

Back to the Future: Discovery Cost Allocation and Modern Procedural Theory

*Martin H. Redish**
*Colleen McNamara***

ABSTRACT

This Article questions the current model of discovery cost allocation, which requires the producing party to bear the expenses associated with its opponent's discovery requests. The current presumption forces the producing party to fund a portion of its opponent's case, transforming discovery costs into a de facto litigation subsidy. This hidden subsidy not only unjustly enriches the requesting party, but it also gives rise to troubling implications from the perspectives of democratic, economic, and constitutional theory. To remedy this foundational problem with the prevailing presumption, this Article presents two constructive proposals. The first proposal suggests a solution: rewriting a new presumption for discovery cost allocation on a clean procedural slate. The second proposal explains how judicial attitudes toward discovery cost allocation can be reshaped under the current procedural framework.

INTRODUCTION

To suggest that discovery costs have been the subject of substantial scholarly and judicial controversy would be an understatement. Over the years, numerous amendments to the Federal Rules of Civil Procedure have been promulgated in an effort to curb discovery costs,¹ and more recently the Supreme Court appeared to toughen the pleading standards expressly because of the burdensome costs that result when vague allegations are allowed to proceed to the discovery stage.²

* Louis and Harriet Ancel Professor of Law and Public Policy, Northwestern University School of Law.

** J.D., expected May 2011, Northwestern University School of Law; A.B., 2003, Williams College.

¹ Of particular relevance are the amendments in 1980 (adding the discovery conference), FED. R. CIV. P. 26 advisory committee's note, at 66 (1980); 1983 (adding, among other things, the certification requirement of Rule 26(g) and the proportionality requirement of Rule 26(b), as well as substantially modifying the discovery conference procedure), *id.* advisory committee's note, at 717 (1994); 1993 (adding automatic disclosure and imposing it ex ante on the use of specific discovery devices), *id.* advisory committee's note, at 718; and 2003 (modification of automatic disclosure), *id.* advisory committee's note, at 172 (2006).

² See *Bell Atl. Corp. v. Twombly*, 550 U.S. 544, 558–59 (2007) (discussing discovery costs

At the time of the Federal Rules' adoption in 1938, however, apparently no attention, at any level of the process, was devoted to the question of to which party, in the first instance, the cost of discovery should be attributed. It appears that it was widely—if only implicitly—assumed that discovery costs were to remain where they fell: a party required to produce discovery requested by another party was—and to this day continues to be—assumed to bear whatever costs it incurred in the course of that production. Thus, since the adoption of the Federal Rules in 1938, the allocation of discovery costs has been governed by the presumption that the party from whom the information is sought—the producing party—must bear the expenses associated with the fulfillment of its opponent's discovery requests.³ Recently, however, several courts have begun to chip away at this presumption. Backed by substantial urging from scholars,⁴ a few pioneering judges have ordered the requesting parties to bear some of the production costs associated with their requests for electronically stored information.⁵

Still, the presumption in favor of imposing on the producing party the costs necessary to respond to its opponent's requests remains largely intact, if not virtually sacrosanct.⁶ Indeed, what is most surprising about the twenty-year debate over rising discovery costs and abuse is the collective failure on the part of most scholars and judges to question the theoretical foundations of our current model of discovery cost allocation. Even the staunchest advocates of cost shifting have designed their proposals around the governing presumption, leaving the inertia in favor of letting costs lie where they fall and, as a result, allowing current cost allocation rules to remain intact in the large majority of cases.⁷

as a contributing factor to the Court's decision to overrule the "no set of facts" test in *Conley v. Gibson*, 355 U.S. 41 (1957), which required complaints to be dismissed for failure to state a claim only if it appeared beyond doubt that the plaintiff could prove no set of facts in support of his claim); see also *Ashcroft v. Iqbal*, 129 S. Ct. 1937, 1950 (2009) (noting that Rule 8 "does not unlock the doors of discovery for a plaintiff armed with nothing more than conclusions").

³ *Oppenheimer Fund, Inc. v. Sanders*, 437 U.S. 340, 358 (1978) ("[T]he presumption is that the responding party must bear the expense of complying with discovery requests . . .").

