quo, you will look bad for having gotten the Soviet Union into a potentially dangerous fix in the first place, but will be able to claim credit for steering your nation through some rocky shoals and thus will not suffer dramatically. If you and Kennedy both refuse to give in, and that mutual stubbornness leads to nuclear war, the resulting loss to you and your nation is likely to be a catastrophe of unprecedented proportions.

The payoff matrix is shown in Figure 4, with "Stare" and "Blink" the shorthand for the stand-fast and concessionary strategies, respectively. The American superiority in nuclear weapons of the early 1960s is reflected in the slightly less catastrophic consequences for Kennedy of a mutual refusal to concede. In addition, since your domestic position is shakier than that of the witty and handsome Kennedy, the other gains and losses to you are more severe than they are for the American.

As with Faculty Meeting, the game of Missile Crisis has a worrisome structure. Both players will be tempted by the promise of a large gain to persist on their risky courses, but if both in fact persist, then disaster looms likely for both participants. In the nuclear shadow, such a structure seems likely in most crisis interactions between the superpowers. Generically, a game of Chicken requires that each player gain more from unmatched stubbornness than from mutual meekness, and that the loss from matched stubbornness exceed the gains from unmatched stubbornness. See Figure 5.

2. Prisoner's Dilemma

I turn now to the Prisoner's Dilemma, which has provided game theorists with as much grist for their mills as footnote 4 of Carolene Products has

---

6 See BARTLETT'S FAMILIAR QUOTATIONS 925 n.7 (E. Morrison 15th ed. 1980) ("'We're eyeball to eyeball, and I think the other fellow just blinked.'") (quoting Dean Rusk during the Cuban Missile Crisis).

provided legal scholars. Interestingly enough, the name of the Prisoner's Dilemma has its genesis in a legal story. Two gentlemen, whom we shall call Guido and George for present purposes, are arrested. They are jointly accused of a crime that, for purposes of explanation, we will assume they did in fact commit. The authorities separate the prisoners immediately upon their arrest and prevent them from communicating with one another.

The district attorney then separately presents each of them with a proposition. The district attorney tells Guido that he will be allowed to go scot-free if Guido will turn state's evidence against an allegedly tight-lipped George. If Guido is silent and George squeals, then the district attorney will see to it that the uncooperative Guido receives a sentence of eight years. (Both prisoners, let us assume, are practical sorts who believe that it is indeed the district attorney who sets their sentences.) Upon further inquiries from Guido, the district attorney reveals as well that he has enough evidence to send George and Guido to jail for two years each if both clam up, and that it would be his duty as a servant of the people to send George and Guido both to jail for five years each if the district attorney should be presented with the abundance of evidence available if both were to squeal. The district attorney then gives George the same information. The payoff matrix is shown in Figure 6, measured in years in prison.
Guido paces the cell and ponders his fate. Suppose, he says to himself, that George decides to turn state’s evidence. If I stay silent, then I unfairly receive eight years in the slammer; if I turn state’s evidence as well, however, then I will face a shorter stay in prison (just five years). It would seem, then, that my best strategy if George turns state’s evidence is for me to turn state’s evidence as well. Furthermore, Guido realizes, my best strategy even if George stays quiet also appears to be to turn state’s evidence: if I sing while George clams up, I get off scot-free, rather than having to spend the two years in jail that would be my reward if we both held our tongues. Ergo, if my best strategy if George squeals is for me to turn state’s evidence, and my best strategy if George doesn’t squeal is also for me to turn state’s evidence, then I should obviously turn state’s evidence. Impressed by his own brilliance, Guido asks for assistance in dictating his detailed recounting of the criminal events at issue.
Prisoner's Dilemma

![Figure 6](image)

At least as the story is told, George undertakes an analogous reasoning process in his own cell. He reaches an analogous conclusion: holding constant Guido’s strategy, I do better for myself to squeal rather than to clam up. George asks for assistance in dictating his recollection of the crime.

The outcome is straightforward. Both prisoners squeal. Both go to jail for five years.

From the prisoners’ point of view, something seems amiss. Both prisoners wind up in jail for five years, despite the fact that one of the outcomes open to them was for both to keep quiet and thereby wind up in jail for only two years. They both chose to squeal because it appeared to be the best response to a given choice of their partner in crime, regardless of which given choice the partner chose. The individual decisions in fact seem quite rational—why keep quiet and get taken for a ride for extra time in jail? The joint outcome of the rational individual’s decisions, however, seem irrational—a longer stay in jail. That is the essence of the Prisoner’s Dilemma: two parties acting individually wind up with an outcome worse for both of them than the outcome that they could obtain through cooperation.

The incentives for a preemptive strike in a modern naval encounter during a grave crisis between the superpowers possess a similar structure. Suppose that you are the commander of the USS Stark, repaired after its earlier misfortune in the Persian Gulf. You are steaming in the balmy Black Sea
with strict instructions to assert your freedom of passage through waters contested by a suddenly belligerent Soviet Union. Assume that relations between the superpowers have worsened dramatically, and that a concentration of Soviet naval combat vessels has been shadowing you for some time, just outside of missile range. Your offensive armament consists of very fast, long-range, highly destructive missiles; your defensive systems, even if turned on, are not likely to be very effective against similar missiles carried by your opponent.

If you wait for your opponent to fire and hit you first, you may be so badly damaged that you are unable to respond. If you fire first, you may harm them badly enough to prevent them from retaliating effectively against you. If you both fire at roughly the same time, you will both suffer significant but roughly equal damage. As you wipe the salt spray from your clipboard, upon which you have sketched the payoff matrix of Figure 7, you are told that the Soviet ships have just turned directly towards your ship and increased speed.

Thus, your options are threefold: (1) fire first and gain a unilateral advantage; (2) fire second and risk being unable to launch because of the damage sustained in the first-strike attack; or (3) do not fire at all and risk going down with your ship. The best outcome, of course, is for neither side to fire, at least for those who wish to live forever. Nonetheless, as in the Prisoner’s Dilemma discussed above, the incentives to fire, and thereby attempt to attain a unilateral advantage over one’s opponent, are strong. Mix into the brew the article of faith among members of the military that a commander of forces may take the necessary steps to protect those forces, and you may have a recipe for a serious international incident.

The generic recipe for the Prisoner’s Dilemma is not especially complex. The gain from taking unanswered advantage of one’s opponent—whether by squealing, shooting, or some other means—must exceed one’s gains from mutual cooperation, while one’s loss from being taken unanswered advantage of must be the worst of the four outcomes.\(^8\) See Figure 8.

### III. Discovery Abuse as a Static Game

Having introduced the basis of game theory and amplified that discussion by examining superpower interactions as illustrations of two types of games, I now analyze discovery under the Federal Rules of Civil Procedure in economic and then in game-theoretic terms.

\(^8\) Axelrod’s tournaments also required the alternation of defect/cooperate and cooperate/defect pairings to produce a lesser total gain for each player than two rounds of cooperation. See R. Axelrod, The Evolution of Cooperation 31 (1984).
Naval Encounter

<table>
<thead>
<tr>
<th></th>
<th>Captain Karamazov</th>
<th>Captain Caine</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hold Fire</strong></td>
<td>(No change,</td>
<td>(Expended a</td>
</tr>
<tr>
<td></td>
<td>No change)</td>
<td>Few Missiles,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Damaged and</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Disadvantaged)</td>
</tr>
<tr>
<td><strong>Let Fly</strong></td>
<td></td>
<td>(Damaged,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Damaged)</td>
</tr>
</tbody>
</table>

Qualitative Evaluation

<table>
<thead>
<tr>
<th></th>
<th>Hold Fire</th>
<th>Let Fly</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hold Fire</td>
<td>(0, 0)</td>
<td>(-12, 10)</td>
</tr>
<tr>
<td>Let Fly</td>
<td>(10, -12)</td>
<td>(-10, -10)</td>
</tr>
</tbody>
</table>

Quantitative Evaluation

FIGURE 7
A. **Impositional Discovery Requests**

It is first necessary to examine briefly the decision of a party to formulate a discovery request. In the economic view of such a decision, a party to litigation will tender a request whenever the benefits to her of doing so exceed the costs of formulating the request. It is possible to separate these benefits into two broad categories: "informational benefits" and "impositional benefits." "Informational benefits" are benefits that the requesting party expects to gain from the information that she receives from the responding party. Factual statements from the responding party can increase the requesting party's ability to hone the legal basis for her case, or help her estimate the value of the stakes in the case and her chances of prevailing on
the merits.9 "Impositional benefits," in contrast, are those benefits that the requesting party expects to gain because her request imposes costs upon the answering party.10 (I will argue shortly that the existence of such costs improves the requesting party's bargaining position in settlement negotiations.) Thus, informational benefits can be considered direct benefits to the requesting party, while impositional benefits may be considered indirect benefits to the requesting party that accrue because the responding party suffers direct costs from the request.11

Impositional benefits from discovery requests exist only if the receiving party is, in deciding whether to settle the case for a given offer or instead proceed to trial, sensitive to the costs that he will incur in taking the case forward. Suppose, for example, that you are a party who has just filed a suit. You have absolute confidence that your case will go to trial in a year and that there will then be a verdict in your favor for $1,000,000. (Let us suppose further that, like the medieval faithful, you believe interest to be immoral and therefore are indifferent between receiving $1,000,000 a year from now and receiving $1,000,000 tomorrow.) The defendant offers you $900,000 to settle the case. You expect your legal fees if you go to trial to be $200,000; moreover, as a basically gentle soul, you get no particular pleasure from dragging the defendant to trial. In these circumstances, you would be wise to settle. Your net benefits from going to trial would be only $800,000 ($1,000,000 from the verdict minus $200,000 in legal fees), while your net benefits from settling the case would be $900,000.

One can use a similar analysis in discussing impositional benefits from discovery requests. Suppose that you are a party who has just filed suit. You have absolute confidence that your case will go to trial in a year and that there will then be a verdict in your favor for $1,000,000. Your initial estimate of your legal fees is a mere $200,000, and you get no particular pleasure out of dragging the defendant to trial. The defendant offers you $700,000. You refuse, since you can receive a net of $800,000 from going to trial. So far, all is essentially as in the previous example, except for the lower offer by the defendant.

Now suppose that the defendant repeats the $700,000 offer and also surprises you with a discovery request that you estimate will cost you another $150,000 in legal fees incurred in responding to the statement. (Assume you also believe that seeking a protective order against the request would be fruitless.) You should rationally cave in with respect to the case as a whole, and accept your opponent's settlement offer. Going to trial, after


10 See id. at 356-57.

11 It is important to note that for purposes of this Article, I assume attorneys will faithfully serve as perfect agents for their clients. For a brief discussion of the problems and the accompanying implications of this assumption see id. at 357-58.
all, will net you only $650,000—the cool million from the verdict minus the $200,000 in fees that you originally expected and the additional $150,000 in fees from responding to the discovery request. From your opponent’s point of view, her discovery request has impositional benefits, because that request has forced you to accept a lower settlement offer than you would have accepted in the absence of the impositional request.

There is, of course, another factor in the tendering of an impositional discovery request: the cost of formulating the request. A rational requesting party will tender a discovery request only if she gains more from tendering the request, in terms both of informational and impositional benefits, than she expends in formulating it. If, for example, a requesting party can expend $1,000 on a discovery request and impose $5,000 worth of legal fees upon her opponent, the request is a rational one for the requesting party to make (even if there are no informational benefits from the request).

As for the correspondence between informational benefits or impositional benefits on the one hand, and legitimate or abusive requests on the other, I think it is safe to say that a request seeking only informational benefits is entirely legitimate while a request intended to bring only impositional benefits is entirely abusive. One of the primary purposes of the Federal Rules governing discovery is to reduce the chances of trial by surprise. To increase the likelihood of settlement through fear and trembling, in contrast, does not seem to be among those purposes. The remainder of this Article shall therefore focus upon the subset of rational requests in which the impositional benefits taken by themselves exceed the costs of tendering the request. I will call such requests “impositionally cost-effective” for short, and will leave the consideration of informational benefits to a day when discovery abuse is a less pressing problem.

The Federal Rules of Civil Procedure are generous both substantively

12 Id. at 361 n.56. The costs of formulating a request are less than the costs of responding to one, since requesting information is generally less expensive than searching for raw information and then shaping it into a suitable response. Id.

13 See FED. R. CIV. P. 26 (Advisory Committee notes to 1983) (“[t]he purpose of discovery is to provide a mechanism for making relevant information available to the litigants.”)

14 Federal Rule 26(b)(1) states:

Parties may obtain discovery regarding any matter, not privileged, which is relevant to the subject matter involved in the pending action, whether it relates to the claim or defense of the party seeking discovery or to the claim or defense of any other party, including the existence, description, nature, custody, condition and location of any books, documents, or other tangible things and the identity and location of persons having knowledge of any discoverable matter. It is not ground for objection that the information will be inadmissible at the trial if the information sought appears reasonably calculated to lead to the discovery of admissible evidence.
and procedurally\textsuperscript{15} to a party seeking information through discovery. A wide variety of discovery requests are therefore likely to be impositionally cost-effective. Jurists and practitioners can speak to this point with more precision than I, but even academics know that it is generally far easier to ask a question than to answer it. At the same time, of course, a requirement of impositionally cost-effectiveness still sets some upper bound on the number of discovery requests that could be undertaken in a particular case. After running out of requests that yield an informational benefit, a party would eventually wind up on a fishing expedition straying so far from the normal, impositionally cost-effective waters that the effort expended in attaching the bait of the request would exceed the rewards of the "catch" from inflicting impositionally costs upon one's opponent. At that point even abusive discovery would cease.

B. The Game of Discovery

I now examine the structure of the games of discovery created by the availability of impositionally cost-effective discovery requests. I assume that parties have a choice between two strategies, "threaten impositionally cost-effective discovery requests" (hereinafter "threaten abuse") and "refrain from threatening impositionally cost-effective discovery requests" (hereinafter "exercise restraint").

