Mechanism Design and Behavioral Economics: Incentivizing Optimal Pre-Trial Discovery

Andrew B. Brod

Abstract: Practitioners and judges recognize excessive civil discovery to be a widespread and serious issue. Since litigants bear only a small portion of the costs of their own discovery requests and face unconscious psychological biases, they regularly seek far more discovery than would be relevant and proportional to proving their case. This in turn massively inflates litigation costs for their opponents. The 2015 Amendments to the Federal Rules of Civil Procedure (FRCP) recognized and confronted this problem directly: they rewrote Rule 26(b)(1) to encourage judges to more actively ensure that discovery requests are “proportional” to the needs of the case. In showing precisely how judges should implement the revised Rule 26(b)(1), I first explore the causes underlying excessive discovery. I draw on insights from classical economics (which understands litigants as purely rational actors) and from behavioral economics (which explores litigants’ psychological biases) to explain the current prevalence of excessive discovery. I then use techniques from mechanism design, a branch of mathematical economics, to develop a new discovery system. Under this system, judges formulate rigorously the “proportionality” standard envisioned in the amended FRCP, which they implement by correcting for the aforementioned causes of excessive discovery. I conclude by exploring questions regarding this proposed system’s real-world feasibility.
Introduction

Excessive pre-trial discovery is widely perceived as a fundamental problem in American civil litigation. By giving parties the opportunity to fully inform themselves about the facts of a given suit, discovery should, when it functions well and serves its true objectives, lead to a more effective system of civil procedure.¹ The discovery process should serve the interests of justice by allowing courts to arrive at more accurate verdicts, and it should improve the overall system’s efficiency by encouraging litigants to settle early.² Many attorneys, however, believe discovery does not live up to these ideals in practice. In a 2008 survey of Fellows of the American College of Trial Lawyers, fewer than half of respondents believed “current discovery mechanisms work

¹ See FED. R. CIV. P. 1 (“[These rules] should be construed, administered, and employed by the court and the parties to secure the just, speedy, and inexpensive determination of every action and proceeding.”).
² See Richard A. Posner, Legal Procedure and Judicial Administration, 2 J. LEGAL STUD. 399, 422-426 (1973) (arguing that discovery should lead to convergent expectations among litigating parties, leading to a greater likelihood of settlement).
well.” The final report summarizing the results of this survey is even more blunt, concluding that “[o]ur discovery system is broken.” A full 82% of respondents to a 2009 American Bar Association survey believed discovery to be too expensive, and over half felt it was commonly abused. At its worst, discovery may allow litigants to impose extensive costs on their opponents, and in the views of some scholars, actually discourages parties from settling. Parties’ self-serving behavioral economic biases compound these problems, causing a divergence in parties’ expectations regarding trial outcomes and further discouraging settlement.

Improving the discovery process has been a recent judicial priority. The 2015 Amendments to the Federal Rules of Civil Procedure (FRCP) introduced prominent changes to Rule 26(b): in an effort to “deal with the problem of over-discovery,” the Amendments introduced to Rule 26(b)(1) a proportionality standard governing the scope of discovery. These changes envision a more active role for judges in reigning in discovery.

8 Samuel Issacharoff & George Loewenstein, Unintended Consequences of Mandatory Disclosure, 73 TEX. L. REV. 753, 753 (1995).
9 FED. R. CIV. P. 26, advisory committee’s notes, 2015 Amendments.
10 FED. R. CIV. P. 26(b)(1).
11 FED. R. CIV. P. 26, advisory committee’s notes, 2015 Amendments (noting that the rule now “contemplates greater judicial involvement in the discovery process”).
ought to make determinations regarding “proportionality” is, of course, an open question, any resolution of which requires a detailed understanding of the root causes of over-discovery.

The existing literature has explored several explanations for the problem of excessive discovery. When litigants fail to internalize the full costs of discovery, for example, they may request more than the socially optimal amount.\textsuperscript{12} Alternatively, strategic abuse of discovery may create an \textit{in terrorem} effect, forcing parties to settle prematurely out of fear of unmanageably excessive and disproportionate costs.\textsuperscript{13} While such insights from classical and informational economics are crucial to understanding the discovery process, they only capture part of the story. By dropping the classical economics assumption of full rationality, behavioral economics can offer a fuller picture of litigants’ actions, focusing instead on the psychological factors governing their decisions. It has long been recognized that parties that choose to litigate in the first place are “irrational” from a traditional economics perspective,\textsuperscript{14} and there exists a compelling case that such “irrationality” should persist in the parties’ choice of what they perceive to be an optimal amount of discovery. Any attempt to improve the discovery process must therefore take such behavioral economic factors into consideration.

It is precisely these insights that I now seek to implement through the theory of mechanism design. This is a branch of mathematical economics that functions as a sort of reverse game theory: by defining the rules of strategic interactions among participants, one attempts to induce behavior so as to optimize the designer’s chosen measure of utility. Most famously, mechanism design has been applied to designing optimal auctions,\textsuperscript{15} constructing a public

\textsuperscript{12} See Jonah B. Gelbach & Bruce H. Kobayashi, \textit{The Law and Economics of Proportionality in Discovery}, 50 GA. L. REV. 1093, 1099-1105 (2016); Cooter & Rubenfeld, \textit{supra} note 7, at 452 (concluding that “the rational plaintiff will not stop conducting discovery until he has gone beyond the point where it becomes abusive”).

\textsuperscript{13} See Kobayashi, \textit{supra} note 6.


good,\(^{16}\) and designing a system of bilateral trade.\(^ {17}\) Here, I explore how a judge might use the techniques of mechanism design to maximize social utility with respect to civil discovery by effectuating the 2015 FRCP Amendments.

There is a sizable literature that examines litigation from the perspective of mechanism design. Demougin and Fluet, for example, demonstrate the optimality of the “preponderance of evidence” standard of proof for incentivizing potential tort-feasors to exert care;\(^ {18}\) Kaplow analyzes the optimality of choosing whether to continue versus terminate at various stages of a legal case;\(^ {19}\) and Silva contrasts “confession inducing” mechanisms to “trial” mechanisms when there is suspicion that a crime has been committed.\(^ {20}\) Of particular relevance is a recent article by Jonah Gelbach, which also focuses on how one might implement the new proportionality standard in civil discovery.\(^ {21}\) While a number of papers examine the law and economics of discovery more broadly,\(^ {22}\) few papers analyze discovery using the tools of mechanism design, and only Gelbach focuses on the proportionality standard that is also the subject of this paper.\(^ {23}\)


\(^{21}\) Jonah B. Gelbach, *Can Simple Mechanism Design Results be Used to Implement the Proportionality Standard in Discovery?*, 172 J. INSTITUTIONAL & THEORETICAL ECON. 200 (2016).


From a technical standpoint, the model presented in this paper differs significantly from Gelbach’s. Much more interesting, however, are its broader conceptual differences. Like its precursors in the mechanism design literature, my model imagines litigants as economic actors and attempts to understand their actions through the formal language of game theory. But it departs from Gelbach and the rest of this literature by recognizing, and accounting for, the bounded rationality of litigants.