⁴ See *infra* note 41 and accompanying text.

⁵ See *infra* Part I.

⁶ In the leading case involving cost shifting, *Zubulake v. UBS Warburg LLC*, 217 F.R.D. 309, 320 (S.D.N.Y. 2003), Judge Scheindlin urged courts to resolve close cases in favor of the governing presumption against cost shifting.

⁷ This description includes one of the authors. See Martin H. Redish, *Electronic Discovery and the Litigation Matrix*, 51 DUKE L.J. 561, 608 (2001) (advocating cost shifting for requests involving electronic information that is stored in a format not reasonably accessible by the producing party). The *Zubulake* court adopted a similar limitation on cost shifting in electronic

In part, at least, this near-total absence of attention to the cost allocation issue likely derives from the inertia that began with the original adoption of the Federal Rules in 1938. The drafters of the original Federal Rules failed to seriously consider moral, economic, or democratic first principles when they apparently assumed, without discussion, that producing parties, rather than requesting parties, would bear the costs of discovery. Indeed, the Advisory Committee on Civil Rules (“Advisory Committee”) seemed more concerned with a different kind of abuse—overreaching oral depositions and unconstrained “fishing expeditions.”⁸ Concerns about excessive discovery costs appeared only at the margins of the debate,⁹ but even these concerns were quickly subsumed by the drafters’ overwhelming conviction that the accumulation of more information by both the parties would lead to a more accurate determination of the merits of the underlying claim.¹⁰

Admittedly, it would have been difficult for the drafters of the original Federal Rules to foresee the advent of the copy machine, let alone the explosion of electronic discovery that occurred at the turn of this century—developments that have no doubt dramatically raised the discovery stakes.¹¹ It is nevertheless surprising that these eminent

discovery. *Zubulake*, 217 F.R.D. at 320; see also Robert D. Cooter & Daniel L. Rubinfeld, *An Economic Model of Legal Discovery*, 23 J. LEGAL STUD. 435, 455–56 (1994) (proposing a two-part rule wherein the defendant would “bear the costs of reasonable compliance up to a level deemed appropriate for this class of cases, beyond which the reasonable costs of complying with further discovery requests would shift to the plaintiff”).

⁸ See Stephen N. Subrin, *Fishing Expeditions Allowed: The Historical Background of the 1938 Federal Discovery Rules*, 39 B.C. L. REV. 691, 724 (1998).

⁹ Judge Edward Finch of the New York Court of Appeals argued that absent some provision for reimbursement of discovery costs,

a poor litigant is at the mercy of a richer opponent. Thus a wealthy defendant may force a poor plaintiff with a meritorious claim to accept an unfair settlement for the reason that the further the litigation is carried the greater the expense which the plaintiff must incur, which continually reduces the amount of his actual recovery unless he ultimately can obtain reimbursement for the costs to which he has been subjected by the unjust defense.

Edward R. Finch, *Some Fundamental and Practical Objections to the Preliminary Draft of Rules of Civil Procedure for the District Courts of the United States*, 22 A.B.A. J. 809, 811 (1936). The chairman of the Advisory Committee, William Mitchell, responded to this critique by accusing the courts of New York City of being filled with lawyers of “low ethical standards” who would bring “many dishonest actions,” while assuring the audience that in the rest of the country, “the rules relating to these subjects are in line with modern enlightened thought on the subject and will not be subjected to abuse.” William D. Mitchell, *Some of the Problems Confronting the Advisory Committee in Recent Months—Commencement of Actions—Effect of Findings of Fact in Cases Tried by Court Instead of Jury, Etc.*, 23 A.B.A. J. 966, 969 (1937).

¹⁰ See Subrin, *supra* note 8, at 739–40.

¹¹ See Maurice Rosenberg, *Federal Rules of Civil Procedure in Action: Assessing Their*

scholars, jurists, and practitioners failed to recognize and incorporate a mechanism to avoid the profound moral and economic problems inherent in their unstated assumption concerning the allocation of discovery costs.