So long as the resources of each party are sufficient to respond to all the impositionally cost-effective requests of its opponent,\textsuperscript{16} discovery abuse takes the form of the Prisoner's Dilemma discussed in Part II. If neither party threatens abusive discovery, then the situation remains the same. If one party successfully threatens abuse and the other refrains, then the

\textsuperscript{15} A 1983 amendment to the Rules qualified this extremely generous standard. The amendment provides:

The frequency or extent of use of . . . discovery methods . . . shall be limited by the court if it determines that: (i) the discovery sought is unreasonably cumulative or duplicative, or is obtainable from some other source that is more convenient, less burdensome, or less expensive; (ii) the party seeking discovery has had ample opportunity by discovery in the action to obtain the information sought; or (iii) the discovery is unduly burdensome or expensive, taking into account the needs of the case, the amount in controversy, limitations on the parties' resources, and the importance of the issues at stake in the litigation. The court may act upon its own initiative after reasonable notice or pursuant to a motion under [Rule 26(c) for a protective order].

\textsuperscript{16} For a discussion of situations in which there is a significant asymmetry in resources or opportunities, see \textit{infra} pp. 606-09 (using model of nuclear deterrence) and Note, \textit{supra} note 9, at 371-72 (using quasi-economic model).
threatening party benefits from an improved bargaining position and the resulting favorable change in settlement offers. The passive, "abused" party's bargaining position deteriorates concomitantly. If both parties threaten to abuse, and pull up short of actual abuse, then their relative bargaining positions do not change, but they have both expended some small amount of resources in making their threats.

Figure 9 shows the payoff matrix for such a game. I assume that each party can obtain $32,000 worth of settlement value through the expenditure of impositionally cost-effective discovery requests. If both refrain from threatening such abuse, there is no change in the situation. If one party successfully threatens abusive discovery and the other exercises restraint, then the abusing party gains $32,000 in settlement value minus, let us say, $1,000 expended in the research necessary to back up the threat with credible details of potential future discovery, for a net gain of some $31,000. The refraining party loses the settlement value of $32,000. If both threaten to abuse and then back down, then both lose the $1,000 expended in backing up the threats, but there is no change in settlement offers for either side. Comparison with Figure 8 reveals that this game is indeed a Prisoner's Dilemma.17

It is well worth noting that the situation changes somewhat if, instead of merely threatening discovery abuse, a party actually follows through on his or her threats. In that case, the threatening party expends resources in formulating the request, and the threatened party expends (greater) resources in responding. Suppose, for example, that only one party makes and follows through on a threat of abuse. If the situation were otherwise as described above, then his expenditures would be $9,000 ($1,000 for making the threat and $8,000 for following through on it). He presumably gains no settlement value if the other party stonewalls, so his net loss is also $9,000. The threatened but unintimidated party's loss would be the $32,000 in expenditures needed to respond to the abusive request. Consider an alternative case in which both parties threaten abuse, and back up their threat with abusive requests. Both parties will then lose $41,000 ($1,000 for making the threat, $8,000 for following through, and $32,000 for responding to the other party's followed-through threat). See Figure 10. As comparison with Figure 10 will show, the overall situation in this case is more like a game of Chicken. Although there are some differences in the precise dynamics of the two situations—Chicken and Prisoner's Dilemma—cooperation in the form of mutual restraint is clearly desirable in both, and I will treat discovery abuse below exclusively as if it were a Prisoner's Dilemma.

17 See Note, supra note 9, at 362-64 (discussing how the discovery rules effectively place the litigants in a Prisoner's Dilemma by allowing a litigant to impose costs on the opposing side).
Discovery Abuse

<table>
<thead>
<tr>
<th></th>
<th>Exercise Restraint</th>
<th>Threaten Abuse Successfully</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exercise Restraint</td>
<td>(0, 0)</td>
<td>($31,000, −$32,000)</td>
</tr>
<tr>
<td>Threaten Abuse Successfully</td>
<td>($31,000, −$32,000)</td>
<td>(−$1,000, −$1,000)</td>
</tr>
</tbody>
</table>

- $8,000 worth of impositionally cost-effective requests
- Yielding $32,000 worth of response expenditures

FIGURE 9

C. The Static Prisoner's Dilemma and the Federal Rules of Civil Procedure

The Federal Rules of Civil Procedure contain a number of measures that are consistent with an acknowledgement that the practice of discovery presents litigants with a Prisoner’s Dilemma. Indeed, an analysis of the Rules’ application and use reveals an increasing sensitivity to the possibility of resolving this seemingly intractable dilemma. Subsection 1 below explores potential solutions to the Prisoner’s Dilemma generally; the subsection following, subsection 2, discusses the specific efforts of the Federal Rules of Civil Procedure to attempt to resolve the dilemma in the case of discovery.
1. Solving the Prisoner's Dilemma

Situations presenting participants with a Prisoner's Dilemma are a recurring social phenomenon. One can, for example, consider the Hobbesian war of all against all, with the resulting impetus for a social contract, to be an illustration of the Prisoner's Dilemma and its socially desirable solution. In the state of nature, each individual has an incentive to do violence to his comrades and take their food. Each individual also suffers if others violently take possession of his sustenance. In any given pairwise interaction, both parties will be better off with mutually peaceful behavior than with mutually aggressive behavior. The result is a Prisoner's Dilemma. See Figure 11. The result is also nasty, mean, brutish and short lives.

If, in contrast, participants enter into a social contract that creates a state with a monopoly on the unpunished use of force, then that state can force a mutually cooperative outcome. The state punishes those who choose the non-cooperative move with sufficient vigor and frequency to make the expected payoff, of an individual who uses violence, negative rather than positive. See Figure 12. Individuals then "voluntarily" choose to cooperate, and life becomes benevolent, kind, cultured, and long lives.
State of Nature

A similar situation arises with respect to the enforcement of legal contracts. Each party would like to be able to breach and get away with it, thereby gaining at the expense of the other party. If a breach were costlessly permitted, then contracts would routinely be breached, and the multiplicity of gains available from structuring economic activity in advance would be difficult to attain. If, in contrast, the state punishes those who breach private agreements, then the cooperative outcome becomes rational for both parties, and joyful businessmen may merrily structure their relationships optimally over time.

When parties first reach a private agreement between themselves, the state employs what I will unimaginatively call "indirect intervention": the state leaves to the individual parties the initial terms of the agreement and steps in only when a breach is alleged. "Direct intervention," in contrast, occurs when the state itself specifies directly the substantive choices available to the parties by setting forth the acceptable limits on behavior and threatening punishment if the parties exceed these limits. Contract law relies almost exclusively upon indirect intervention, whereas criminal law almost exclusively involves direct intervention. Labor law is an example of a
mixture of direct and indirect intervention. Although the state’s best choice of its method of intervention is a potentially complex question, I will assume that there are just two important factors in determining the mode of intervention: (1) the informational costs attendant upon the state’s determination of the party’s preferences; and (2) the moral costs attendant upon the state’s decision to allow the parties free reign.

High moral costs argue in favor of direct intervention; high informational costs make indirect intervention more attractive. For example, the moral cost of a non-fraudulent transaction of goods between two private parties is generally considered to be quite low; moreover, the particular structure of such deals is an issue on which the parties have information far superior to that generally in the possession of the state. Thus, non-fraudulent transactions of goods do not present an appropriate situation for direct intervention. In contrast, murder presents a perfect situation for direct government intervention, because the act of murder is generally considered to be such a morally reprehensible “exchange” that information costs are essentially
irrelevant. In the realm of labor law, some moral cost is attached to various potentially unfair exchanges (i.e. collective bargaining agreements), and much of the information relevant to these exchanges is difficult for the government to obtain. Labor law thus predictably contains a mixed bag of statutory prohibitions and governmental attempts to facilitate bargaining.

2. Static Solutions to Discovery Abuse

While discovery abuse frequently generates frustration in observers, few of these criticisms rise to the level of moral outrage. The information relevant to the exchange (discovery requests) generally does not rest in the repository of common knowledge, but rather inheres to the litigating parties. Since the moral content of discovery is low, and the information costs to the state of ferreting out the truth are high, one would expect indirect intervention by the state (in the form of the judge) to be the guiding principle of the Federal Rules of Civil Procedure. Indeed, virtually all of the interventions in the Rules share some flavor of an indirect intervention, in that they require the responding party to bring the alleged abuse to the attention of the court. In addition, several of the interventions also allow or require an explicit agreement between the parties before the state will step in.

The interventions that most nearly resemble purely indirect interventions occur under the auspices of Rule 29 and Rule 36. The former allows what is essentially a contractual option: the parties may reach any agreement on discovery matters that they desire, unless the court specifically prohibits the agreement.18 This procedure is a useful lubricant to the discovery process, but one must assume that it does not, in practice, remove a great deal of opportunity for one party to impose significant costs upon the other. Typically, agreements under Rule 29 cover medical examinations, transcription fees, or the timing of a deposition—not trivial issues, but hardly the sort of request that most would consider prototypical abuse.19

Rule 36 is an indirect intervention with a bit more force behind it. Under this rule, parties may request of one another an admission that a particular fact is true.20 A party who refuses to admit to a particular assertion is liable to sanctions, under Rule 37(c), for the requesting party’s expenses in proving the admission true if it is later shown that the responding party’s failure to admit the specified assertion was unreasonable.21 Once again, however, one

18 FED. R. CIV. P. 29 provides in pertinent part: "Unless the court orders otherwise, the parties may by written stipulation . . . modify the procedures provided by these rules for [the] methods of discovery . . . ."
19 See Note, supra note 9, at 364 n.71.
20 FED. R. CIV. P. 36.
21 Id.
should be cautious about the magnitude of impositional costs saved by conscientious application of Rule 36.

More typical, or at least more numerous than these purely indirect interventions, are interventions that still require a party to bring the "violation" of discovery rules to the attention of the state, but that do not require a pre-existing agreement between the parties. These rules therefore do not represent purely indirect interventions, but do retain much of their flavor. The enforcement mechanisms for these examples require that any party not obeying the code of discovery etiquette pay the costs thereby incurred by its opponent. If, for example, a party is brazen enough to skip a scheduled deposition or to ignore completely an interrogatory, then Rule 37(d) allows the court to hold the party given the request liable for the costs that its inaction has imposed upon the requesting party—once the responding party has brought a motion asking for such an action by the court.22 Courts may similarly impose costs on those litigants who, viewed in the light of the information already in their possession, display a deep reluctance to answer interrogatories in good faith.23 Entreaties from a party may also lead a court to charge an uncooperative party for the costs to the other party resulting from the former's failure to obey the order stemming from a discovery conference.24 Indeed, any stand by either side on a discovery issue that winds up in court and is not "substantially justified"25 can result in a decision by the court to require the uncooperative party to refund to its opponent any reasonable costs resulting from the opponent's need to obtain an order to settle the matter in its favor.

Courts can not only impose costs as sanctions, but can also intervene directly and place substantive limits on the information that a requesting party may pursue. The discovery conference described in Rule 26(f), for example, can result in substantive limits rather than monetary penalties; it is also one of the rare weapons in the battle against discovery abuse that the trial court can employ without having to wait for one of the parties to request its use. In addition, a 1983 amendment to the Rules set forth some new and useful "express criteria"26 for judges to use in determining when to employ substantive limits: it imposes on judges an obligation to limit discovery that is either redundant or unduly burdensome under the circumstances of the

22 FED. R. CIV. P. 37(d); see Note, supra note 9, at 366 nn.78-79 (discussing actions sufficient to avoid FED. R. CIV. P. 37(d) sanctions).
23 See Note, supra note 9, at 366 n.80 (asserting that the courts have construed FED. R. CIV. P. 37(d) to mean that parties need not answer interrogatories that the requesting party could perform equally or more cheaply).
24 FED. R. CIV. P. 26(f); FED. R. CIV. P. 37(b)(2).
26 FED. R. CIV. P. 26(b)(1).
action. Moreover, it allows them to employ these measures either on their own initiative or at the behest of the responding party.

Rule 33(c) provides an even more specific direct intervention of the court in the discovery process: a trial court, prompted by a request from the responding party, may allow a responding party to answer an interrogatory with the documents necessary to ferret out the answers rather than with the answers themselves. This direct regulation of the interchanges of discovery obviously reduces the impositional costs associated with such a request, thus reducing the impositional benefits of such interrogatories to the requesting party—and with them the benefits of abusive discovery. Currently, however, the court may apply this rule only when the requesting party can produce the answers from the documents with no more effort than the responding party. As long as the responding party is more efficient at producing the information—which is likely to be true whenever the information at issue was initially within the responding party’s possession—Rule 33(c) will not hinder abusive litigants from using interrogatories to impose costs upon their opponent. Indeed, the reported cases in which trial courts have used this tool frequently involve such matters as citations to literature and the weather at the time of an accident, both matters that seem unlikely to cut to the heart of discovery abuse.

All these interventions to reduce impositional costs are laudable, though limited. The sanctions on costs, for example, prevent a party from making an outrageously impositional request and then forcing the responding party to foot the bill for persuading the court that the stance is not substantially justified. Some impositional requests may therefore never see the light of the discovery day. But, if the abusive requesting party fails in his efforts, he suffers only the relatively minor penalties of paying for his opponent’s costs of preventing the discovery (and, even then, only if his stand is not substantially justified). If, however, he succeeds in his efforts, the abusive party forces his opponent to undertake what seems likely to be a far costlier proposition—the response to the abusive requests. Clearly, then, the benefits of a successful impositional discovery request are likely in many instances to outweigh the mild court-ordered penalties mandated by the Federal Rules of Civil Procedure’s efforts at indirect intervention.


FED. R. CIV. P. 26(b)(1).

FED. R. CIV. P. 33(c).

See Note, supra note 9, at 367 n.87.