Survey data makes clear that many litigants themselves recognize the ineffectiveness of the large majority of requested discovery materials. And yet, “many lawyers believe that they should—or must—take advantage of the full range of discovery options offered by the [FRCP]” since “zealous advocacy . . . demands no less.” A central goal of Gelbach’s mechanism—as in much of the mechanism design literature—is to constrain parties’ abilities to act in their own self-interest, permitting only discovery requests that are efficient after internalizing costs to other parties. My own mechanism fully recognizes and accounts for this dynamic. But my

---

24 In Gelbach’s model, a plaintiff and a defendant draw correlated signals from a discrete type space about the value of a particular discovery request. Gelbach specifies a functional form by which the litigants transform their signals into a change in expected payoff. He then carries out a Crémer-McLean construction to find a mechanism that is ex post efficient, incentive compatible, individually rational, and unsubsidized. Gelbach, supra note 21; Jacques Crémer & Richard P. McLean, Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions, 56 Econometrica 1247 (1988). My model, by contrast, focuses exclusively on the interactions between the plaintiff and the judge, so that any role for the defendant is suppressed. My model also does not specify a particular functional form for the plaintiff’s value for discovery, but rather permits a broad class of value functions. Additionally, my model uses a continuous type space, which (alongside the fact that there is only one signal drawn) renders the Crémer-McLean construction inapplicable. While the mechanism I construct is also efficient, incentive compatible, individually rational, and unsubsidized, I rely on different tools for its derivation.


26 See, e.g., Beisner, supra note 5, at 4. As early as 1980, “practitioners estimated that 60 percent of discovery materials did not justify the cost associated with obtaining them.” Id. And this problem has only gotten worse, with a survey of Fortune 200 companies finding that, in 2008, “the average number of discovery pages to the average number of exhibit pages (that is, pages actually utilized in some fashion at trial) . . . was 1,044 to 1.” Id. This should not be surprising with the advent of electronic discovery, which has led to an “exponential growth” in documents produced—and to a concomitant increase in the burden of producing, and analyzing, such vast amounts of information. See id. at 2.

27 Final Report, supra note 4, at 7.

28 See Gelbach, supra note 21, at 210; supra note 12 and accompanying text.
mechanism also recognizes litigants’ tendencies to seek discovery in excess of what is beneficial to their own case, and in that sense, it has a more cooperative interpretation than does Gelbach’s mechanism. Insofar as litigants experience biases that distort their assessments of an optimal amount of discovery, my mechanism actually helps them to act in their own best interests, rather than serving (at least from the litigants’ perspective) a purely adversarial function. Thus, by incorporating behavioral economic considerations into an analysis of classical economic incentives, my implementation of the FRCP proportionality standard improves the efficiency of the civil justice system while simultaneously appealing to the self-interest of litigating parties.

My paper proceeds as follows. Part I briefly reviews the 2015 FRCP Amendments, examining how changes to Rule 26 reflect an awareness of, and an attempt to mitigate, problems with excessive discovery. Part II then explores several categories of explanations for excessive discovery, focusing in particular on explanations rooted in behavioral economics. Finally, in Part III, I address these problems by “implementing” the 2015 FRCP Amendments: I construct an optimal discovery mechanism that simultaneously addresses classical and behavioral economic causes for excessive discovery, forcing litigants to appropriately internalize discovery costs and “debiasing” them from their “irrational” psychological inclinations. I derive key properties of such a mechanism, and I explore considerations for its real-world feasibility.

I. Judicial Oversight and Amendments to the Federal Rules of Civil Procedure

The Federal Rules of Civil Procedure were amended in 2015 in large part to address perceived issues with excessive discovery. As amended, Rule 26(b)(1) reads in relevant part:
Unless otherwise limited by court order, the scope of discovery is as follows:

Parties may obtain discovery regarding any nonprivileged matter that is relevant to any party’s claim or defense and proportional to the needs of the case, considering the importance of the issues at stake in the action, the amount in controversy, the parties’ relative access to relevant information, the parties’ resources, the importance of the discovery in resolving the issues, and whether the burden or expense of the proposed discovery outweighs its likely benefit.  

This language regarding proportionality was originally introduced into the Federal Rules of Civil Procedure in 1983, but subsequent amendments minimized it in other rule subparts. The 2015 FRCP Amendments recognized a “shortfall in judges’ and parties’ use of the proportionality standard,” as a result of which this language was moved to its current, far more prominent location. A 2016 American Bar Association report suggests that this change could be “profound”: “The restoration of the proportionality factors as a qualifier . . . can only be viewed as an effort to encourage the parties to narrow their discovery requests, or as the Notes [to the 2015 Amendments] state, ‘deal with the problem of over-discovery.’” Moreover, the Notes to the 2015 Amendments state explicitly that the new rule “contemplates greater judicial involvement in the discovery procedure.”

---

29 FED. R. CIV. P. 26(b)(1) (emphasis added to indicate new language).
31 See Gelbach & Kobayashi, supra note 12, at 1097-99.
32 Marinelli, supra note 30.
33 Id. (quoting FED. R. CIV. P. 26, advisory committee’s notes, 2015 Amendments).
process,” in contrast to the largely self-regulating manner in which discovery currently functions.34

From this brief discussion alone, it is clear that the judicial system perceives excessive discovery to be a significant problem, one which a more involved judge would ideally mitigate. But this raises immediate questions regarding how a judge might in fact perform such a role.

How should one define proportionality? Before anything else, a judge must make a baseline assessment of what constitutes proportionate versus excessive discovery. Some degree of discovery is beneficial insofar as it enables meritorious litigants to achieve a just outcome, whereas over-discovery imposes costs on all parties that exceed the requestor’s likely benefit. Consistent with standard practice in law and economics, one might interpret Rule 26(b)(1) to be commanding judges to define some sort of social welfare function,35 which an optimal amount of discovery would maximize.

Moreover, how should one then assess proportionality? This is particularly challenging because of the presence of imperfect information. A judge may wish to compare the benefits and detriments of a given amount of discovery, but the judge is far less capable of estimating such benefits than is the litigant that requests discovery in the first place.36 On the one hand, a judge is much less familiar with the details of a given

36 The assumption that a judge is information-constrained compared to the litigating parties is common to the theoretical literature on civil and criminal procedure. See, e.g., Gelbach, supra note 21, at 203 (characterizing as private information litigants’ beliefs about the value of discovery, such that courts know only the probably distributions of these beliefs); Kaplow, supra note 19, at 620 (setting up a model where a tribunal, reacting to a signal about a defendant’s private information, must decide whether or not to proceed with a case against the defendant); Spier, supra note 20, at 200 (noting that litigants have private information about their damages and liability, about which the court observers a signal); Cooter & Rubinfeld, supra note 7, at 458 (noting that a court is “poorly placed” to “estimat[e] the expected value of the information to the party making the [discovery] request”);
case—especially at the pre-trial discovery stage—than is the party bringing suit; this suggests that the judge owes a large degree of deference to the litigant requesting discovery. At the same time, however, the judge cannot simply take a litigant at his word that the proposed amount of discovery is optimal. As discussed in much greater detail in Part II, a party’s choice of the optimal amount of discovery likely does not correspond to the socially optimal amount of discovery. Since the requestor typically bears a relatively small proportion of the costs of discovery, he has an incentive to choose an amount that exceeds the social optimum. Well-studied behavioral biases compound this effect, as parties often do not have a realistic estimate of the extent to which discovery will help their case. An optimal discovery process must therefore understand, and correct for, both classical and behavioral economic incentives before granting any discovery request.

To a student of informational economics, the above description is strongly reminiscent of classic problems of mechanism design. A social planner seeks to maximize a social welfare function. To do so, however, the planner must rely on private information held by other parties. These other parties have incentives to misstate their private information, leading the social planner to optimize social welfare incorrectly. The planner must therefore set up a “mechanism” that induces other parties to reveal their private information while simultaneously taking into account these parties’ incentives for strategic decision-making. It is precisely this task that I undertake in Part III, though first, I examine in detail litigants’ behavior during the discovery process, clarifying the factors that a judge-cum-social planner must take into consideration.