As to the direct interventions involving substantive limits on the discovery that may be pursued—and also with respect to determining when to impose sanctions under the indirect-intervention methods—the judge faces some inherent difficulties in placing herself in the shoes of the requesting party to see if the request is, in fact, abusive. After all, the judge knows little about the information available to either the responding or the requesting party. If the responding party were to produce fully the information sought and display it to the judge, the responding party would already have incurred all the costs that the responding party was trying to avoid. And in the absence of such production, the responding party has some leeway with which to exaggerate the burden of response, and a wise judge will thus to some extent discount the protestations of that responding party.

The requesting party provides no more reliable guidance, as that party possesses the incentive to exaggerate, in a complementary fashion, the ease with which the responding party can produce the information. Indeed, even a party making a legitimate request is likely to have only a faint idea of what the responding party must do to respond. It is therefore not surprising that, although the writers of the Federal Rules have taken some steps to lessen discovery abuse that are consistent with viewing discovery abuse as a static Prisoner’s Dilemma, few observers of the scene appear satisfied with the current state of affairs. In hopes of illuminating the dark glass of discovery, the next section examines discovery abuse from various dynamic perspectives. These perspectives place significantly less emphasis upon the role of a central, i.e. judicial, authority in analyzing the problem. Such an approach is more promising than a static perspective in light of the fact that the central authority in the case of discovery abuse is inherently deprived of important information necessary to solve the Prisoner’s Dilemma effectively.

IV. Dynamic Game Theory

I have thus far examined the Prisoner’s Dilemma and discovery abuse only from a static perspective—that is, one in which the game consists of a single, simultaneous choice of strategies by the two players. Discovery, however, is generally a dynamic process. One side’s request, or threat to make requests, lands on the doorstep of the other; the responding party answers that request or threat, and may then compose its own requests or threats. The ping-pong game of discovery can continue, with several balls in the air, right up until trial. Parties can therefore choose their strategy for a particular request dynamically—that is, in response to an earlier choice of strategy by their opponent. This Part examines three dynamic perspectives not directly addressed to discovery. I apply them to the problem of discovery abuse in Part V.

The first dynamic model, which I call “the theory of cooperation,” is an abstract, game-theoretic model of how a cooperative—i.e. mutually restrained—solution to the Prisoner’s Dilemma can evolve even without a
central authority to impose or enforce such a solution. The second dynamic model that I set forth, which I call "the classical theory of nuclear deterrence," examines the conditions under which the promise to use nuclear weapons against one's opponent can lead to a cooperative solution to the Prisoner's Dilemma. The third dynamic theory examined in this Part, which I call "the revisionist model of nuclear escalation," asserts that a crisis between the superpowers could lead to nuclear war even if the conditions for stable deterrence under the classical theory are met.

A. The Theory of Cooperation

In *The Evolution of Cooperation*, Robert Axelrod sets forth a general theory of the conditions under which the dynamics of the Prisoner's Dilemma can lead to a stable and cooperative outcome, even in the absence of a central authority to impose cooperative solutions upon the parties, such as the judicial interventions discussed in Part III. In light of the many demands made on a judge's time and the inherent difficulty that judges face in accurately determining the precise benefits and burdens of a particular discovery request, scholars and jurists concerned with discovery abuse should be interested in encouraging such favorable dynamics.

1. Axelrod's Assumptions, Methodology, and Results

In considering the dynamics of the Prisoner's Dilemma, Axelrod begins with the implicit assumption that, whether through conscious choice or genetic programming, individuals choose their strategy in a particular move of the multi-move game through the use of what is known as a "meta-strategy." This meta-strategy does not vary from move to move, though the strategies that the meta-strategy determines for each move can change. Axelrod expressly assumes that each member of a community can identify all other members of the community and can perfectly remember her previous interactions with them.

A meta-strategy generally chooses a strategy for a particular move by applying its rules to the history of interactions between the two parties to

---

33 I argue in Part V that this model implies that recent developments in the general structure of litigation and the practice of law have aggravated the problem of discovery abuse. See infra pp. 616-23.

34 I argue in Part V that this model implies that certain asymmetries between the parties in their resources or their opportunities to conduct discovery affect the likelihood that one or both of the parties will abuse discovery. See id.

35 I argue in Part V that this model implies that discovery is a process with some inherent tendencies toward the litigative Armageddon of extensive abuse. See id.

36 R. AXELROD, suprano note 8.

37 Id. at 11.
that move. For example, a meta-strategy might simply specify that the choosing individual will cooperate if the comrade she encounters cooperated in either of their previous two interactions, and will otherwise defect. Alternatively, a meta-strategy might specify that the choosing individual cooperate if the current opponent has cooperated more frequently in the past than it has defected. A meta-strategy can employ a random element, even to the extent of ignoring past interactions completely and simply choosing randomly whether to cooperate or defect.

Axelrod ran two round-robin "tournaments" of such meta-strategies, with the entries solicited from professional game theorists. With the assistance of a computer as bookmaker and bookkeeper, Axelrod let each meta-strategy interact for exactly 200 moves with every other meta-strategy. He awarded them points according to the particular manifestation of the Prisoner's Dilemma shown in the payoff matrix of Figure 13.

In the first such tournament, a meta-strategy called TIT FOR TAT was the winner. TIT FOR TAT always cooperated the first time it encountered a given comrade; after that, it cooperated in the current play if the comrade had cooperated in their previous encounter, and defected in the current play if the comrade had defected in their previous encounter. In other words, TIT FOR TAT offered its cheek for the first encounter, but otherwise adhered to standards of Old-Testament justice.

The result was that TIT FOR TAT did not typically do much better than its opponent in its overall score against any particular opponent, but did very well in comparison to the field as a whole. This apparent paradox is not really so puzzling. The cooperative outcome has a large but equal yield for the two cooperators. "Winning" in Axelrod's tournament was a matter of a meta-strategy's own overall yield, not of the number of times that a meta-strategy bested its opponents. TIT FOR TAT, for reasons I will discuss more fully below, encouraged cooperation and thereby shared the rich bounty of cooperation time after time. The tournament proved to be an environment in which it was more important to keep Rome from burning than to outfiddle the other meta-strategies in each pairwise encounter.

After transmitting to potential entrants the results of the first tournament, Axelrod held a second tournament. In a much larger field, the TIT FOR TAT meta-strategy again placed first. Axelrod also conducted an "ecologi-

---

38 Id. at 31-40, 43-48 (discussing the various meta-strategies employed by players involved in a Prisoner's Dilemma).
39 Id. at 30.
40 Id. at 8, 31.
41 Id. at 31. Axelrod capitalizes all the letters of the names of all the meta-strategies. I follow his convention here.
42 Id.
43 The second tournament involved 62 entries as compared with just 14 in the first tournament. Id. at 31, 41.
44 Id. at 42.
Axelrod's Dilemma

<table>
<thead>
<tr>
<th></th>
<th>Cooperate</th>
<th>Defect</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cooperate</strong></td>
<td>(3, 3)</td>
<td>(0, 5)</td>
</tr>
<tr>
<td><strong>Defect</strong></td>
<td>(5, 0)</td>
<td>(1, 1)</td>
</tr>
</tbody>
</table>

FIGURE 13

cal" tournament that sought to examine the dynamics of interaction when less successful meta-strategies "died out" over time. The initial pool of meta-strategies—the "first generation"—was apparently entered as in the second tournament described above, with each of the entries represented once. At the close of the first 200 round-robin, the meta-strategies "reproduced" so that their resulting percentage presence in the meta-strategy pool was the same as their percentage share of total points in the points pool of the previous 200 round-robin. For example, if TIT FOR TAT had garnered 4,000 points out of a total of 100,000 points awarded in the first 200 rounds, then 4% of the "second generation's" meta-strategies would be TIT FOR TAT—not just the single TIT FOR TAT meta-strategy that was present in

45 Id. at 48-53 (discussing the ecological tournament and the numerous principles which dictated the course it would take to produce a "winner").
the first generation. The ecological tournament proceeded for 1,000 generations. In the end, the TIT FOR TAT meta-strategy was at the top of the evolutionary heap, just as it had topped the other heaps in the other tournaments. The TIT FOR TAT meta-strategy thereby showed not only its ability to interact with a wide variety of different meta-strategies for a given generation but also its ability to continue to flourish as its competitors more frequently became those who had succeeded in previous generations.

2. Axelrod’s Inferences

On the basis of the success of TIT FOR TAT, as well as of an analysis of the other meta-strategies entered and their success or lack thereof, Axelrod reached some general conclusions about which kinds of meta-strategies are likely to do well in the multi-move, or “iterated,” Prisoner’s Dilemma. Most important for our purposes is his conclusion that successful meta-strategies are likely to be characterized by four adjectives: “nice,” “retaliatory,” “forgiving,” and “clear.”

A “nice” meta-strategy never defects first. TIT FOR TAT, for example, cooperates on the first move and does not defect until its opponent has defected. It is therefore nice. RANDOM, a meta-strategy that makes its choices at random, is not nice, because it might in its randomness defect against even an opponent who had an “always cooperate” meta-strategy.

A “retaliatory” meta-strategy is one that immediately defects after an “uncalled for” defection by its opponent. TIT FOR TAT always and immediately defects if its opponent has defected, and is therefore retaliatory. An “always cooperate” meta-strategy is nice, but it is not retaliatory. Other players can easily take advantage of such an easygoing strategy.

A “forgiving” meta-strategy is one that is not significantly more likely to defect just because its opponent defected a long time ago. A “forgiving” meta-strategy therefore has a relatively high propensity to cooperate even if the other player has sometimes defected in the past. TIT FOR TAT’s Hartpencian “memory,” for example, stretches exactly one move into the past; it forgets, and thereby forgives, quite quickly. Contrast TIT FOR TAT with Captain Ahab’s meta-strategy, in which a single act of cetological insubordination leads him to a relentlessly unforgiving strategy in future interactions with his antagonist.

46 Id. at 52.
47 Id. at 53.
48 Id. at 20, 54.
49 Id. at 33 (noting that this “niceness” quality distinguished the high-scoring entries from the relatively low-scoring participants).
50 Id. at 44. Axelrod himself acknowledges that the meaning of “uncalled for” is not precisely determined. Id.
51 Id.
52 Id. at 36.
Finally, a meta-strategy is "clear" if its behavior is easy to predict on the basis of a small number of previous interactions. TIT FOR TAT looks only one move back into the past, has a simple rule for translating the past into a present action, and has no random element. It is therefore a clear meta-strategy. In contrast, RANDOM, along with meta-strategies that use some complex weighting of a number of past strategy choices by the opponent, are unclear. As most of us know, the next move of either a madman or a genius is often difficult to predict. A clear meta-strategy, in contrast, allows one player to adapt to the other's patterns of action.

TIT FOR TAT possesses all four of the characteristics that are desirable in a meta-strategy for an iterated Prisoner's Dilemma—niceness, forgiveness, retaliation, and clarity—and in fact did outstandingly well in the various tournaments that Axelrod created. All things considered, then, the ideal meta-strategy for an iterated Prisoner's Dilemma appears to bear close relation to that typically adapted by a four-year-old child in the presence of her grandparents: cooperative while paid attention, and prone to an immediate but quickly forgotten tantrum if ignored. The reasons behind the success of such meta-strategies are not difficult to understand, at least with the benefit of hindsight. In the Prisoner's Dilemma, mutual cooperation is the goal, while being taken advantage of is the fear. Nice meta-strategies indicate that you lean towards cooperation. Forgiving meta-strategies confirm that leaning. Retaliatory meta-strategies punish a meta-strategy that takes advantage of you. Clear meta-strategies show the other party that you know the difference between cooperation and the taking of advantage. What is more surprising, perhaps, is the robustness of the TIT FOR TAT meta-strategy: it retained its supremacy despite competing with some other sophisticated meta-strategies designed to take advantage of it. The first conclusion that one might glean from this analysis is that TIT FOR TAT is probably not a bad strategy in many examples of the iterated Prisoner's Dilemma, though there is no guarantee that a particular configu-

---

53 Id. at 20.
54 Id. at 20.
55 Id. at 54.
56 Id. at 53. For example, TESTER was a meta-strategy designed to look for soft, exploitable players, but back off if the other player shows it won't be exploited. Towards this end, it defects on the very first move to test its opponent's response. If the opponent defects, TESTER quickly apologizes by immediately cooperating and playing TIT FOR TAT for the rest of the game. If the opponent does not defect, TESTER cooperates on the second and third moves but defects every other move for the rest of the game.

When matched against TIT FOR TAT, TESTER would quickly apologize to TIT FOR TAT in order to obtain mutual cooperation. By attempting to take advantage of TIT FOR TAT, TESTER simply hurts itself. Id. at 53.
57 Id.
ration of other strategies will not make a more successful and resilient meta-strategy, besides TIT FOR TAT, the best choice.

The second conclusion that one might reach is that one would be wise to examine the circumstances in which TIT FOR TAT or its cousins is a poor strategy. Axelrod undertakes this examination by focusing on what he calls "the shadow of the future." The general notion of the shadow of the future is that, in different circumstances, the future may cast a shadow of varying lengths—and, as it turns out, the longer that shadow, the more fertile is the field for cooperative solutions to an iterated Prisoner's Dilemma.

Axelrod sometimes uses the term the "shadow of the future" to mean what economists mean by the discount rate, or what financiers glean from the interest rate: how much less does the future inherently matter to the players than does the present? Most people consider even the utterly reliable promise of a particular future gain to be less valuable than the receipt of the same magnitude of gain in the present. The importance to the success of cooperative meta-strategies of this differential valuation of present and future is not so difficult to see. If the future is of almost no importance relative to the present, then a non-cooperative meta-strategy can choose the exploitative, non-cooperative one-move strategies in the early moves and make a quick killing against its cooperative comrades. The risk of its thereby forcing the game into a low-yield, mutually non-cooperative mode carries little weight if the future is heavily discounted. The non-cooperative meta-strategy can outperform its cooperative competitors even in the long run, because the "long" part of the run is worth so little.

If, in contrast, the future is worth nearly as much as the present, then cooperation is wiser. A failure to cooperate in early interactions with a comrade is likely to lead to future failures to cooperate that deprive a player of what would be a heavily weighted, larger pie; and early gains from taking advantage of another's cooperativeness would be unlikely to be sufficiently large to make up for the later losses from being unable to cooperate.

Those who seek empirical confirmation of this phenomenon might consult the bevy of learned articles bemoaning the rapaciousness of corporations in which short-term performance is the gauge of managerial success. One might also note the surly uncooperativeness of the adherents of punk rock and remember that the Ramones, an early punk-rock band, had a cult hit with a song whose most comprehensible lyrics are: "No future, no future, no future for me."

In addition to this discount-factor view of the "shadow of the future," Axelrod also sees the likelihood of future pairwise interactions between a given pair as part of the shadow of the future. In real-life interactions, as in

---

58 Id. at 12, 124.
59 Id. at 126-32.
60 See id. at 12-13, 126-27.
61 Id. at 126-32.
62 Id. at 20, 126-30.
Axelrod's tournament, the precise ending of the overall "tournament" is generally not known to the participants, but one can sometimes guess that one will not likely run into another player again. Credible rumors may fly that a particular academic is thinking about leaving her field for even greener pastures, or one may have a chatty roommate on a plane flight from one obscure airport to another. If the likelihood of future interaction between the players is low—as in these examples—then the shadow of the future effectively shortens, and the attractiveness of exploitation rises. The circling sharks, for example, know that a quick bite is not likely to lead to future retaliation because there is unlikely to be a future interaction.

Evidence for the proposition that the likelihood of future interactions affects the cooperativeness of current interactions is not difficult to find. Compare New York City with Mayberry, R.F.D. In the latter, the blissful cooperation of one resident with another is plausible not only because Andy Griffith has a kind face, but also because small towns—in the real world or on television—provide the opportunity for repeated interactions between individuals. And in the days when one did not stray far from one's place of birth, those interactions were likely to continue far into the future. The incentives for cooperation in such a situation are high, since one's reputation is likely to be well known to those with whom one later interacts and one will have many future opportunities to reap the gains of a good reputation. In New York City, in contrast, the number of actors is huge and their mobility is great. The probability of future interactions is small; the incentives for cooperation are therefore low. Unsurprisingly, then, the strangers who call you "buddy" don't usually mean it.

The importance that attaches to the likelihood of future interactions between players has a bright side for those who might otherwise fear that virtue is never rewarded in the big, bad world out there. Like a knot of luminaries at a cocktail party, a small number of cooperative players can create an immediate environment so rich that they do not need to care about the larger picture. Two cooperative players can each gain a great deal from their efforts, and if the likelihood of their own particular future interaction is high, they will come to cooperate (assuming that both their meta-strategies are forgiving), and thereby do quite well for themselves even in the midst of a number of more exploitative comrades.\(^\text{63}\) Axelrod shows that such meta-strategies can improve upon the gains reaped by an unmitigatedly hostile meta-strategy even when those employing TIT FOR TAT have only five percent of their interactions with one another and spend the rest of their time interacting with uncooperative actors.\(^\text{64}\)

Empirical confirmation on this point, at least of the experimental sort, is once again not difficult to call to mind. Immigrant communities in the United

\(^\text{63}\) Id. at 63-69 (explaining how the higher points awarded for mutual cooperation will allow cooperative players to prevail over those who defect every move).

\(^\text{64}\) Id. at 64-65.
States can often succeed, in spite of the prejudices of the general citizenry, by interacting frequently with one another. In such a situation, cooperative meta-strategies become profitable, and diamonds will be traded on a handshake.

One might also note that initially uncooperative individuals should eventually be won over to the cooperative meta-strategy out of sheer self-interest. The cooperative environment is rich once it has been established, and the fact that it is best "plundered" by the rather unpiratical meta-strategy of cooperation will make it in the interest of even the naturally invidious to undergo a conversion to communitarianism. Rigorous empirical evidence on this point is perhaps lacking, but, in this Age of Anecdote, perhaps a personal example will suffice. I came to law school aware of the practice of pooling one's notes with others, parcelling out various portions of the course among the pool's members for outlining as the exam approached, and sharing the resulting outlines. As an economics major well schooled in the "free-rider" problem and as a recent reader of newspaper articles discussing the cut-throatedness of the modern law student, I was dubious about the quality of output that such a cooperative project would produce. Nonetheless, I soon came to see the advantages of such an arrangement and underwent a self-interested conversion to communitarianism, for I found that—at least in a small group of outliners at a small law school—the cooperative solution of high-quality work was in fact stable.

I shall apply these lessons about the "shadow of the future" to discovery abuse in Part V, but first I will examine another theory of dynamic behavior in the Prisoner's Dilemma that results in a stably cooperative situation in the absence of a central authority: nuclear deterrence.

B. The Classical Theory of Nuclear Deterrence

Madness and analysis may intersect more frequently in seminar rooms, where people talk about nuclear war with graphs and matrices, than in the psychiatrist's office. Nonetheless, nuclear weapons do not seem likely to go away in the next few years, and in any case my personal-defense mechanisms make me more comfortable with the game-theoretic analysis of nuclear weapons than with psychoanalysis. I undertake the neo-Faustian endeavor of analyzing nuclear weapons well aware however, of both their uniquely horrifying, non-rational aspects and of the outer limits of abstraction.

1. The Utility of the Deterrent Promise

The game theory of nuclear deterrence begins with the Prisoner's Dilemma. The reasoning is quite similar to that which we used in the USS

---

65 For the seminal game-theoretic work on deterrence, see T. Schelling, The Strategy of Conflict (1960).
Stark example in Part II, though I undertake a more extended analysis here.

If we view the nuclear world as one in which there are two decisions, "launch nuclear weapons against the opponent" and "do not launch nuclear weapons against the opponent," then the nuclear world presents a Prisoner's Dilemma. Nuclear weapons can wreak tremendous devastation on both civilian and military targets. If the United States were to launch an unanswered nuclear strike against the Soviet Union, the obliteration of the United States' principal rival would probably gain for the United States an increased ability to pursue its interests around the world. The Soviet Union would clearly suffer horribly, mostly from the damage inflicted upon its citizens but also in some small measure from the reduction in its ability to advance its interests on the international scene. If the Soviet Union were to launch an unanswered nuclear strike against the United States, the converse would apply: some gain for the Soviet Union and a tremendous loss for the United States. If each side simultaneously launches nuclear weapons against the other, then both sides suffer horrendous losses, and neither increases its international influence relative to the other. If neither side launches its nuclear weapons, then the status quo prevails, with each party far happier than if it had been the target of nuclear weapons, but less happy than if its rival had been eliminated without loss to itself. The payoff matrix is shown in Figure 14.

Comparison with Figure 8 will reveal that the situation is indeed a Prisoner's Dilemma, and one can recapitulate in this particular situation the general reasoning used to describe the incentives for both sides to adopt the non-cooperative strategy. Examine the issue from the perspective of the Soviet Union, for example. If the United States does not launch, then the Soviet Union is better off to take the opportunity to launch its weapons and wipe its greatest rival from the face of the earth. If the United States does launch, then the devastated Soviet Union does better to launch its weapons and prevent an untouched United States from achieving world dominance. Under the logic ascribed to the participants in the Prisoner's Dilemma, the Soviet Union should therefore launch. The same logic applies with respect to the decision of the United States; logically, it too should launch.

Fortunately for us, the Soviet Union apparently does not adhere to this logic. Fortunately for the Soviet Union, the United States does not appear to adhere to this logic. Furthermore, the decision not to launch has been a

---

66 See supra pp. 578-79.

67 I assume that the moral horror of the rest of the world at such a unilateral action by the United States would be insufficient to overcome its fear that the United States would treat any recalcitrant nation as it had treated the Soviet Union. I also assume that no "nuclear winter" would ensue to make the situation of the United States virtually as precarious as that of the Soviet Union. Traditional theories of nuclear deterrence have yet to evolve to take account of this possibility.
consistent one over a wide variety of constellations of nuclear weapons systems. Between the mid-1950s and the mid-1960s, for example, the United States possessed hundreds of nuclear warheads and hundreds of long-range bombers capable of carrying them to the Soviet Union, while the Soviets had only a small number of warheads and no truly suitable delivery vehicles. Nonetheless, the United States refrained from launching a first strike, even during the Cuban Missile Crisis. Today both sides possess large numbers of warheads and delivery vehicles, including submarine-launched ballistic missiles (SLBMs) and highly accurate intercontinental ballistic missiles (ICBMs). In this very different situation, neither side has come close to launching its weapons deliberately against the other.

---

The explanation, at least in game-theoretic terms, lies in the dynamics of the situation, which are not captured in the traditional, static payoff matrix. The heart of nuclear deterrence is a simple, conditional promise to the opponent: if you use nuclear weapons against me, then I use nuclear weapons against you. This promise depends upon the passage of time: if you use nuclear weapons against me, then I will use nuclear weapons against you. The game is essentially only one move long, but the strategies for that one move are chosen sequentially, not simultaneously. And that, as Robert Frost might say, has made all the difference.

I will not here discuss whether the proper analogy to the deterrent promise is a mutual suicide pact, two scorpions in a bottle, a pair of blow-gunners with slow-acting-poison darts, or something else entirely. The important point for our purposes is that each side attempts to persuade its opponent that there will never be an outcome in which the opponent launches and the persuader does not. If both sides make this promise and both sides believe the opponent's promise, then the game shifts to an either/or choice for both: either "maintain the status quo," or "kill and be killed." See Figure 15. In this situation, the rational choice is clearly to maintain the status quo. The knife-edge of nuclear terror becomes something closer to a ledge.

2. The Credibility of the Deterrent Promise

This deterrent strategy is perfectly consistent with the TIT FOR TAT meta-strategy that Axelrod found so successful in the computer tournaments of the iterated Prisoner's Dilemma that he conducted. In its general shape, the deterrent promise is the same as TIT FOR TAT. If your opponent "defected" in the "previous" move by launching a nuclear strike against you, then you "defect" in your "next" move by launching a strike against him. If your opponent "cooperated" in the "previous" move by not launching such a strike, you do not launch a strike in this move. See Figure 16.

More generally, the deterrent meta-strategy fits three of Axelrod's four general qualities of a successful meta-strategy. Deterrence is "nice" in Axelrod's terms because a nation pledges not to launch first. The two sides maintain stability—cooperation—through the threat of a second strike described as "retaliatory" in the argot of both Axelrod and nuclear strategists. The deterrent promise is explicit and simple, and therefore "clear." Whether it is "forgiving" is essentially irrelevant, however, in light of what will happen to the world after the first "defection"; the theory of deterrence in fact provides no useful answer for any who survive the first exchange and wonder, "What next?"

The enormous damage that would result from a nuclear war argues against waiting for the two sides to signal to one another their meta-strategy of deterrence through a series of iterations of peace and nuclear war. Despite the attractiveness of TIT FOR TAT in general terms, therefore, the analysis of nuclear deterrence focuses not only on the dynamics of the strategy, but
also on efforts to convince one's opponent in advance that the TIT FOR TAT of deterrence is in fact one's meta-strategy.

The strategist's code word for this hyper-clarity is "credibility." One might view the credibility of the deterrent promise as having two components, one psychological and one physical. The psychological credibility of the deterrent promise rests on the belief that a nation that has been devastated by nuclear attack will want to respond by devastating the other nation. The physical credibility of the deterrent promise depends on the likelihood that a nation devastated by nuclear attack can respond by devastating the other nation.

I can offer little insight into the psychological credibility of the deterrent promise. The thought of responding to nuclear explosions upon one's soil by visiting nuclear devastation upon the originator's soil strikes many as point-
The Deterrent Promise

(Mutual)

<table>
<thead>
<tr>
<th>Does Either Side Launch?</th>
<th>Who Launches First?</th>
<th>Does the Other Side Launch Back?</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>Yes</td>
<td>(-800, -800)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>No</td>
<td>(100, -1000)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>Yes</td>
<td>(-800, -800)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>No</td>
<td>(-1000, 100)</td>
<td></td>
</tr>
<tr>
<td>(0, 0)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

FIGURE 16

less at best and despicably immoral at worst. Nonetheless, the modern historian could note that civilized nations have recently thirsted for revenge when faced with insults far less dramatic than nuclear war. For example, Britain went to war with Argentina over the latter’s occupation of tiny islands no closer to London than is Alaska. The United States, which we like to think of as naturally reticent to go to war, had little difficulty doing so after the day of infamy at Pearl Harbor, or in going to war over the sinking of a few American ships in World War I. A nuclear attack upon one’s homeland would seem a more justified ground for vengeance.

In any case, it is the physical credibility of the deterrent promise that has received the most attention from the seminar-room set. This aspect of credibility also lends itself more naturally to an analysis of discovery abuse. I therefore treat the physical credibility of deterrence in some depth.

The physical credibility of the deterrent promise rests upon the ability of a side that has been the subject of nuclear attack to launch its own weapons in response. If it can do so, it possesses a “secure second-strike capability.” If it cannot, it must possess either a willingness to launch first or an excellent ability to bluff.

During the nuclear era, the common wisdom has been that there would be some measurable military advantage in launching one’s nuclear weapons.
Bilateral 1st-Strike Capabilities

<table>
<thead>
<tr>
<th>Does Either Side Launch?</th>
<th>Who Launches First?</th>
<th>Does the Other Side Launch Back?</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>US</td>
<td>No</td>
<td>(100, -1000)</td>
</tr>
<tr>
<td></td>
<td>USSR</td>
<td>No</td>
<td>(-1000, 100)</td>
</tr>
<tr>
<td>No</td>
<td></td>
<td></td>
<td>(0, 0)</td>
</tr>
</tbody>
</table>

FIGURE 17

before the other side (a "first strike") rather than waiting until the warheads of your opponent had landed on your territory and then launching a counterstrike (a "second strike"). The common wisdom has also been that such an advantage would hardly be sufficient to prevent a devastating second strike from one's opponent.