Ron Siegel & Bruno Strulovici, Judicial Mechanism Design 2 (October 2020) (unpublished manuscript) (assuming that a defendant in a criminal trial privately knows his guilt among other private information, and so a judicial process must generate evidence to determine the defendant’s guilt).
II. **Explanations for Excessive Discovery in Litigation**

A. **Behavioral Economics**

A sizable literature analyzes discovery from the perspective of a rational, economic actor. Under such theories, a litigant requests precisely the quantity and type of discovery that is likely to best serve the litigant’s case. While perfect rationality on the part of individual litigants is insufficient to guarantee a socially optimal amount of discovery (see the discussion in Sections II.B-C), it does serve as a useful framework for analyzing litigants’ behavior.

Behavioral economics challenges this classical view. No longer can one take for granted the rationality of litigants; rather, bounded rationality distorts their decision-making processes, leading litigants to engage in discovery that detracts not just from social welfare, but from their own private welfare as well. Most literature studying this phenomenon analyzes the role of behavioral economics in litigation more generally—focusing, in particular, on parties’ decisions on whether or not to settle—but one can readily extend these insights to the discovery process. I focus on three insights from behavioral economics, all of which should lead litigants to request more discovery than a rational actor would otherwise find optimal.

*Self-serving bias* is the tendency to conflate what is fair with what is in one’s own self-interest.\(^{37}\) Parties with such biases interpret information in a highly self-serving manner: Babcock and Lowenstein attribute to this phenomenon people’s overwhelming tendencies to rate themselves in the top fifty percent in a wide variety of skills and positive qualities, including driving, ethics, and health.\(^{38}\) They also see such self-serving informational biases as causing


\(^{38}\) Id. at 111 (citing a variety of studies).
people to overestimate their individual contributions to joint projects.\textsuperscript{39} And in another study, Princeton and Dartmouth students watched the same football game between their colleges and counted the penalties committed by each team; the students displayed self-serving biases by recognizing vastly different penalty counts for each team, strongly favoring their own university’s side.\textsuperscript{40}

In the context of litigation, this bias shows up through litigants’ expectations that neutral parties would be favorably disposed toward their case, and that a full consideration of a litigant’s grievances represents the only “fair” resolution to a dispute.\textsuperscript{41} Litigants have an unrealistic perception of what constitutes a fair trial outcome, they believe neutral judges will favor their arguments, and they are unwilling to accept less than what they perceive as a fair outcome.\textsuperscript{42} Since litigants pursue “fairness” more so than a strict maximization of monetary value, such biases considerably decrease the occurrence of settlement, causing litigation procedures to drag on far beyond what would be economically rational for either party.\textsuperscript{43}

The implications of self-serving bias on discovery are immediate. First, a decreased rate of settlement makes it much more likely for litigants to engage in a costly and lengthy discovery process.\textsuperscript{44} Moreover, litigants’ tendencies to interpret information favorably to their case should incentivize the collection of further such information. The literature shows that information

\begin{small}
\textsuperscript{39} \textit{Id.}
\textsuperscript{40} Id.
\textsuperscript{41} See Linda Babcock et al., \textit{Biased Judgments of Fairness in Bargaining}, 85 \textit{AM. ECON. REV.} 1337, 1337 (1995).
\textsuperscript{42} See \textit{id.}; George Loewenstein et al., \textit{Self-Serving Assessments of Fairness and Pretrial Bargaining}, 22 \textit{J. LEGAL STUD.} 135, 142, 152-153 (1993); see also Kathryn E. Spier, \textit{Litigation, in 1 HANDBOOK OF LAW AND ECONOMICS} 259, 278 (A. Mitchell Polinsky & Steven Shavell eds., 2007) (discussing some of the literature related to self-serving biases in litigation).
\textsuperscript{43} See Loewenstein et al., \textit{supra} note 42, at 139, 152.
\textsuperscript{44} See, e.g., Issacharoff & Loewenstein, \textit{supra} note 8 (arguing that behavioral economic biases lead to a divergence in expectations among litigants).
\end{small}
obtained through discovery “is likely to be interpreted egocentrically by disputants.”\footnote{Maya Steinitz, \textit{The Litigation Finance Contract}, 54 WM. & MARY L. REV. 455, 493 (2012) (citing George Loewenstein & Don A. Moore, \textit{When Ignorance Is Bliss: Information Exchange and Inefficiency in Bargaining}, 33 J. LEGAL STUD. 37, 37 (2004)).} By assigning a higher subjective benefit to the contents of discovery, a biased litigant would feel that new information that comes to light should increase either the probability of winning or the expected monetary gain in the event of a victory at trial.\footnote{Elizabeth F. Loftus & William A. Wagenaar, \textit{Lawyers’ Predictions of Success}, 28 JURIMETRICS J. 437, 450 (1988).} Finally, since a party believes neutral judges will agree with the merits of its case, it should also believe judges would agree with the importance of extracting additional information from the party’s opponents, thus driving toward a “fair” outcome for the suit.\footnote{See, e.g., Donald C. Langevoort, \textit{Behavioral Theories of Judgment and Decision Making in Legal Scholarship: A Literature Review}, 51 VAND. L. REV. 1499, 1511 (1998) (noting parties’ willingness “to construe the fairness of the situation in a self-serving fashion” and citing a number of other articles on litigants’ psychological biases).} Regardless of the exact causal explanation, it is clear that self-serving biases should drive litigants to request a larger amount of discovery as compared to an unbiased party.

\textit{Confirmation bias} is the “seeking or interpreting of evidence in ways that are partial to existing beliefs, expectations, or hypothesis in hand.”\footnote{Raymond S. Nickerson, \textit{Confirmation Bias: A Ubiquitous Phenomenon in Many Guises}, 2 REV. GEN. PSYCHOL. 175, 175 (1998).} Once again, the widespread existence of this bias is well-documented in the psychology literature.\footnote{See \textit{id.} at 175-184 (summarizing a large literature on the subject).} One would expect litigants to be especially vulnerable to this phenomenon: a party bringing suit generally has strongly held beliefs and hypotheses about the merits of its case; the ubiquity of confirmation bias suggests that litigants would seek out information to affirm their pre-existing beliefs, rather than to establish the “truth” of a given matter. Indeed, an empirical study by Wistrich and Rachlinski demonstrates that lawyers are subject to considerable confirmation bias, which in turn materially
impacts the discovery process. The researchers redesigned a classic study pioneered by Peter Wason, whereby participating lawyers were asked to pretend they represented a female plaintiff claiming gender discrimination. From a given selection of files, the participants were asked to choose for discovery one or more files that would help to prove the plaintiff’s allegations. They were specifically instructed to avoid excessive discovery. Nonetheless, a large majority of participants chose the incorrect combination of files, often opting for a superfluous amount of discovery.

Bounded self-interest also plays a key role in the litigation and discovery process, pushing people to litigate more, and to request more discovery, than traditional economic models might otherwise predict. In litigation, this takes the form of spitefulness and a desire for fair treatment that go beyond pure monetary considerations. The satisfaction of punishing another party that has behaved “unfairly” is itself often perceived as a material benefit, one which leads litigants to tolerate higher costs and to reject early settlements. Similarly, economic actors are concerned with upholding their own reputations and self-image, leading to “principled” retaliation when, as in litigation, a party’s sense of fairness has been violated. Such behavior manifests itself in a number of well-studied aspects of litigation. For example, litigants often place a high value on having their “day in court,” or, in other words, having the opportunity to share publicly their side of the story in a dispute. Notions of vengeance and vindictiveness also often

52 See Wistrich & Rachlinski, supra note 32, at 600.
53 See Jolls, Sunstein & Thaler, supra note 25, at 1493-96.
54 Id. at 1495.
55 Id. at 1494-95.
play a key role, especially in tort and libel suits: the literature shows that litigants derive considerable emotional benefits from harming their opponents, sometimes more so than from receiving compensation for their own injuries.\[^{57}\] Still other litigants are motivated by a desire to have society formally allocate blame for an incident on their opponents and have themselves recognized as blameless.\[^{58}\] In all cases, the trend is clear: more litigation and more discovery, driving toward emotional vindication that a “rational” economic actor should not theoretically consider.