This situation is stable, but bears close watching. If the common wisdom were incorrect (or the strategic balance changed sufficiently), then a disarming first strike might become possible—one in which a first strike in fact incapacitated the attacked side and prevented a second strike. See Figure 17. This would be a dangerous situation indeed, because a side that can launch a disarming first strike could ignore the other side's deterrent promise if it actually launched a first strike. The Prisoner's Dilemma then becomes something much less stable. The situation becomes especially unstable if both sides can conduct a disarming first strike, since each then not only has the incentive to strike first and reap the resulting benefits, but also fears losing this advantage by waiting too long and falling prey to the enemy's disarming first strike. See Figure 18.

Because of the importance of having a secure second-strike capability, and because the revisionist theory of escalation discussed in the next subsection builds upon the classical theory of nuclear deterrence, I will discuss at some length the factors that go into the benefits from a first strike and the
odds that a second strike will be sufficiently probable to pose a physically credible threat as part of the deterrent promise.

Striking first in a nuclear war is likely to provide several advantages in comparison with striking second. If the side striking first can destroy more than one nuclear warhead of its opponent with each of the nuclear warheads that it launched, then the side striking first would be ahead of the game. In fact, a significant number of nuclear delivery systems have typically been so geographically concentrated that a single nuclear delivery vehicle could destroy several of them. This has generally been true of strategic bombers, which have not been placed in super-hardened shelters and which fly quite slowly relative to incoming weapons. In addition, the deployment in the 1970s of thousands of nuclear missiles each carrying several warheads

---

69 Such weapons are denominated "MIRVs," an acronym for multiple independently-targetable re-entry vehicles.
meant that each side could launch a single “delivery system” and destroy several of the multi-warhead weapons of the other side, thereby producing a doubly effective impact on the other side’s weapons.

Wholly apart from the weapons systems that might be destroyed in a first strike, top civilian leaders are generally both much more vulnerable to, and much fewer in number than, nuclear weapons. Thus, a first strike might be able to paralyze completely the forces of the other side by killing the top political leaders of a nation70 or by disabling communications systems for the transmission of their orders.71

The side striking first is also likely to achieve, at least in part, some military advantages that traditionally adhere to the party initiating an attack. One of these is surprise. While the early-warning systems of the attacked side are likely to prevent complete surprise, the psychological shock of being under nuclear attack may nonetheless generate the confusion or paralysis that often accompanies ignorance in military systems. Another advantage of initiating an attack is the ability to pre-plan one’s own attack while forcing one’s opponent to improvise its response with whatever weapons and personnel survive the first wave of attacks. In addition, there are a number of technical difficulties in carrying out a retaliation in what defense analysts, scaling the heights of euphemism to their peak, call a “nuclear environment.” Communications systems may malfunction, for example, or weapons may have to fly through huge clouds of dust.

Striking first in a nuclear war, then, is likely to convey a host of advantages in comparison with striking second. Any party that had actually decided to fight a nuclear war would therefore seem likely to cede to the opponent whatever the opponent might learn from watching the arc of the attacker’s missile tracts terminate on his territory. Nonetheless, for several reasons, one should not necessarily fear for the physical credibility of the deterrent promise. Pre-planned attacks sometimes fail to achieve the complete destruction of their targets. At Pearl Harbor, for example, the Japanese achieved complete strategic surprise, but the U.S. aircraft carriers that were a crucial target of the Japanese attack aircraft were out of port through a stroke of luck. In addition, many nuclear weapons systems have over time become much more difficult to destroy—more “survivable,” in the vernacular. ICBMs, for example, initially sat above the ground and required elaborate preparations before they could be launched. These characteristics made

---

70 A nuclear attack that seeks to destroy those who possess launch authority, rather than the weapons that they have authority to launch, is known as a “decapitation attack.” See P. Bracken, supra note 68, at 232-37 (discussing factors involved in launching a successful decapitation attack); Steinbruner, Nuclear Decapitation, 45 Foreign Policy 16 (Winter 1981-82) (analyzing the United States’ ability to react to a decapitation attack).

71 This type of attack, though closely related to a decapitation attack, is most properly termed a “command-and-control” attack.
them quite vulnerable to attack by the other side. ICBMs now sit in holes of super-hardened concrete and can now be launched within minutes of a decision to do so. Submarines, which are difficult to detect underwater, began to carry long-range nuclear missiles in the 1960s; they have assumed much of the role formerly played by weapons aboard more vulnerable aircraft carriers. Bombers can now carry long-range, hard-to-detect cruise missiles rather than lumbering all the way to the targets themselves. Finally, one must remember that the costs of failing to destroy even a single weapon of one's opponent can be extremely high. Allowing a handful of Japanese soldiers to escape the attentions of U.S. soldiers after the end of World War II in the Pacific cost a handful of lives; allowing a single Soviet ICBM to escape destruction in a U.S. first strike could cost the nation New York City.

The balance between the gains from a first strike and the costs from the resulting second strike therefore clearly tip in favor of restraint at present. Nonetheless, one must always be vigilant in ensuring that both sides feel comfortable in their ability to strike second, lest the uncomfortable side be tempted to launch a disarming first strike and thereby render the relative peacefulness of the nuclear age to date nothing more than a prelude to unprecedented mass destruction.

C. The Revisionist Theory of Nuclear Escalation

In the current world of a secure second-strike capability, deterrence theorists can rest relatively easy. What I will call the "revisionist theory of nuclear escalation," in contrast, draws upon some of the specifics of the tension between first-strike and second-strike capabilities to argue that the military establishments of the superpowers have, with the consent of their political leaders, created a world in which the nuclear beast might break the chains of deterrent logic. This school of thought, led by Paul Bracken, asserts that the superpowers' tightly coupled intelligence-and-alerting systems could, through their interaction with one another and the pressures to delegate launch authority, lead to a nuclear war—even if the top political leadership of the United States and the Soviet Union did not enter the underlying crisis with any desire to engage in nuclear war. I first discuss the historical developments that have allegedly led to this situation, and then recount the chain of speculations asserting that this situation could lead to nuclear war in a crisis.

1. A Volume of History

A person of average means born in 1910 in the United States would have been transported about in her infancy by a horse-and-buggy that trotted from town to town at something like 10 miles an hour. By her 40th birthday, she could see the U.S.A. in her Chevrolet at something like 50 miles an hour. By the time she was eligible for Social Security, she could take a plane on the same journey at over 500 miles per hour.
A similar revolution in the delivery of nuclear weapons has occurred during their comparatively short lifespan. In 1945, propeller bombers trundled along at about 300 miles an hour. Jet bombers could travel at roughly twice that speed. The development of land-based ballistic missiles in the late 1950s and early 1960s led to a much greater compression of decision times: flight times from the United States to the Soviet Union, or vice versa, dropped from several hours to about thirty minutes. The further development of submarine-launched ballistic missiles shortened the flight time to less than fifteen minutes. Indeed, cruise missiles may now be small and stealthy enough to be fired from a submarine off the coast of an enemy superpower and arrive in the enemy’s capital with virtually no warning at all.

The superpowers have tried to meet this revolution in weapons “transpor-tation” by improving the scope and sensitivity of their early-warning systems through the use of another oft-discussed technological revolution of the 20th century: computing power. Both superpowers have constructed far-flung early-warning networks that tie together satellites, radars, and “data fusion centers”—throbbing, non-stop conventions of computers—to search for incoming nuclear delivery vehicles. This extraordinary effort to ensure “tactical warning” (an indication that weapons are actually in the air) has been paralleled by efforts to improve “strategic warning” (an indication that the opponent is moving towards actually launching weapons into the air). Spy satellites, spy planes, and spy trawlers prowl the globe looking for information about the other side on the ground or in the ether, assisted by cryptologists and roomfuls of intelligence analysts.

The superpowers have taken steps not only to see that they know when something is brewing within the nuclear forces of the other side, but also to respond to such situations. Indeed, the superpowers have constructed a near-real-time “vertical integration” of their warning and weapons systems that allows rapid protective responses by their nuclear forces, thereby reducing the chances that the enemy could succeed in a disarming first strike while catching them with their nuclear-force pants down.

These measures, which I will describe shortly, contrast significantly with the alertness of nuclear forces in their normal peacetime posture. Despite what one might have inferred from various films of the 1950s, even the Strategic Air Command does not constantly keep every one of its airplanes tautly poised to strike back at the enemy aggressor. In fact, the Strategic Air Command no longer keeps any weapons-carrying planes in the air, and generally even refrains from keeping many planes on any meaningful alert at

---

72 P. BRACKEN, supra note 68, at 6.
73 Id. at 26-28, 38-39.
74 For an overview of Soviet early-warning intelligence gathering techniques, see id. at 46-47.
75 See id. at 18-22, 31, 55.
76 Id. at 205.
all. Similarly, a great number of ballistic-missile submarines spend any given day in port rather than hiding and patrolling in the open ocean. To keep both the weapons and the warning system constantly ready is, with the exception of the ICBM force, simply considered too expensive even for those who are willing to spend hundreds of dollars for a ten-dollar claw hammer.

Instead, the United States, and presumably the Soviet Union, has developed an elaborate system of alerts that it can use to increase the readiness of its forces if it receives strategic or tactical warning that something is up. The United States uses the DEFCON (Defense Condition) system. Ballistic-missile submarines in port can prepare for departure at DEFCON 3, and then actually leave when ordered to even higher alert at DEFCON 2.77 Submarines at sea are much more difficult to find, and thus to destroy, than submarines in port; "surging" the submarines from port to sea increases the number of submarines that are relatively invulnerable, and therefore presumably discourages an adversary's first strike. The Strategic Air Command might raise the number of bombers on "strip alert" if it moved to DEFCON 3, as it did for twenty-four hours during the 1973 Yom Kippur crisis.1 This would increase the rapidity with which such planes could flee their bases if the early-warning system detected incoming strikes against them, and, as with surging the ballistic-missile submarines, would presumably diminish the incentives for the Soviet Union to conduct a first strike. The military could also bolster the early-warning system itself, by shifting any dual-purpose surveillance systems into their early-warning mode.79

The Soviets might undertake similar measures in a serious crisis—though to date, and in contrast to the United States, the Soviets have never alerted their nuclear forces beyond normal peacetime readiness. The Soviet Union keeps its bombers at a very low state of alert in peacetime,80 and it generally keeps a higher percentage of its ballistic missile submarines in port than does the United States.81 Clearly, Soviet military leaders might view this relaxed state of readiness as leaving the Soviet Union vulnerable to an American first strike. In an urgent crisis, these military leaders might convince their political masters to take steps to increase the readiness of the Soviet nuclear forces as a means of discouraging any American contemplation of a first strike.82

78 Id. at 88.
79 Id. at 84.
80 P. BRACKEN, supra note 68, at 63.
81 Bracken states that 80% of Soviet nuclear submarines are routinely in port, compared to only 30% for the United States. Id.
82 Indeed, the Soviet Union might even go to what might be called the "ultimate alert" level—a "launch under attack" posture, in which strategic weapons would be launched as soon as the Soviets received tactical warning of a U.S. attack. This posture would certainly reduce the attacker's expected gains from a first strike for the
In addition to improving the second-strike capability of their weapons systems, the superpowers might also take measures to insure the integrity of their command-and-control systems. The airplanes that would serve as airborne command posts can be placed in a higher state of readiness. The military could fully staff ground-based communications facilities and various military commanders could fan out from their usual peacetime locations. More importantly, the highest civilian authorities might delegate the authority to launch nuclear weapons, either to their civilian deputies or to officers in the military.

This is an especially important—and sensitive—point. Suppose, for example, that only the President of the United States can, both legally and physically, authorize the release of nuclear weapons in peacetime. Suppose further that the President considered during a grave crisis that the Soviet Union might actually contemplate launching a nuclear attack on the United States. If in such a situation the President were incinerated in a nuclear blast, or even simply incapacitated by a rifle bullet, the United States would be temporarily incapable of responding to a Soviet nuclear attack. Indeed, in a nuclear attack on Washington, all the logical choices for the President’s successor—or at least all those with whom it might be possible to establish communications—could perish as well. The incapacity might therefore last longer than the time needed simply to transfer the Presidency one link down the chain, and in the nuclear era any delay might cause the unanswered destruction of a nation.

As a means of countering any Soviet incentive to strike first in hopes of destroying the American command structure, the President could in a grave crisis provide top-ranking military officers with the physical means to authorize the release of nuclear weapons. Such officers might be authorized to launch if, for example, they were to obtain confirmed evidence of a nuclear attack on Washington and the President could not be reached for twenty-four hours. As with the measures to protect nuclear weapons systems from a first strike, such a delegation of authority would increase the chances that the United States could respond decisively to a nuclear attack, thereby presumably reducing any incentives the Soviet Union might have to launch a weapons that were the targets of that first strike may already be in the air on their way to the attacker’s homeland. The first-strike missile would end up destroying only empty missile silos. The danger, of course, is that the side adopting a launch-under-attack posture will misread its early-warning systems and start World War III by mistake. In addition to serving the second-strike capability of their weapons systems, the superpowers might also take measures to insure the integrity of their command-and-control systems.

For an extended discussion of a significant delegation that many assert has occurred, see id. at 197-200, 224-30.
first strike.\textsuperscript{84} Such measures also, of course, risk that the President will lose control of the geographically dispersed nuclear forces.\textsuperscript{85}

2. A Page of Logic

The alerting and delegations that I have described above all have a "defensive" purpose: the party undertaking those actions seeks to reduce the chances that an attacker could launch a successful first strike by reducing the gains from such an attack. Under the classical theory of nuclear deterrence, such actions are therefore desirable. But because of the irreconcilable ambiguity of any action taken to reduce the risk to one's own forces and the extent to which the superpowers keep an electronic eye on one another, the revisionist theory of nuclear escalation holds that these "defensive" moves in a crisis are potentially quite dangerous. Especially in the heat of a crisis, it may be virtually impossible to distinguish a defensive action taken by your opponent to reduce the chances that you will launch a successful first strike against it from an offensive action meant to improve its own chances of launching a successful first strike against you.\textsuperscript{86} Suppose, for example, that the United States were to disperse and alert its strategic bombers. As discussed above, such a move serves a defensive purpose: the temptation to attack pre-emptively is presumably lessened by the reduced gains that the pre-emptive attacker could expect by striking at more vigilant, more widely dispersed forces. But such a move would also serve an offensive purpose: the more dispersed, more vigilant forces are now that much closer to being able to take off on offensive missions. Moreover, the forces protected by the move would be more likely to survive any Soviet strike taken in response to a U.S. first strike (assuming of course that some of the forces alerted and dispersed did not participate in the initial wave of the U.S. attack).