**B. Externalized Costs**

Returning now to the perspective of classical economics, a litigant’s choice for the quantity of discovery is simple. A party compares marginal costs and benefits associated with a given discovery request; it then requests additional discovery until these two quantities are equal. Intuitively, one would expect that an initial amount of discovery would be especially beneficial to a litigant’s case. Procuring the first wave of documents and depositions reveals crucial information that allows a litigant to formulate his arguments more effectively, considerably increasing the likelihood of winning and the associated economic benefit. However, subsequent discovery requests should be less and less useful. Eventually, it will reach the point that the requesting party’s own discovery costs—preparing the discovery requests, reviewing the documents received, bearing the opportunity costs of activities it cannot pursue due to ongoing litigation, and so forth—exceed any further gain that a party can expect from discovery. Thus, there should be some finite, privately optimal quantity that a party ultimately requests.


Unfortunately, what constitutes privately optimal discovery is almost certainly not socially optimal. The socially optimal amount of discovery equates marginal benefits and costs for all parties, whereas the requesting party considers only its own private costs and benefits. The problem is that of externalized discovery costs: analogously to a factory owner that pollutes the environment through the production process, the requesting party derives almost the full benefit of any discovery, but it creates substantial costs for other parties, a factor that the requesting party does not consider in its optimization calculations. In practice, discovery imposes far larger costs on one’s opponent than it does on the requesting party. Costs usually fall where they lie, so that the party receiving the discovery request must accept the full expense of searching for, identifying, and delivering all documents necessary to comply with the request. This may be multiple times the cost that the requesting party faces to process this new information, creating enormous cost externalities.

C. Strategic Litigant Behavior

Finally, one can interpret discovery as an example of strategic litigant behavior in the context of game theory. Many researchers model litigation as a formal game in which discovery plays a crucial role in signaling, which in turn maximizes one’s utility through the collection of information, or through inducing the opposing party to settle. Among the most important of these strategic considerations is the possibility of overwhelming one’s opponent with unmanageable discovery requests. Such a litigation strategy seeks to force an in terrem

59 Gelbach & Kobayashi, supra note 12, at 1103.
60 Id.
61 Id.
62 See Klerman, supra note 22, at 362-64 (reviewing a large literature on the subject).
settlement of largely groundless claims, solely through the threat of enormous discovery expenditures.\(^{63}\)

While no less important than behavioral biases or externality costs in understanding the discovery process, discovery as strategic litigant behavior is significantly more difficult to generalize, and so I address this topic only briefly. The main issue is that game theoretic conclusions identified in the literature are very sensitive to modeling assumptions: who has private information, who makes settlement offers, to what extent discovery impacts the probability of winning at trial, and so forth.\(^{64}\) Due to the large variety of such models and the complexity of assimilating them into broader insights, I generally limit my discussion \textit{infra} of game theoretic dynamics to those generated by my own model.

\section*{III. Toward an Optimal Discovery Mechanism}

\subsection*{A. Objectives and Features}

Having discussed in detail the causes of excessive discovery, I now attempt to derive a solution by drawing on the theory of mechanism design. Recall that the task of a mechanism designer is to define the rules of strategic interactions among economic actors, inducing these actors to behave in a manner that optimizes the designer’s measure of utility.\(^{65}\) In this case, the mechanism designer is a judge, encouraged by amended Rule 26(b) to stringently enforce proportionality in discovery. The judge’s goal is to choose precisely the amount of discovery that optimizes social utility, setting social marginal benefit equal to social marginal cost. To do this, however, the judge must overcome the challenges identified in Part II. \textit{A priori}, the judge does

---

\(^{63}\) See Kobayashi, \textit{supra} note 6.

\(^{64}\) See Klerman, \textit{supra} note 22, at 362-63.

\(^{65}\) See \textit{supra} notes 15-17 and accompanying text.
not know what the optimal amount of discovery might be, as the judge is much less familiar with the particulars of a given case than is the party that requests discovery. The judge must therefore elicit from this party some measure of how much that party values discovery. The judge cannot, however, accept uncritically the requesting party’s statements regarding its private information; for all the reasons discussed in Part II, the requesting party will overstate how much it values further discovery, pushing the total quantity above the social optimum.

A successful discovery mechanism must therefore address the requesting party’s incentives for strategic dishonesty. It must “debias” this party by taking into account the psychological biases encouraging the party to request excessive discovery. Moreover, it must force the requesting party to properly internalize the full social costs of additional discovery. Only then can the judge properly calculate the socially optimal amount.

What other properties should such a mechanism possess? First, it must be realistic and readily implementable. A judge does not have unlimited power to induce litigants to behave in a certain manner. Rule 26(c) explicitly permits judges to alter the allocation of discovery expenses, and so the model will rely on precisely that power to influence litigants’ behavior, rather than granting the judge some greater authority not envisioned in the FRCP. An optimal discovery mechanism only works, moreover, if it is simple enough to implement. Neither the judge nor the litigating parties should have to rely on overly burdensome calculations or difficult-to-estimate information; rather, the mechanism should involve a fairly simple procedure that transforms the requesting party’s value for discovery into an optimal quantity.

---

66 See supra note 36 for a number of works that make similar assumptions.
67 As discussed in note 83 infra and the accompanying text, the properties introduced in this section are frequently identified in the mechanism design literature as desirable characteristics of mechanisms.
68 FED. R. CIV. P. 26(c).
Next, parties should be incentivized to participate in the mechanism. Known in the mechanism design literature as *individual rationality*, this is a reasonability constraint that mandates non-negative expected utilities for all parties involved in the mechanism. Otherwise, a mechanism designer could simply extract an arbitrary amount of money from participants without promising anything in return—decidedly not a realistic picture of the American civil court system. Finally, the mechanism should require *no subsidy*. Courts (and by extension, taxpayers) do not currently subsidize litigants’ discovery motions, and they have little institutional capacity to do so. Any proposed discovery mechanism should insist that all costs of discovery be allocated among the litigants themselves, and moreover, that total discovery costs decrease as a result of the mechanism.

My model is the first to address all of these issues in broad generality. In Sections III.B-C, I formally derive and solve an optimal discovery mechanism. I then discuss in Section III.D its consequences and implications.

**B. Introducing the Model**

The participants in my model are a litigant-plaintiff and a judge in a civil action. I assume for clarity that the suit and discovery process are one-sided: an individual plaintiff has sued a company for negligence, claiming tort damages, and the only meaningful discovery is what the plaintiff might request from the defendant.69

---

69 This one-sided simplification of the discovery process is similar to the approach in Joel Sobel, *An Analysis of Discovery Rules*, 52 LAW & CONTEMP. PROB. 133 (1989) and to the discussions in Cooter & Rubenfeld, *supra* note 7. Since my model is concerned only with judicial management of discovery requests (as opposed to more general questions about the broader litigation process), I focus strictly on the game theoretic dynamics of each discovery request in isolation. I also note that, since the defendant’s actions do not play an explicit role in this model, I assume informally that any discovery costs are small enough compared to the claimed damages that the defendant is willing to move ahead with the litigation process. I thus sidestep the possibility of in terrorem settlements.
Discovery is a one-dimensional quantity $q \in \mathbb{R}^+ \cup 0$. It influences the plaintiff's utility, at least in expectation, by increasing either the probability of winning $P(q)$ or the amount the plaintiff would win assuming the court rules in his favor, $G(q)$. Therefore, the baseline expected additional winnings for the plaintiff given an amount of discovery $q$ is $P(q)G(q)$. However, since I assume both $P$ and $G$ are increasing in $q$, and since I do not further differentiate these functions or take a position on which one $q$ is most likely to affect, I define the function $v(q) = P(q)G(q)$, which I use for the remainder of the exposition. I further assume that $v(0) = 0$ (i.e., that a quantity of 0 discovery has no impact at all on the plaintiff’s expected winnings), that $v(q)$ is twice differentiable, and that $v'(q) > 0$ and $v''(q) < 0$. This captures the intuition that further discovery should help the plaintiff, but that it has diminishing marginal returns.