This issue is particularly worrisome because of what Bracken calls the "tight coupling" of the two sides' warning-and-alert systems.\textsuperscript{87} The Soviet Union keeps close tabs on the United States, and vice versa. Any significant military action that the United States takes is likely to be detected by the

\textsuperscript{84} See \textit{id.} at 202-04. For example, the Commander-in-Chief of the Strategic Air Command is reported to have the authority to send the strategic bombers into the air if he believes that the Soviet Union has actually launched a nuclear attack. \textit{Id.} at 69.

\textsuperscript{85} I have here expressly discussed only the factors involved in delegation with respect to strategic nuclear war. It is important to remember that nuclear weapons are also currently deployed in Europe. These weapons are likely to be close to the front lines even in a conventional conflict with the Soviet Union and its Warsaw Pact allies. For a variety of reasons, there might be tremendous pressure to delegate authority for their use to relatively low levels of the military. See \textit{id.} at 129 (discussing the special problems associated with nuclear war in Europe).

\textsuperscript{86} \textit{Id.} at 64.

\textsuperscript{87} The following two paragraphs draw heavily upon P. BRACKEN, \textit{supra} note 68, at 54-65.
Soviet Union—indeed, the implicit purpose of many strategic moves during a crisis may be to show U.S. firmness and resolve by making sure that the Soviet Union does see what the United States does. When the Soviet Union detects that action, it could well infer movement towards a first strike by the U.S. and take some roughly equivalent action that the United States would in turn detect and interpret as increasing the threat of a Soviet first strike. The frightening prophecies of the two sides could become self-fulfilling if the interactions of the two sides spiral out of control and into nuclear war after leaders delegate authority to launch, or if malfunctions of early-warning systems that are routinely dismissed in peacetime are taken seriously in the caffeine fumes of crisis.

This revisionist theory of nuclear escalation does not take an expressly game-theoretical perspective, but it is nonetheless closely related to a problem with an iterated Prisoner’s Dilemma that Axelrod acknowledges. With a TIT FOR TAT meta-strategy, a move by one side that the other perceives as uncooperative can lead to an endless cycle of non-cooperative moves, first by one side and then by the other. Suppose that Players 1 and 2 both employ a TIT FOR TAT meta-strategy. Player 1 cooperates, but because of a mistaken reading of Player 1’s move as uncooperative, Player 2 defects on the next move. Angered, Player 1 retaliates on the next move, as the dictates of the TIT FOR TAT meta-strategy demand. Because Player 2 has a memory of only one move back under TIT FOR TAT, it then adopts an uncooperative strategy in response to Player 1’s retaliation. And, as Kurt Vonnegut said, so it goes. The players are now rigidly locked in a game of brinksmanship with one another, at least as long as they both adhere to the TIT FOR TAT meta-strategy.

The revisionist theory of nuclear escalation similarly fears that an echoing cycle of higher alert levels, eventually leading to nuclear war, will result because each party perceives the others’ strategy as uncooperative even though each intends its own move to be cooperative (or at least no worse than precautionary). Dampening the cycles can be difficult in the stress of crisis, especially in light of the asymmetrical vulnerability that either side would suffer if it were first to reduce its alert level.

V. Discovery Abuse as a Dynamic Game

Having set forth these theories of cooperation, deterrence, and escalation, I will now apply them to discovery abuse. The theory of cooperation is useful in providing a number of reasons why discovery abuse seems likely to

---

88 Bracken also discusses the possibility that such an interlocking spiral of alerts could lead to conventional anti-submarine warfare against nuclear-missile submarines, or the dispersal of NATO’s nuclear weapons, thereby leading to further escalation and eventually to nuclear war. Id. at 64-65.

89 P. Bracken, supra note 68, at 224.
have become worse over the past few decades. The classical theory of nuclear deterrence illuminates the effect that asymmetries in the availability of resources and opportunities for discovery have on the potential for discovery abuse. The revisionist theory of nuclear escalation provides some insight into the intractability of discovery abuse. Having examined these theories in the context of discovery abuse, I offer some suggestions for reform.

My method is fairly simple. Within the framework of each of the three theories discussed above, I set forth below a number of stylized statements about the current or future shape of the practice of law and of discovery. I do not assert that the list is exhaustive, but I have tried to capture the essence of most of the commonly held beliefs about modern litigation and its practice. And, despite the informality with which they are rendered, I have also striven mightily to ensure that each statement stands independently of the others. I cannot promise that the statements universally reflect reality, of course, but I have heard something like them uttered by more than one practitioner. After each statement, I discuss its implications for discovery abuse in light of the relevant theory.

A. Discovery Abuse and the Theory of Cooperation

Before I begin the treatment of various specific statements and their possible relationship to the theory of cooperation, I should note that Axelrod's work has one fairly clear implication for attorneys enmeshed in discovery: use a TIT FOR TAT meta-strategy. While doing so does not necessarily guarantee a litigant's success, TIT FOR TAT's carrot of cooperation and stick of retaliation is certainly likely to increase a litigant's chances of a favorable outcome. Moreover, other attorneys might recognize the increased benefits flowing to both parties from a mutual adoption of a TIT FOR TAT meta-strategy in a particular litigation, and adopt a similar strategy themselves. These continually cooperative attorneys, even if a small percentage of the local bar, can encourage the entire local bar to alter its discovery practices. Cooperation is unlikely to replace abuse overnight—and I discuss in the next subsections some reasons why cooperation is unlikely ever to replace abuse completely—but every little bit of cooperation helps.

"The courts are so jammed up that the cases take forever to resolve."

An increase in the time between the filing and expected resolution of the average suit is likely to increase discovery abuse. Part of the shading of the

---

90 As a result, no single statement attempts to cover too many factors at once. Readers may therefore wish to skim through the list of statements to see what I have included.
"shadow of the future," after all, comes from the size of the discounted value of future payoffs. The benefits of cooperation accrue largely in the future; an uncooperative party living in a world with a high discount rate can make a quick killing by exploiting his opponent and then living with the heavily discounted retaliation. An action brought in an overcrowded court presents a similar situation. For a suit of a given level of complexity, delay in the courts means that the case unfolds in slow motion. The benefits from cooperation will accrue whenever the parties (or the court, if one of the parties seeks judicial intervention) finally get around to the next discovery request, and these heavily discounted future benefits will therefore be relatively small. This is not because the discount rate itself is particularly high, but rather because the benefits are particularly far off in the future. The current request, in contrast, presents the opportunity for a large, undiscounted, immediate gain if one is fortunate enough to choose the uncooperative strategy while one's opponent chooses the cooperative strategy. The retaliation that such a move may provoke, in contrast, will result in a loss of benefits only in the distant future. Après l'abuse, la déluge.

"There are a heckuva lot of lawyers out there."

The widely publicized proliferation of attorneys should also concern those observing discovery abuse. As the population with which one interacts becomes larger, the frequency with which one interacts with any particular colleague diminishes. The shadow of the future, in its second meaning, thereby becomes less important. With a diminished likelihood of interacting with the same attorney again, a cooperative strategy—viewed this time with respect to repeated interactions in future cases rather than with respect to multiple discovery requests in the same case—becomes less attractive. When the probability of future interaction decreases, so do the benefits from cooperation. Abuse is therefore likely to swell as does the bar.

"The firms just keep getting bigger."

If attorneys believe that a particular attorney's firm has some independent value in predicting the attorney's behavior, then larger firms may result in more abuse—even if attorneys stay with a particular firm for a constant length of time or the number of attorneys admitted to the bar remains

---

91 I assume here that the reason for increasing delays is a result only of increases in the numbers of actions filed. The increase in the time necessary to resolve a case may stem not simply from an increasing number of cases, but also from an increasing complexity of cases. I treat complexity separately. See infra text accompanying note 98.

92 Though "firm" is a word generally used in conjunction with private partnerships, I do not mean by using it to exclude attorneys in public practice.

93 See infra pp. 621-22.
constant. The arguments necessary to reach this conclusion are neither simple nor unassailable, however, and the conclusion must therefore remain suspect. I address the question nonetheless because an increase in firm size appears to be one of the important trends of the past few decades in legal practice.

If one attempts to force the issue of firm size to fit into the theory of cooperation, one probably does best to treat it as a matter of how precisely one can identify one's opponent. Axelrod's tournaments assumed that parties can unfailingly identify one another when they encounter each other, yet Axelrod noted that the fidelity with which individuals can identify one another is a real-life factor in the success of cooperative strategies.\(^9\) The more reliable the identification, the more confident a player can be that his cooperation will be met with cooperation (if previous interactions have been cooperative) or that he should be wary and defensive (if previous interactions have been uncooperative). Identifying an individual as part of a law firm might provide a useful external cue to that individual's behavior in discovery. For whatever reasons, firms have a reputation for "playing hardball" or "being professionals." To the extent that such reputations translate into reliable guesses about the discovery behavior of a firm's attorneys, these firm identities can be substituted in part for actual interactions between litigating opponents as a means of predicting an opponent's behavior during discovery. For example, an attorney who knew nothing about her opponent except that his firm had a reputation for being aggressive and contentious would be wise to expect little cooperation. Firms would have to live with their reputations, and only those firms that could cultivate a reputation for cooperativeness would be able to share reliably in the resulting mutual gains.

To go from a belief that firms are a useful partial predictor of an attorney's behavior to the belief that interactions with larger firms increase discovery abuse, however, requires two more assumptions: (1) larger firms generally have fuzzier identities; and (2) the negative impact of such fuzziness on one's ability to discern a law firm's discovery behavior outweighs the positive impact flowing from the need to keep track of fewer firms' identities.

The increase of fuzziness in a firm's image with an increase in its size seems likely to be true on average. Large corporations have typically gone through an "entrepreneurial" phase in which their identity was clearly linked to their founder (e.g. Henry Ford or H. Ross Perot) before reaching their current, "managerial" phase (e.g. Ford Motor Company or Electronic Data Systems), in which their success depends more upon adapting to the problems that face all large corporations. Similarly, the identity of law firms is often closely associated at first with its named partners until, as the founding partners retire or die, its identity becomes less sharply drawn.

\(^9\) R. Axelrod, supra note 8, at 20.
There exists an even more significant reason for law firms’ practices becoming more and more similar as they expand in size and scope: they must compete at the national level not only for business but also for new recruits. While some firms attempt to make virtue out of necessity by trumpeting their “diversity,” the important lesson for present purposes is that large organizations tend to come to resemble one another.

This fuzziness in the identity of the large firm presumably carries over to its discovery practice as well. Whatever differences there once were from firm to firm, the differences shrink with the growth in firm size. As firms become poorer relative predictors of behavior, the likelihood of cooperation in discovery affairs shrinks with the difficulty of identifying one’s opponent precisely. The two attorneys (or teams of attorneys) cannot rely upon firm identity as a substitute for previous interactions, and must therefore begin the game as if they had never interacted before. An attorney with a TIT FOR TAT meta-strategy must thus convince her opponent of that fact through a series of interactions—which may result in unproductive retaliations, and could even lead to echoing uncooperativeness—rather than simply “beginning” with the knowledge in her opponent’s mind that cooperation will in the long run profit that opponent, and therefore begin with a stable, cooperative relationship. Old friends do not need to test each other’s loyalty every time they meet. (Nor do old enemies, however. Yet another caveat to this line of argument, therefore, is the possibility that attorneys who must begin the discovery process with poorly-defined identities, as if they had never interacted before, will be more likely to obtain cooperation because they can bury an uncooperative past and begin anew).

The final assertion in the argument that interactions with larger firms result in more discovery abuse is the that losses from fuzziness of identity outweigh the gains from having fewer firms with which to interact. One can isolate the favorable impact of this latter factor by comparing a legal market with ten 10-person firms to a legal market with two 10-person firms. In the second situation, with fewer firms with whom to interact, the probability of repetitive interaction increases when compared to the first situation. With an increased probability of interaction, the shadow of the future lengthens, thus increasing the incentives for cooperation. If the growth in firm size also meant a reduction in the number of firms, then there would be some positive effects on discovery. Unfortunately, I cannot offer any reliable guidance as to whether this positive increased-interaction factor in fact outweighs the negative fuzzy-identity factor, and so I must admit that I cannot be certain that larger firms necessarily lead to more discovery abuse.

“We at Dewey, Screwham have a truly national practice.”

The national scope of legal practice in the United States has probably increased the amount of discovery abuse over the amount of abuse that would occur if legal practice were geographically parochial. A Boston law-
yer who battles another Boston lawyer has a relatively high probability of encountering that opponent again in future case in the city's courts. Since the shadow of the future is therefore relatively dark, the cooperative solution to the discovery problem—for both sides not to abuse discovery—becomes relatively attractive. In contrast, a Boston lawyer who clashes with an Atlanta lawyer in the Peachtree State may well never oppose that attorney again. Since the shadow of the future is difficult to see, the incentives for both sides run in favor of abuse.

One should note, however, that a national practice is not inevitably bad. Suppose that only one lawyer in the country handles personal injury cases that involve limb injuries which result from using a Brand X lawnmower as a hedge trimmer, and that Brand X company's insurer always uses the same lawyer to defend against such suits. Both lawyers will then have a national practice, but the probability of their future interaction is extremely high; thus, the incentives to refrain from discovery abuse in this situation remain strong.

"If you're smart, you'll stay at a big firm for a couple of years to polish your résumé and then take off before they make you work there for the rest of your life."

A decrease in the time that attorneys spend with a given firm is roughly analogous to an increase in the national character of a firm's practice. From the point of view of both the firm-hopper and the opponent, their future interactions become less likely to whatever degree affiliation with a firm (or residence in a geographical area) is a valid predictor of what sort of cases one will handle or the region in which they will be handled in the future.

I note also that an exodus from litigation practice as a whole, as well as migration from firm to firm, may exacerbate the problem of discovery abuse. Some of my law-school classmates have decided that one can live as nobly in other professions as in the litigator's, and have taken up another vocation. If this represents a general trend, then those remaining in the rough-and-tumble world of appearing before judicial officials will know, when considering any given action with respect to discovery, that they are less likely to encounter their current opponent in the future than if they both were certain to continue in a litigation-oriented practice. Once more, the shadow of the future dims and the incentives for cooperation diminish.

One must qualify this conclusion with a caveat similar to that which I noted in discussing firm size, however. With an exodus from litigation practice, the pool of practicing lawyers shrinks. This increases the likelihood of future interactions with those remaining in the pool, and thereby increases the incentives for cooperation. If one could identify those likely to leave from those likely to stay, one could treat them differently: cooperate with those destined to become Of Counsel after a lifetime of service to a firm, and abuse those destined to leave the legal market altogether. Otherwise, how-
ever, it is difficult to judge the precise impact upon abuse which increased
exits from the litigation market has. I note, however, that as long as new
law-school graduates pour into the marketplace at least as quickly as the
aging law-school graduates pour out, the size of the overall pool will not
decrease and the likelihood of future interactions with the average attorney
therefore will decrease, thus, resulting in a combination of circumstances
that leads unambiguously to less cooperation.

"Patents and copyrights? They've got their own little world over there."

Substantive specialization within the practice of law is likely to lessen the
problem of discovery abuse. This phenomenon is the functional converse of
several of the developments that I have discussed above. If patent lawyers
are forever butting heads against other patent lawyers, rather than against
other lawyers generally, then the pool of individuals with whom patent
lawyers interact is small. The probability of interacting with any individual in
the pool is therefore large, just as wolves who roam in a particular pack
interact frequently with those in the pack. With that increase in the probabil-
ity of interaction comes a lengthening of the shadow of the future and a
concomitant increase in the likely rewards of refusing to engage in discovery
abuse. If patent lawyers are rational, then they will engage in less discovery
abuse in this situation and save their venom for absent-minded inventors
who forget to pay their legal fees.

"In the newly competitive law-firm environment of the late 1980s, some
large but unnimble firms may actually go under."

The reality of actual competition in the market for legal services is bad,
not only because of the indignities involved in watching a lawyer's ad follow
a used-car salesman's ad on television, but also because discovery abuse is
likely to increase. This increase in abuse would not, however, stem from the
fact that law firms will compete against each other. The model of discovery
abuse that I have set out already assumes that the lawyers accurately
represent the interests of their clients. The increased effort to please the
client that might come with more competition among law firms is already
taken into account.

Rather, the likely increase in abuse stems from the fact that the possibility
that one's opponent (or oneself) might go out of business decreases the
chances of future interaction between the two firms. If firm-to-firm identities
are a significant part of predicting an opponent's behavior, then the possibil-
ity that some firms will exit the market due to a business failure will thus
shorten the shadow of the future and increase the incentives for each
participant to take advantage of the other.
"See that stack of papers on the table next to the signed picture from Lyndon Johnson? That's 500 g—damned pages of discovery requests."

Large discovery requests are bad, regardless of the overall level of complexity in the case. Let us assume, for analytical purposes, that every case generates, in some Platonic sense, a particular number of questions that can be asked with the tools of discovery. One can parcel out those questions in either a small number of chokingly large requests or a large number of bite-sized requests. From the point of view of a believer in the theory of cooperation, a small number of large requests is undesirable. Decreasing the number of interactions between the parties, while holding constant the stakes, always decreases the incentives for cooperation in an iterated Prisoner's Dilemma.

Under the Federal Rules of Civil Procedure, for example, each side can tender all of its discovery requests in a single document. The outcome of that game, as we should recall from the Prisoner's Dilemma, would almost invariably be the mutually uncooperative outcome. Having played his hand of discovery requests, the requesting party has little to offer the responding party in terms of future cooperation. The responding party thus has little to lose by recalcitrance or retaliation; aware of this, the requesting party may as well be abusive from the start.

The situation is similar, though less stark, for any small number of interactions. A large number of small interactions, in contrast, provides ample opportunity to clearly demonstrate that one is retaliatory but forgiving. A wise attorney opposing an attorney with a multiple-interactions strategy will, as noted earlier, find it in her client's best interest to adopt a similarly cooperative strategy herself.

"The cases keep getting more complicated."

Considered from the theory of cooperation point of view, the effect on discovery abuse of increasing complexity is difficult to determine. The advantage of complex cases, from this perspective is that they are likely to involve more discovery, thus providing a greater opportunity for the kind of ongoing interactions in which a cooperative relationship can evolve. After all, the parties in a complex case—at least one in which each side has a significant opportunity to employ discovery against the other—know that the case is likely to involve a good deal of lawyering on both sides. An

---

95 The Rules do not contain any substantive limits on the way in which a discovery request may be tendered. See Fed. R. Civ. P. 26.
96 See supra pp. 575-79.
97 See R. Axelrod, supra note 8, at 129-32.
98 In contrast, considered from the point of view of the theory of escalation, increasingly complex litigation is likely to exacerbate problems of discovery abuse. See infra text accompanying notes 104-05.
attorney can therefore adopt the TIT FOR TAT meta-strategy confident that she will have a welter of sequential opportunities to choose one strategy or another, and thereby have the opportunity to demonstrate clearly to her opponent that her strategy is forgiving but retaliatory.

Unfortunately, if the welter of opportunities occurs over a longer period of time (as they typically do in complex cases) than they do in simple interactions, then the complexity of a case may also have some negative effects upon discovery. As discussed above, a case that takes a long time to resolve presents the parties with a relatively dim shadow of the future, and the incentives for cooperation diminish concomitantly. The net effect of these two factors related to complex cases—more interactions on the one hand and interactions occurring over a longer period of time on the other—is therefore indeterminate.

B. Discovery Abuse and the Theory of Nuclear Deterrence

The theory of deterrence provides a relatively simple way to conceive of the effects of asymmetries in the parties' resources upon discovery abuse.

"This is just a couple of mom-and-pop companies going at it. Don't go billing 200 hours on it, okay?"

Discovery abuse is less likely when the two parties have limited resources to spend on litigation. The acts of filing and answering discovery requests consume resources. In many cases, of course, the resources necessary to tender and respond to all impositionally cost-effective discovery requests are only a small fraction of the total resources available to the parties for litigation. In other cases, however, one or both of the parties involved may find that the resources needed to make and respond to all impositionally cost-effective discovery requests represent a significant fraction of their total available resources. When both parties to a suit have limited resources to spend in discovery, the situation is roughly akin to the model of stable nuclear deterrence discussed in section B of Part IV.

Let us assume for purposes of simplicity that both sides have identical

---

99 A more complex case need not inevitably take more time to complete, of course. Calendar time is the important factor in determining the discount rate, not billing hours on one's timesheet. A team of attorneys who work twenty hours a day on a complex case might finish it before a single attorney working a mere eight hours a day on a simple, less time-consumng case finishes hers.

I similarly ignore the fact that a more "complex" case is not always a case with a greater number of discovery opportunities. A case might be legally complex but factually simple. If such a case took a longer time to complete because of its complexity, the fact that it offered no increased opportunity for interactions in discovery would mean that the incentives for abuse unambiguously increased in comparison to a simpler, less time-consuming case.
resources and identical amounts of impositionally cost-effective discovery requests that would produce identical amounts of discovery expenditures to which they must respond. Let us also assume that requesting and responding to all impositionally cost-effective requests would bankrupt both of the participants, but that merely making the requests would not.

If one party in this situation actually “launches” an unanswered impositionally cost-effective request against her opponent, and the other party launches no such request, the requesting party will then inflict decisive damage on her opponent. Under the assumptions we have set forth, her opponent must either refuse to answer the request (and presumably forfeit the suit) or answer the request and go bankrupt. In most cases, however, the defendant is likely to be able to launch his own impositionally cost-effective request before the court compels him to respond to the plaintiff’s request. This holds true for two reasons. First, impositionally cost-effective requests are typically easier to create than they are to answer. Second, the courts generally move slowly enough on discovery matters that the defendant could, in all likelihood, delay proceedings on sanctions against him until he had tendered his own request against his opponent and sought similar sanctions against her.

If all these assumptions are correct—and I have admittedly constructed a lengthy chain, though I do not think it a weak one—then this situation is analogous to the stable world of nuclear deterrence. If both attorneys refrain from launching abusive discovery requests, the status quo persists and both attorneys (and their clients) are relatively happy. If one side launches abusive discovery requests, the other side will indeed suffer gravely. But the probabilities indicate that the “attacked” side will be able to mount a retaliatory strike. Thus, the real, dynamic outcome of launching an abusive discovery request becomes a disaster for the initially abusive party as well as the initially abused party. In this situation, we would expect that rational parties with a modest ability to peer into the future would solve the Prisoner’s Dilemma on their own and refrain from abusing discovery.

“Our client is a large conglomerate that has its undercapitalized opponent in this litigation by the b—s.”

One might contrast stable deterrence with a situation in which one, and only one, of the parties has the resources to make and respond to all impositionally cost-effective discovery requests. The party flush with resources can then launch its abusive discovery requests without fear of harmful retaliation from the impoverished party. This is the equivalent of an effective first-strike capability: one side can “shoot” at the other and do great damage, while the attacked side can do little to the aggressor in return.

100 See supra note 12 and accompanying text.
Thus, discovery abuse is likely to be more common where the parties have asymmetrical resources available for litigation purposes.

Similarly, discovery abuse is likely to be prevalent in situations where there exists an asymmetry in the opportunities for discovery. Just as possessing a greater number of survivable weapons makes one more confident about launching a first strike, possessing resources less vulnerable to being drained by responses to the opponent's discovery requests should make one more confident about undertaking discovery abuse. The relationship between the two parties is destabilized by the lack of credible, mutual retaliatory threats.

The simultaneous existence of asymmetrical resources and asymmetrical discovery opportunities may in some cases actually combine to produce a more stable situation. In a shareholder's derivative suit, for example, the corporation may possess greater resources than the shareholder, but the shareholder may possess many more opportunities to request discovery from the corporation than the corporation can request from the shareholder. A "balance of terror" might result. Two wrongs would thereby make a right, or at least make the net effect on the likelihood of discovery abuse relatively difficult to predict.

"This case is basically just two big corporations p—ing in each others' tents."

Let us now assume that both parties have the resources available to tender and receive all of the impositionally cost-effective discovery requests available in the suit. Both sides will therefore be worse off, but not disasterously so, if both abuse discovery. This situation calls to mind the static Prisoner's Dilemma discussed in Part III, and abuse is concomitantly likely in the absence of the evolution of a cooperative solution to the problem as discussed in section A of Part IV. The situation is roughly equivalent to a "limited nuclear war" (an oxymoron that ranks with "deliberate speed," though one can hardly compare the two in any other way), in which the two sides lob abusive requests at selected targets of discovery until either one side caves in or runs out of requests to be launched. I note, though it pains me to do so, that the problem in this instance is that in some sense the parties do not have available to them enough opportunities for discovery abuse. If the two parties believed that abuse could lead to mutual disaster, instead of merely viewing abuse as an unpleasant but necessary expenditure of resources, perhaps discovery abuse in complicated cases between two large corporations with plentiful resources would disappear.

101 I discuss below, in the application of the theory of escalation to discovery abuse, some of the factors that make such an evolution less likely than it was in Axelrod's tournaments. See infra pp. 626-28.
The uneasiness with which I present this conclusion, however, is dwarfed by the uneasiness with which one must acknowledge the fundamental fact of superpower relations since World War II: against a backdrop of an abundant store of nuclear weapons, the United States and Western Europe on the one hand, and the Soviet Union and Eastern Europe on the other, have fired nary a single shot at one another in anger. In light of the history of Europe and the ideological gulf that divides the superpowers, this is an extraordinary achievement. In addition, the post-war era has been stable despite the instability that generally exists when two large powers face one another. Understandably, it is nonetheless difficult for most to welcome nuclear weapons with open arms.

C. Discovery Abuse and the Theory of Nuclear Escalation

"It's not abuse. It's thorough research into the facts of a very complicated case."

Although law firms possess sophisticated software in their plushly carpeted offices, they generally lack the satellite surveillance systems employed by the superpowers to spy on each other. Nonetheless, the dynamics and ambiguities of going toe-to-toe with one's opponent in litigation bear distinct similarities to the dynamics of being eyeball-to-eyeball with one's opponent in a crisis between the superpowers.

The driving ambiguity in the case of discovery stems from the inevitable difficulties in determining whether an informational request made by a partially ignorant opponent is tendered out of ignorance or spite—that is, whether the request is a plea for enlightenment on a subject about which the requesting party is legitimately entitled to some information or is rather an impositional salvo in an abusive strategy. As discussed above, a superpower motivated by a desire to avoid war might disperse its bombers because it fears a first strike from its opponent, while a superpower motivated by a desire to initiate war might disperse its bombers to ready some of them for its own attack and to increase the chances that the remainder will survive any retaliation by the victim. In a similar manner, a litigant motivated by a desire to use the discovery rules purely to gather information might tender a wide-ranging discovery request because he lacks information on a wide variety of complicated factual matters, while a litigant motivated by a desire

102 While one might ascribe this situation to the division of the ambitious and militarily competent German state, most would probably agree with the assertion that nuclear weapons have also had a great deal to do with this peace in Europe.

103 I do not address the question of whether we would feel more or less comfortable arming a pair of Wall Street law firms with nuclear weapons than we feel about their possession by the President of the United States and the General Secretary of the Soviet Union's Communist Party.
to abuse the discovery rules might tender a complicated and wide-ranging discovery request because such requests impose substantial impositional losses on his opponent. The judge is likely to have a difficult time discerning the litigant's true motivations, unless she possesses some sort of X-ray discovery vision that allows her to see the naked truth.