For any given quantity of discovery, the plaintiff faces the costs of processing new information, and the defendant faces the costs of searching for, identifying, and delivering all requested materials. For simplicity and clarity, I assume these costs are linear in $q$, so that the plaintiff faces costs $c_p q$ and the defendant faces costs $c_d q$, where $c_p > 0$ and $c_d > 0$. I assume moreover that $c_d \gg c_p$ to capture the intuition that discovery should be much more costly for the

---

70 In reality, there can be multiple types as well as magnitudes of discovery, but multidimensionality in mechanism design is generally a very hard problem. See TILMAN BÖRGERS, AN INTRODUCTION TO THE THEORY OF MECHANISM DESIGN 29 (2015). This one-dimensional structural assumption is very common in the theoretical literature on discovery. See, e.g., Steven Shavell, Sharing of Information Prior to Settlement or Litigation, 20 RAND J. ECON. 183, 185 (1989); Sobel, supra note 69, at 137; Joel L. Schrag, Managerial Judges: An Economic Analysis of the Judicial Management of Legal Discovery, 30 RAND J. ECON 305, 310 (1999).

71 Analogously to my setup, Cooter and Rubinfeld model the benefits of discovery as a concave function. See supra note 7, at 450-52.

72 A linear cost structure is not essential to my conclusions, though it allows for me to solve explicitly for the quantity of discovery as a function of its value to the plaintiff; I use linearity in my exposition because this explicit solution is helpful for understanding the intuition underlying my results. Indeed, my results hold for any cost structure that allows for a unique solution to the first order condition in part (a) of the proof of Proposition 2 infra, and which is increasing in the quantity of discovery requested. Any continuous monotonic cost function (subject to the regularity condition in the following paragraph), for example, works for the purposes of this model.
defendant company than for the individual plaintiff, and I assume that neither the plaintiff nor the defendant is budget-constrained.

I now introduce the plaintiff’s private information $\theta > 0$, which is known to the plaintiff but not to the judge. This parameter (generally called the plaintiff’s type in the mechanism design literature) represents the plaintiff’s (partially subjective, partially objective) value for discovery, so that the plaintiff’s expected gain from a quantity $q$ of discovery is $\theta v(q)$. A plaintiff’s net benefit from discovery is then $\theta v(q) - t(q)$, where $t(q)$ is the amount the plaintiff must pay for a quantity $q$ of discovery. Observe that a larger $\theta$ results in a larger absolute willingness to pay for discovery $\theta v(q)$ and a larger marginal willingness to pay $\theta v'(q)$ for all $q$. While the judge does not know $\theta$, he does know that the plaintiff draws $\theta$ from a probability distribution with compact support $[\overline{\theta}, \bar{\theta}]$, and that this distribution has a cumulative distribution function $F$ with density $f$. For technical reasons, I also assume $f(\theta) > 0$ for all $\theta \in [\overline{\theta}, \bar{\theta}]$. I then make the additional assumption that $\lim_{q \to \infty} \theta v'(q) < c_p$; since the highest-type plaintiff’s willingness to pay eventually falls below marginal cost, this regularity condition forces a solution with a finite amount of discovery for all plaintiff types.

Finally, I introduce the judge’s utility function. The judge seeks to maximize social welfare, or the difference between total social benefits and costs: $U(q) = \theta_j v(q) - c_p q - c_d q$, where $\theta_j$ represents the judge’s value for discovery. There are several key points in my setup of the judge’s utility function. Notice first that, in contrast to the plaintiff’s utility function, the

---

73 See Gelbach & Kobayashi, supra note 12, at 1103.
74 See supra note 36 and accompanying text.
75 Note that, prior to the FRCP Amendments, the plaintiff’s net benefit would be $\theta v(q) - c_p q$. This model allows the judge to choose a different function $t(q)$.
76 This probabilistic treatment of parties’ private information is very common in the mechanism design literature. See, e.g., Gelbach, supra note 21, at 210; Mnookin & Wilson, supra note 23, at 227-28 (applying a similar probabilistic treatment to their own mechanism design-based analysis of pretrial discovery).
judge takes into account both the plaintiff’s discovery costs and the defendant’s discovery costs. The judge thus internalizes all social costs, addressing the problem identified in Section II.B.

Next, observe that the plaintiff’s payment \( t(q) \) does not enter into the judge’s utility function. Since the judge seeks to maximize social welfare and not, for example, expected revenue from the litigants, the judge does not care how much the plaintiff actually pays. The plaintiff’s payment is purely a tool to induce optimal discovery.\(^7\) Lastly, the judge takes a “one-sided” view of social benefits, considering only the extent to which discovery benefits the plaintiff’s case, and not the extent to which it damages the defendant’s case. I define the Rule 26(b)(1) proportionality standard in this manner on the grounds that any action taken by a litigant should always attempt to benefit his own case to the detriment of his opponent. Given the adversarial system of litigation in the United States, it would make little sense for a judge to limit discovery on the grounds that discovery might reduce an opponent’s probability of winning, or increase the amount an opponent might be forced to pay should he lose. The only costs for which the judge should control are those “external” to the core processes of litigation, which for purposes of this discussion are the direct costs associated with realizing a certain quantity of discovery.

The parameter \( \theta_j \) requires particular attention, as it is through this parameter that I capture the impact of the plaintiff’s behavioral economic biases. Similar to my setup for the plaintiff’s parameter \( \theta \), I assume that \( \theta_j \) represents the realization of a random variable with compact support \([\underline{\theta}, \overline{\theta}]\), a cumulative distribution function \( F_j \), and a density function \( f_j \) that is strictly positive over \([\underline{\theta}, \overline{\theta}]\). While the judge lacks certain case-specific knowledge and therefore

\(^7\) This is a standard assumption in a mechanism design problem for a public good, where the mechanism designer is concerned only with social efficiency and not with the distribution of payments among the participants. See, e.g., d’Aspremont & Gérard-Varet, supra note 16. An extension of this model might consider the implications of a court that demands some payment for its adjudication, and thus has an additional profit-maximizing motive. Since I assume court fees are small compared to the amount at stake in the litigation, however, I do not explore this question.
does not know the true value of $\theta$, having adjudicated a large number of similar cases, the judge
has developed a good estimate of the distribution from which $\theta$ should have been drawn by a
“rational” litigant.\textsuperscript{78} This distribution is $F_j$, and the extent to which $\theta$ exceeds the value it should
have taken were it drawn using $F_j$ is a measure of the plaintiff’s psychological bias toward
excessive discovery.\textsuperscript{79} Formally, we say that $F$ stochastically dominates $F_j$ in the sense that
$F(\theta) \leq F_j(\theta)$ for all $\theta \in [\theta, \bar{\theta}]$. The judge can accordingly transform $\theta$ into $\theta_j$ by computing
$\theta_j = F_j^{-1}(F(\theta))$, thus addressing the issues identified in Section II.A.\textsuperscript{80}

C. Solving the Model

A solution to the model described above is a general mechanism, subject to the
conditions of individual rationality and no subsidy, that maximizes social welfare when played
by the plaintiff.\textsuperscript{81} To tackle such a tremendously broad and unwieldy problem, I invoke the
revelation principle, which allows me to focus without loss of generality on a much smaller
subset of mechanisms called direct mechanisms.\textsuperscript{82}

For the purposes of this model, a direct mechanism consists of a function
$q: [\theta, \bar{\theta}] \rightarrow \mathbb{R}^+ \cup 0$ and a function $t: [\theta, \bar{\theta}] \rightarrow \mathbb{R}$, and the plaintiff’s corresponding strategy is a
function $\sigma: [\theta, \bar{\theta}] \rightarrow [\theta, \bar{\theta}]$. Such a mechanism has the following interpretation: the plaintiff

\textsuperscript{78} The judge does have some information from which to make this assessment. He will have read the parties’
preliminary briefs and other pre-discovery filings (including, for example, any motions for summary judgment or
other motions to dismiss).

\textsuperscript{79} See Figure 1 \textit{infra} for a graphical illustration of this phenomenon.

\textsuperscript{80} To the best of my knowledge, this paper is the first to use stochastic dominance in this setting to characterize
behavioral economic biases.

\textsuperscript{81} Strictly speaking, the plaintiff plays the game of incomplete information specified by the general mechanism. To
avoid unnecessary exposition, however, I do not formally define either general mechanisms or formal games.

\textsuperscript{82} For a formal statement and proof of the revelation principle, see BÖRGERS, supra note 70, at 10. In Börgers, $q$
represents a probability, whereas in this model, it represents a quantity. This does not in any way affect the proof.
declares a type $\theta$, pays a value $t(\theta)$ to the court, and receives in return a quantity of discovery $q(\theta)$.

The revelation principle implies that any social optimum achievable under a general mechanism is also achievable simply by choosing functions $q(\theta)$ and $t(\theta)$ in a direct mechanism, such that the plaintiff always tells the truth about his own type. This last condition is called incentive compatibility.

Having settled on direct mechanisms as a solution for the model, I can now formally define the other key constraints of individual rationality and no subsidy, and I can restate the definition of incentive compatibility in a more useful form.

Individual rationality, the idea that all plaintiff types must be willing to participate in the mechanism, is the condition that $\theta v(q(\theta)) - t(\theta) \geq 0 \ \forall \ \theta \in [\theta, \bar{\theta}]$, or in other words, that all plaintiff types have non-negative utility from participating.

The no subsidy condition requires that the court not pay for any part of the plaintiff’s discovery efforts, and that total discovery costs decrease as result of the mechanism. Since I impose no requirements about the distribution of costs between the plaintiff and the defendant, we guarantee the no subsidy condition so long as total discovery costs decrease. Thus, we require that $qc_p + qc_d < q_{ni}c_p + q_{ni}c_d$, where $q_{ni}$ is the amount of discovery borne by the parties.

---

83 The idea of the plaintiff declaring his type may seem somewhat less natural than directly requesting a certain quantity of discovery, but this setup merely assists with the exposition and does not detract from the feasibility of the model. The final solution implicitly defines $t$ as a function of $q$, or in other words, allows the judge to set a price $t$ for any quantity $q$ that the plaintiff might request. This is how such a mechanism would likely be implemented in practice.

84 The three conditions of individual rationality, no subsidy, and incentive compatibility are commonly identified in the mechanism design literature as desirable properties of a mechanism; they feature prominently in, for example, Myerson and Satterthwaite’s famous 1983 impossibility theorem. See supra note 17. In this case, it turns out that all three conditions are satisfied by the social welfare-maximizing mechanism I construct. Gelbach similarly identifies these three conditions as properties of his own discovery mechanism. See Gelbach, supra note 21, at 218-19 (describing his model as a “Nirvana mechanism” for having these key properties, though noting that informational constraints may limit our ability to realize these properties in practice).

85 After all, a plaintiff could simply choose not to participate in the mechanism at all (i.e., request no discovery), and thus guarantee himself a payoff of 0. Notice that I do not restrict the mechanism by requiring individual rationality for the defendant; surely, if the defendant had his way, he would not participate in the lawsuit in the first place!
under a non-interventionist regime. Because of the linear structure of discovery costs, the no subsidy condition is fulfilled if the optimal discovery mechanism results in a net decrease in discovery.\(^{86}\)

Finally, I restate incentive compatibility in the following (obviously equivalent) form:

\[
\theta v(q(\theta)) - t(\theta) \geq \theta v(q(\hat{\theta})) - t(\hat{\theta}), \text{ where } \hat{\theta} \text{ is any (potentially dishonest) type that the plaintiff might report to the judge. Standard arguments from mechanism design, which I reproduce for completeness in the Appendix, show that incentive compatibility is equivalent to two joint conditions.}
\]

**Proposition 1**

A direct mechanism is incentive compatible if and only if:

(i) \(q\) is (weakly) increasing in \(\theta\), and

(ii) for every \(\theta \in [\theta_1, \theta_2]\),

\[
t(\theta) = t(\theta) - \theta v(q(\theta)) + \theta v(q(\theta)) - \int_{\theta_1}^{\theta} v(q(x))dx
\]

*Proof:* See Appendix.

*Note:* For the purposes of this model, Proposition 1 is purely a technical result. Its importance lies in the fact that the mechanism designer’s choice is now simplified even further: so long as \(q\) is increasing, the mechanism is uniquely determined by the choice of \(q\) and the utility of the lowest type, \(\theta v(q(\theta)) - t(\theta)\).

---

\(^{86}\) As noted in note 72, this conclusion follows for any continuous monotonic cost structure, not just a linear one.
At last, I turn to the judge’s optimization problem. We recall that the judge’s utility function is \( \theta_j v(q) - c_p q - c_d q = F_j^{-1}(F(\theta))v(q) - c_p q - c_d q \). But since the judge does not know the plaintiff’s private information \( \theta \), the judge must choose \( q \) to maximize expected social utility (subject to individual rationality, no subsidy, and incentive compatibility), given by the following expression:

\[
\mathbb{E} \left[ \int_{\theta} \left[ F_j^{-1}(F(\theta))v(q) - c_p q - c_d q \right] f(\theta) d\theta \right].
\]

If we are able specify \( q \) (and \( t(\theta) \)) subject to the three aforementioned constraints, Proposition 1 guarantees that we have identified a direct mechanism that achieves the socially optimal amount of discovery.

**Proposition 2**

Let

\[
q(\theta) = \left[ v' \right]^{-1} \left( \frac{c_p + c_d}{F_j^{-1}(F(\theta))} \right)
\]

and

\[
t(\theta) = \theta v \left( q(\theta) \right).
\]

These functional specifications maximize expected social utility subject to (a) incentive compatibility, (b) individual rationality, and (c) no subsidy.

---

\(^{87}\) See supra text accompanying note 77.
Proof of (a): Expected social utility maximization and incentive compatibility

Incentive compatibility requires that \( q \) be increasing and that \( t \) satisfy the functional form given in Proposition 1. Since \( t \) does not enter into the judge’s utility function, we take the functional form of \( t \) as given and we maximize the utility function only with respect to \( q \). We temporarily disregard the requirement that \( q \) be increasing; we can then maximize the judge’s utility function pointwise by choosing \( q \) to optimize social utility for each \( \theta \) separately. Accordingly, we derive the relevant first order condition by differentiating the bracketed part of the integrand:

\[
F_j^{-1}(F(\theta))v'(q(\theta)) - c_p - c_d = 0,
\]

and so

\[
q(\theta) = [v']^{-1}\left(\frac{c_p + c_d}{F_j^{-1}(F(\theta))}\right),
\]

as desired. Given the properties of \( F \), \( F_j \), and \( v' \) discussed in Section III.B, it is readily verified that \( q(\theta) \) is indeed increasing in \( \theta \), and so this choice of \( q \) maximizes expected social utility subject to incentive compatibility. To avoid degenerate cases, I assume \( v \) is such that \( \theta v(q(\theta)) - c_p q - c_d q \) is non-negative for all \( \theta \in [\theta, \bar{\theta}] \), meaning the judge should never deviate from the above choice of \( q \).