This difficulty in discovery, it should be emphasized, will not go away with some judicious tinkering with the Federal Rules of Civil Procedure. Ignorance may or may not be bliss, but ignorance is definitely going to be with us for a long time with respect to divining a requesting party's purposes in formulating a discovery request. The requesting party will almost always appear to the responding party to have asked some irrelevant or excessively far-ranging questions, in light of the fact that the responding party generally knows in advance which questions are relevant to itself (while the requesting party does not) and in view of the generally unfavorable light in which attorneys on opposite sides of a case tend to view one another's motives. The fact that the responding attorney has to go to the trouble of tracking down the information necessary to respond is unlikely to aid matters. The judge may be more objective, but she lacks the requisite information.

"The cases keep getting more complicated."

Under the theory of cooperation, we found that the increased complexity of cases may or may not foster cooperation among opponents. From the point of view of the theory of escalation, in contrast, increased complexity is likely to be unambiguously damaging to the prospects for cooperation in discovery. Just as the complexity and coupling of the superpowers' warning-and-alert systems are damaging to the prospects of their maintaining control in a deep crisis, the complexity of a particular lawsuit appears to militate against cooperation in discovery. Complicated cases generally require a significant amount of discovery in order to learn about the gravamen of the suit. It follows that opportunities in such cases for either side to misinterpret legitimate requests—which will of necessity be more complicated and wider-ranging—may grow as well. The party interpreting the initial request as an impositional one may then retaliate by formulating an abusive request, and thus the cycle of escalation begins. Ambiguity leads to uncooperativeness in discovery as it does in superpower crisis.

"The firms just keep getting bigger."

Under the tenets of the revisionist theory of nuclear escalation, an increase in the average size of law firms is likely to exacerbate the problems of discovery abuse. The theory of escalation is concerned not only about the complexity of the respective superpowers' intelligence-aid-warning systems,

[104 See supra text accompanying note 98.]
but also about the degree to which high-ranking politicians can actually control their military forces. As the "chain of command" lengthens in a system in which military officers view politicians as interlopers, and in which behavior at the lowest ranks can have important effects on the course of world events, the ability of top political leaders to bend their military establishments to their will assertedly decreases.\(^{105}\)

A large firm presents similar difficulties in maintaining control in comparison to keeping control of a small firm. The capacity of top leadership in a large firm or government office to fine-tune decisions carried out by subordinates far down the chain of command is simply less than at a four-person firm. Too much goes on in the offices of those under the U.S. Attorney for the Southern District of New York for one person to comprehend and control, even if the job did not encourage politico-legal tendencies; similarly, too much goes on at a Wall Street firm for the senior partners to understand even if they are still capable of working twenty hours a day. The foot soldiers in the discovery war are certainly not senior partners. The possibility that some shots will be fired, even if the top management does not desire them, is therefore real.

D. Reforms to Alleviate Dynamic Discovery Abuse

These applications of dynamic theories to discovery abuse suggest some possible reforms,\(^{106}\) many of which are not of the traditional tinker-with-the-Rules sort. I set forth these reforms in this section. I treat as a significant, but not dispositive, issue the likelihood that there would ever be sufficient support within the legal community to implement the various suggested reforms.

1. Limit the Length and Rate of Discovery Requests

One reform suggested by these dynamic theories of discovery abuse, and one kindred in spirit to the traditional redraftings of the Federal Rules governing discovery, would be to set a limit on the length of discovery requests. Such a limit would increase the number of interactions in a suit of

\(^{105}\) The counterargument is that modern political leaders have modern tools of communication available to them, and that such tools allow those at the top of the chain to control directly those towards the bottom of the chain. President Kennedy, for example, was able in the Cuban Missile Crisis to speak to the captains of the ships enforcing the blockade of Cuba as Soviet ships approached. See G. Allison, ESSENCE OF DECISION (1971).

\(^{106}\) Axelrod proposes in a very general way some reforms flowing from his theory. See R. Axelrod, supra note 8, at 124-41. Bracken also sets forth several suggestions for changes in international relations. See P. Bracken, supra note 68, at 238-47.
given complexity, thereby lengthening the shadow of the future and encouraging cooperation in discovery.\textsuperscript{107}

Presumably, the Rules would need to set forth such a limit on the rate at which a party can file discovery requests as well as the length of each request. Otherwise, a party could simply send along 500 one-page requests on a single day under a "limited"-length rule, and no change in the incentives for cooperation would occur.

The tradeoff that results with a length-and-rate limit, however, is the possibility that the limitations will wind up delaying the case excessively. A party in a complex case might, for example, wish to request hundreds or even thousands of answers through interrogatories; if the rate-and-length limit is "too" constraining, then that process might take even longer than it would have when the parties had greater incentives to abuse discovery. This may be undesirable, especially since delay itself increases the incentives for discovery abuse.\textsuperscript{108} One might agree upon length-and-rate limits on a case-by-case basis, however, or set up an experimental program of such limits in selected courts.

2. Time Limits on Discovery Requests

Similar considerations argue for the consideration of a time limit on discovery requests: if discovery finishes more rapidly, the future casts a longer shadow at the beginning of the process, and the incentives for cooperation increase. The countervailing consideration in the case of time limits is a bit different from the concern with delay in length-and-rate requests—indeed, time limits would be one way to make length-and-rate requests even more effective in encouraging cooperation. Rather, the factor to weigh in the balance against time limits is the possibility that so little time might exist for completing discovery that the parties would wind up sacrificing a great deal of the informational benefits of discovery in order to avoid suffering its impositional costs. Presumably justice hurried can be as unfair as justice delayed. Nonetheless, time limits seem likely to be worthy of consideration in any future set of reforms to discovery procedures.

3. Restrict the Size of Law Firms

An increase in the average size of law firms implies, under certain reasonable assumptions, an increase in discovery abuse under either the theory of cooperation\textsuperscript{109} or the theory of escalation.\textsuperscript{110} Arresting the growth of firms might therefore reduce discovery abuse. Nonetheless, I would wager that, in

\textsuperscript{107} See supra text accompanying notes 95-97.
\textsuperscript{108} See supra text accompanying note 99.
\textsuperscript{109} See supra pp. 616-23.
\textsuperscript{110} See supra pp. 626-28.
this great nation of ours, such limits are unlikely to pass muster with state or federal legislatures stocked mainly with lawyers, or with judiciaries composed almost entirely of attorneys. Similarly, proposing a limit on the mobility of lawyers, either in their freedom of movement from firm to firm or from geographic locale to geographic locale, might reduce discovery abuse,¹¹ but is likely to cause an outcry far greater than has ever accompanied descriptions of discovery abuse. Restrictions in the movements of the factors of production, whether they be capital or labor, generally clog the arteries of capitalism in an undesirable fashion. Some irate attorney with itchy feet would doubtless also file a constitutional challenge to such limitations as restrictions on his freedom to travel.¹¹²

4. Bail Out Firms Bankrupted by Discovery Abuse

The theory of cooperation suggests that the specter of bankruptcy chases away the shadow of the future,¹¹³ and thereby lessens the possibilities for cooperation. One might therefore consider the possibility of bail-outs for law firms. I doubt that such a possibility will become a riveting issue until a spate of large firms fail, or firms become so large that a single failure will be as allegedly disruptive as the demise of Chrysler. In any case, given the breadth of the relevant issues in such a situation, the effect of a bail-out on discovery abuse via the theory of cooperation is likely to flare only briefly in the heat of the debate.

5. Limit the Number of Practicing Attorneys

As discussed previously, increases in the number of practicing attorneys are likely to exacerbate discovery abuse.¹¹⁴ Limiting the number of practicing attorneys would benefit those who remain in (or enter) the pool. Of course, the decrease in the supply of legal services should, all other things remaining equal, also lead to a rise in the price of their legal labors. One might thereby muster the support of the current bar for, say, more rigorous bar examinations. The general biliousness of the person on the street towards lawyers further implies that public opinion might be channeled in favor of reducing the number of attorneys. Nonetheless, such a move would have a number of systemic effects—for example, a rise in torpidity of our courts due to limits on the time available to those who wear dark suits rather than dark robes. Reformers backing such a proposal should therefore at least examine those systemic effects before lending their support to limits on the pool of practicing attorneys undertaken simply to reduce discovery abuse.

¹¹¹ See supra pp. 617-19.
¹¹³ See supra p. 621.
¹¹⁴ See supra p. 617.
6. Reduce Court Delays

Court delays increase the incentives for discovery abuse.\textsuperscript{115} Reducing delay is, therefore, an idea with some merit to those who wish to reduce discovery abuse. Reducing court delays generally seems a proposition with roughly the breadth of support granted to leaner beef, homier apple pies, and reductions in the likelihood of nuclear war. Butchers, bakers, and policymakers nonetheless differ on the proper methods of achieving their respective goals, and so too do attorneys. The proper way to reduce delay, and how to pay for the costs incurred in doing so, is a difficult question, and one that I will not attempt to address. One should nonetheless be sure to include reductions in discovery abuse among the benefits produced from a reduction in court delays.

7. Encourage Specialization of the Bar

Encouraging an increase in the specialization of the bar would have several effects besides its influence on discovery abuse, though I think on balance this specialization might well be desirable. The specialized bars should be better behaved with respect to discovery abuse simply because their numbers are likely to encounter one another more frequently, and therefore can more reliably benefit from cooperative, i.e. non-abusive, discovery strategies.\textsuperscript{116} The state bars might require attorneys to pass particular examinations before undertaking certain types of cases, just as those who practice medicine must successfully complete examinations in pathology or surgery before undertaking the practice of those specialties. The creation of specialized courts might have similar benefits: the judges of such courts would be able to interact more frequently with the specialized attorneys, and thus obtain a better idea of which lawyers were likely to abuse discovery. Indeed, a special concern for discovery abuse would lead the bar to focus, within specializations, on those types of cases most likely to lead to discovery abuse, such as anti-trust litigation. As in the case of reducing court delay, however, specializations of these sorts may raise questions beyond the horizons of discovery abuse, or even of procedural abuses generally.

8. Preserve “Second-Strike Capability”

The theory of deterrence suggests that the legal system should attempt to ensure both litigants the equivalent of a secure second-strike capability, though without going so far as to make them insensitive to the attacks of their opponents. This requirement is not as easy to conceptualize within a domestic legal system as it is within the realm of international relations, but some measures come to mind.

\textsuperscript{115} See supra pp. 616-17.
\textsuperscript{116} See supra p. 621.
The first is mostly exhortatory: judges should be especially zealous in applying discovery sanctions against a party who, because of its relative advantage in resources or in opportunities for discovery, has little to fear from the retaliatory capacity of its opponent. Judges should also be on the lookout for abuse in cases where both parties' resources are sufficient to absorb all impositionally cost-effective requests—where the threat is symmetrically undisturbing rather than asymmetrically quite threatening. Many judges seem likely to exercise such vigilance already, however, at least if they have much intuition about the generalities of equity or the specifics of discovery abuse.

The second measure in this category is more concrete, though it applies only to asymmetries in resources, not in opportunities for discovery abuse: providing court-appointed attorneys to civil litigants of limited resources. Such a measure would lessen the incentives for a resource-rich party to attempt a disarming first strike against the resource-poor party, since such a provision increases the staying power of the party possessing fewer resources. Such appointments are clearly not without their social costs, however, and a court would doubtless wish to assure itself that its appointments went to deserving parties rather than to the frivolous filers (or defenders) who sometimes enter the federal-courts fray.

9. Reduce the Complexity of Cases

The complexity of cases is perhaps the factor least amenable to reform of all these relatively intractable issues. For all its complaints about the results, society and its elected representatives have chosen to allow a whole variety of complex litigation, such as anti-trust suits and Chapter 11 bankruptcy actions, to flop into the federal courts. Society outside the courtroom—whether in terms of communications, corporate transactions, or international security—has become significantly more complex in the past forty years or so. If the trend continues, as it seems likely to do, one should be pessimistic about the ability of the Congress and the courts to resist the invasion of complexity into their halls and chambers to keep cases simple merely by fiat.

10. Clarify the Intent of Discovery Requests

Unfortunately, similar pessimism should adhere to any hopes that the ambiguous signals sent by a discovery request—whether the request is one seeking information or intimidation—can somehow be rendered clear. After all, discovery is about discovering—discovering information in the possession of the other party. The responding party will almost inevitably be unable to tell, with much certainty, whether the requesting party needs the information; the requesting party will almost invariably be incapable of determining whether responses to its questions will be simple or difficult for
the responding party to provide. This informational curtain is likely to prove less permeable than the iron one.¹¹⁷

VI. CONCLUSION

Those faced with apparently intractable problems typically seek the refuge of calling for greater education with at least as much verve as scoundrels take refuge in patriotism. I join the former crowd by noting that, in many ways, all one can hope to do about discovery abuse is to demonstrate to attorneys the difficulties created by their abusive ways. Just making them aware that a cooperative solution to discovery abuse can evolve under certain conditions, and that such a solution benefits them and their clients, may give them a nudge towards adopting a cooperative discovery strategy. An awareness that it is inherently difficult to determine whether discovery requests are either informational or impositionally abusive may encourage the responding party to keep an open mind, and thereby delay the onset of abusive requests by that party. And finally, if perhaps most forlornly, one might hope that attorneys could be convinced to relax and adopt a somewhat more cooperative attitude towards their opponents simply because the stakes in litigation are not as high as they might be. After all, in contrast to nuclear war, discovery and its use (or abuse) is not the end of the world.

¹¹⁷ See, e.g., “Poland and Hungary Lead the March West,” Washington Post, April 16, 1989, at A1, A3 (“The Iron Curtain that once stood between Hungary and Austria in effect has moved east: Budapest has announced that its 20-year-old forces and alarms along the Austrian border will be dismantled by 1991, while the Romanian border is now tightly sealed and guarded.”).