Note: This expression for \( q \) behaves in the manner we would intuitively expect given the objectives of our mechanism. It is decreasing in the costs of both the plaintiff and the defendant, since our mechanism forces the plaintiff to internalize both parties’ costs when requesting discovery. It is increasing in the plaintiff’s type, since a higher type more highly values discovery and therefore would request a greater amount of it. And finally, it is decreasing in the
functional composition of $F_{j}^{-1}$ and $F$. This term should be understood as “shifting back” $\theta$ by some amount; the more $\theta$ is shifted, the more the court corrects for the plaintiff’s psychological biases. I illustrate this effect in the below figure, where the length of the shift (the black arrow) is a measure of the plaintiff’s biases. We therefore see that $q$ accounts for, and offsets, the plaintiff’s classical and behavioral economic incentives for excess discovery, as desired.

![Figure 1: I depict the relationship between $F$ and $F_{j}$ and demonstrate how the judge “shifts” the plaintiff’s type to offset the plaintiff’s behavioral economic biases.](image)

**Proof of (b): Individual rationality**

Satisfying individual rationality is straightforward. As I show in the Appendix, the plaintiff’s utility function $\max_{\theta \in [\underline{\theta}, \bar{\theta}]} \left[ \theta v(q(\hat{\theta})) - t(\hat{\theta}) \right]$ is differentiable almost everywhere\(^{88}\) with derivative $v(q(\theta))$. Since both $q$ and $v$ are non-negative and strictly increasing, this implies that the plaintiff’s utility is increasing in his type. It therefore suffices to ensure individual rationality for the lowest type, which we do by setting $t(\theta) = \theta v(q(\hat{\theta}))$. Note that this also fully defines the payments made by all plaintiff types: $t(\theta) = \theta v(q(\theta)) - \int_{\underline{\theta}}^{\theta} v(q(x))dx$, with $q$ defined as

\(^{88}\) I use this expression in the technical sense, with respect to the Lebesgue measure.
above. The now-complete specification of \( q \) and \( t \) completes the derivation of the optimal discovery mechanism.

Note: Observe that our choice of \( t(\theta) \) sets the utility of the lowest type equal to 0, making him indifferent about participating in the mechanism. Since our other assumptions imply that the lowest-type plaintiff gains the least utility from discovery, this choice of \( t(\theta) \) therefore guarantees that all types have non-negative utility. Other choices of \( t(\theta) \) are theoretically possible, but the choice of some arbitrary positive value for the minimum utility is exogenous to the model.

\textit{Proof of (c): No subsidy}

Finally, I show that this mechanism results in a net decrease in discovery costs compared to a regime of no judicial intervention, and therefore requires no subsidy from the court.\(^{89} \) Under the traditional discovery paradigm, the plaintiff would maximize \( \theta v(q_{ni}) - c_p q_{ni} \).\(^{90} \) In this non-interventionist regime, the plaintiff simply compares his private benefits and costs: he only takes into consideration his own costs for discovery, and he does not account for his behavioral economic biases. We differentiate this expression and solve the first order condition, which implies that the plaintiff chooses \( q_{ni} = [v']^{-1} \left( \frac{c_p}{\theta} \right) \). We now compare \( q_{ni} \) to the expression for optimal discovery derived above, reprinted here for convenience:

\(^{89} \) \textit{See supra} note 86 and accompanying text.  
\(^{90} \) \textit{See supra} note 75.
Discovery requested without judicial intervention

\[ q_{ni}(\theta) = [v']^{-1}\left(\frac{c_p}{\theta}\right) \]

Discovery requested under the proposed mechanism

\[ q(\theta) = [v']^{-1}\left(\frac{c_p + c_d}{F_j^{-1}(F(\theta))}\right) \]

Figure 2: A comparison of quantities of discovery requested under non-interventionist and interventionist regimes of civil discovery.

As expected, the mechanism design approach has materially decreased the amount of discovery requested. The numerator is much smaller since \( c_p < c_p + c_d \), and in particular since \( c_p \ll c_d \).

The denominator is also larger: because \( \theta \) stochastically dominates \( \theta_j \), \( \theta \geq F_j^{-1}(F(\theta)) \). Since \( v' \) and therefore \( [v']^{-1} \) are decreasing functions, we conclude that the equilibrium quantity of discovery is strictly smaller using the optimal discovery mechanism, resulting in correspondingly lower costs for the parties.

Note: These contrasting expressions for \( q_{ni} \) and \( q \) reinforce the discussion following the proof of part (a) of Proposition 2. Observe that, the higher the defendant’s costs, the less discovery the plaintiff requests as compared to the traditional discovery paradigm, since the plaintiff must now internalize these costs. Similarly, the more \( \theta \) exceeds \( F_j^{-1}(F(\theta)) \)—or, in other words, the more the plaintiff’s behavioral economic biases cause his perceptions to depart from those of a purely “rational” economic actor—the less discovery the plaintiff requests under the new mechanism as compared to the old system.

D. Discussion

At first blush, the new discovery mechanism appears extremely promising. It addresses all of the most common problems perceived to affect discovery today. First, judges have great flexibility to address behavioral economic biases by estimating a distribution from which a
“rational” plaintiff’s value for discovery should be drawn. Moreover, judges incorporate all social costs of discovery, balancing the benefits to the plaintiff with the potentially enormous discovery costs faced by the defendant. Collectively, these factors materially lower the equilibrium level of discovery, resulting in lower costs for the collective parties and requiring no subsidies from the court. And the mechanism achieves this goal in a relatively straightforward manner. The plaintiff’s strategy is simple and requires no involved calculations: he need only reveal his true value for discovery. This corresponds to a unique quantity of discovery and corresponding payment, which the judge can reallocate among the litigants in any manner the judge chooses.

That said, there are some limitations to the model, of which informational concerns are the most prominent. How realistic is common knowledge about the various parameters of the model? It seems reasonable that both the plaintiff and the judge should know $c_p$ and $c_d$. The hourly rate for document review—or the effective hourly rate, for lawyers taking the case on contingency—should not vary enormously among legal teams with comparable resources and a comparable degree of sophistication. The plaintiff obviously knows $c_p$, the judge can either make an informed estimate of $c_p$ or research the prevailing rate, and both the judge and the plaintiff should be able to research the prevailing value of $c_d$. The “objective” component of the benefit to the plaintiff, $v(q)$, is more difficult to assess. Parties can reasonably disagree about how much a given quantity of discovery is likely to improve the plaintiff’s outcome. The model mitigates this somewhat by treating $v(q)$ as just the expected improvement due to discovery, but it may still not be entirely realistic to treat $v(q)$ as common knowledge. Finally, we turn to the distributions of $\theta$ and $\theta_j$. It is very reasonable that a judge, having overseen a large number of similar cases and having been exposed to preliminary briefs from the litigants, can come up with
a reasonable estimate of $F_j$. But it is far harder to justify why $F$ should be common knowledge—how should the judge know the precise nature of the behavioral economic biases influencing the plaintiff? Even if the judge could quantify those biases, it is far from obvious how the judge might translate them into a probability distribution, let alone one that agrees with the distribution the plaintiff has in mind.

The reader should note two crucial points about this discussion of informational constraints. First, it illustrates a sense in which behavioral economic biases may be more difficult to mitigate than the impact of classical economic incentives on excess discovery. Whereas a judge can readily estimate classical externality costs, and perhaps even the distribution of $\theta$ for a purely rational litigant, the impact of psychological biases should be considerably more difficult to measure.

Second, while information requirements are indeed a limitation to my model, such concerns are necessarily present in any treatment of the FRCP proportionality standard. As a preliminary matter, the objection that the distribution of private information should not be common knowledge is common across all classical mechanism design models.91 In discussing his own model of civil discovery, for example, Gelbach also identifies informational concerns as the largest practical impediment to the adoption of his mechanism.92 But this observation does not invalidate mechanism design as an approach. As one observer noted in reviewing Gelbach’s model, “Gelbach’s deepest point is that the informational requirements are not specific to mechanism design. . . . Law’s instructions to legal decision-makers implicitly require a lot of information. For example, the new FRCP 26(b)(1) implicitly assumes that the judge in any given

---

91 See Dirk Bergemann & Stephen Morris, *An Introduction to Robust Mechanism Design*, 8 FOUND. & TRENDS MICROECONOMICS 169, 172-73 (2012). The very active field of robust mechanism design, developed in large part by Bergemann and Morris, is an attempt to reduce or eliminate reliance on such strong informational assumptions.

92 See supra note 21, at 204
case has the information required to assess what is ‘proportional to the needs of the case.’”

Thus, what distinguishes my approach from models not based in mechanism design—and indeed, from the generic language calling for proportionality in FRCP 26(b)(1) itself—is that only the mechanism design approach “explicitly makes the best of whatever information [we] assume the actors to have.” Mechanism design does not create informational requirements, but rather makes explicit the requirements already present in Rule 26(b)(1).

The other major change this mechanism introduces compared to the current practice of discovery is the allocation of discovery costs. Whereas today the defendant bears the vast majority of discovery costs, this discovery mechanism envisions the plaintiff now covering a large portion of the costs. This may raise questions about access to justice. Consider, for example, situations in which a single plaintiff is suing a large company-defendant. It is certainly conceivable that such a plaintiff would be unable to finance the enormous costs of the discovery needed to win the case. Such worries, however, are less pressing nowadays than they may have been some decades ago. The last twenty years have seen the emergence of a well-capitalized plaintiffs’ bar; plaintiffs’ firms should be willing to invest money upfront and pay for discovery that will lead to a sufficiently large payoff later on. Moreover, the tremendous growth of litigation finance companies over the past fifteen years has led to a robust market for financing lawsuits. And concerns about resource constraints largely do not apply to situations in which well-capitalized corporations sue each other, suggesting a natural set of cases in which this new discovery mechanism might be readily introduced.

---

93 Holger Spamann, *Can Simple Mechanism Design Results be Used to Implement the Proportionality Standard in Discovery?: Comment*, 172 J. INSTITUTIONAL & THEORETICAL ECON. 227, 228 (2016).

94 Id.


IV. Conclusion

The 2015 FRCP Amendments envision a much higher degree of judicial involvement in the civil discovery process. In response to the widely perceived problem of excessive discovery, the amended rules insist on a proportionality standard, whereby judges balance all parties’ interests in deciding whether to allow a given discovery request. This task is difficult: a judge has limited knowledge about the potential value of a given amount of discovery, and he cannot fully rely on litigants’ own reports of the value of discovery. Because of behavioral economic biases and externalized discovery costs, litigants have a strong incentive to over-report their value of discovery.

Even so, using the techniques of mechanism design, I show it is in fact possible to improve the civil discovery process. Through a relatively simple procedure, judges can choose the amount of discovery that optimizes social welfare. This mechanism corrects for litigants’ misaligned incentives without any subsidy, and it encourages litigants of all types to participate in this new system. Although limited somewhat by informational constraints, this mechanism offers a compelling blueprint for judges seeking to implement the 2015 FRCP Amendments and thus dramatically reduce the prevalence and costs of excessive discovery.
Appendix

I present an abbreviated proof of Proposition 1. The core ideas in this proof are not novel and can be found in any textbook on mechanism design. However, since I refer to these ideas elsewhere in Part III, I include the proof for completeness.

We seek to show that incentive compatibility is equivalent to conditions (i) and (ii) in Proposition 1. We begin by showing that incentive compatibility implies \( q \) is increasing in \( \theta \).

Consider any two \( \theta \) and \( \hat{\theta} \) with \( \theta \geq \hat{\theta} \). From the definition of incentive compatibility,

\[
\theta v(q(\theta)) - t(\theta) \geq \theta v(q(\hat{\theta})) - t(\hat{\theta}) \quad \text{and} \quad \hat{\theta} v(q(\theta)) - t(\theta) \leq \hat{\theta} v(q(\hat{\theta})) - t(\hat{\theta}).
\]

Subtracting these inequalities and simplifying, we find \( v(q(\theta)) \geq v(q(\hat{\theta})) \), and so \( q(\theta) \geq q(\hat{\theta}) \) as desired.

For condition (ii), we examine the plaintiff’s general utility function

\[
u(\theta) = \max_{\hat{\theta} \in [\theta, \bar{\theta}]} \left[ \theta v\left(q(\hat{\theta})\right) - t(\hat{\theta}) \right].
\]

Observe that \( u(\theta) \) is affine, and therefore convex, for any choice of \( \hat{\theta} \). The maximum of a family of convex functions is convex, and convex functions are differentiable almost everywhere. By incentive compatibility, this maximum occurs at \( \hat{\theta} = \theta \), so that \( u(\theta) = \theta v(q(\theta)) - t(\theta) \). At any point where \( u(\theta) \) is differentiable, direct computation of the right- and left-hand derivatives \( \lim_{\delta \to 0} \frac{u(\theta+\delta)-u(\theta)}{\delta} \) and \( \lim_{\delta \to 0} \frac{u(\theta)-u(\theta-\delta)}{\delta} \) (respectively) reveals that \( u'(\theta) = v(q(\theta)) \). Finally, since \( u(\theta) \) is convex, it is also absolutely continuous and therefore equals the integral of its derivative: \( u(\theta) = u(\underline{\theta}) + \int_{\theta}^{\theta} v(q(x))dx \). Substituting \( u(\theta) = v(q(\theta)) - t(\theta) \) and rearranging yields the desired expression.

\[97\] See, for example, BÖRGERS, supra note 70, at 11-14, upon which this version of the proof is based.
A short computation reveals that these two conditions are also sufficient to imply incentive compatibility. Consider some $\theta$ and $\hat{\theta}$ with $\theta > \hat{\theta}$. If $q$ is increasing, then it must be true that $\int_{\theta}^{\hat{\theta}} v(q(x))dx \geq \int_{\theta}^{\hat{\theta}} v(q(\hat{\theta}))dx = (\theta - \hat{\theta})v(q(\hat{\theta}))$. From condition (ii), we know that $\int_{\theta}^{\hat{\theta}} v(q(x))dx = u(\theta) - u(\hat{\theta})$, so $u(\theta) - u(\hat{\theta}) \geq (\theta - \hat{\theta})v(q(\hat{\theta}))$. Substituting $u(\hat{\theta}) = \hat{\theta}v(q(\hat{\theta})) - t(\hat{\theta})$ and rearranging, we see that $u(\theta) \geq \theta v(q(\hat{\theta})) - t(\hat{\theta})$, the definition of incentive compatibility. Analogous reasoning applies for $\theta < \hat{\theta}$.