In light of both this fundamental failure by the original drafters of the Federal Rules and the near-total avoidance of the issue by scholars, an analysis of the underlying principles of discovery cost allocation is sorely needed. Such an analysis reveals that this presumption has itself likely given rise to many of the problems that continue to plague the operation of our discovery system in at least certain categories of suits.

More than seventeen years ago, in a response to a cost-shifting proposal made by two legal economists, Professor Edward Cooper asked why the authors would require payment of discovery costs incurred by the defendant, regardless of the outcome of the case, when the plaintiff is not obliged to pay other costs incurred by the defendant.¹² This Article represents the first scholarly attempt to tackle that foundational question. It is our belief that enormously important differences exist between discovery costs incurred in responding to an opponent's discovery requests on the one hand, and the normal costs inherent in preparing a legal defense on the other. This belief leads to the conclusion that the traditionally accepted assumption about the allocation of discovery costs (an assumption, it should be noted, that has nowhere been expressly articulated in rule or statute, much less justified) is wrongheaded as a conceptual matter and disastrous as a practical matter.

Before beginning our analysis, we should make clear our recognition of the practical futility of any effort to reverse the long-established assumption that discovery costs are, at least as a presumptive matter, properly attributable to the producing party. We cannot, of course, go back in time to alter either what we consider an unfortunate mistake at the Federal Rules' creation or the consistent failure of courts or scholars to question it since. Such an effort would seriously tinker with the spacetime continuum—if, indeed, it were even possible

Impact, 137 U. PA. L. REV. 2197, 2208 (1989) (“In the 1930s, when the Rules were conceived and written, and for decades afterward, . . . [t]here was by today's standards no crisis of volume to draw [the drafters'] attention to systemic problems. A 'retail' view of the administration of justice was a natural approach. And because large, complex, protracted cases were uncommon, the monolithic, uniformly-applicable Rules often did not appear inadequate for purposes of controlling the movement of large, complicated law suits through the courts.”).

¹² Edward H. Cooper, *Discovery Cost Allocation: Comment on Cooter and Rubinfeld*, 23 J. LEGAL STUD. 465, 473 (1994).

as a matter of physics. Still, the advantage of scholarly inquiry is that it is not restricted by even well-settled practice or precedent, the way a brief in court is. Scholars thus have the luxury of being able to think outside the box, questioning previously unquestioned first principles. Moreover, while it would be unrealistic to believe that our foundational reconsideration of this core assumption about the discovery process will result in its abandonment, it does not follow that the value of such reconsideration would be purely academic. Although courts are, concededly, highly unlikely to reverse the traditional presumption of cost allocation, acceptance of our arguments could well lead to a significant increase in judicial willingness to shift discovery costs—with, most likely, very positive results for the litigation system.

Our argument is quite simple. If one strips away the long accepted assumption as to how the American system allocates costs among litigants, the actions of the parties to a lawsuit in the discovery process would be most appropriately seen as analogous to a quasi-contractual relationship between the adversary litigants. Under the theory of quantum meruit, a party to a quasi contract is legally entitled, as a matter of fundamental principles of economic justice, to be reimbursed for any benefit he confers on another person at that person's express or implied request. Since the seventeenth century, courts have consistently recognized that it is unjust to allow one person to profit at the expense of another, even when no formal, legally enforceable underlying agreement exists between them. We therefore liken the discovery process to a quasi contract and argue that it is morally untenable to allow the requesting party to retain the benefit of its opponent's labor without, at the very least, reimbursing the costs of discovery incurred by the producing party.¹³

It is true that the analogy is by no means perfect. In the classic quasi-contractual context, the parties act of their own free will. In contrast, in the discovery context, the producing party acts in the shadow of potential Rule 37 sanctions if it fails to produce requested discovery.¹⁴ Its actions are therefore hardly voluntary, in any meaningful sense of the term. Moreover, unlike the paradigmatic quasi-contractual situation, the value of efforts by the producing party on behalf of the requesting party extends well beyond merely the costs incurred in the production. But if anything, these distinctions only tend to exacerbate the inherent unfairness of failing to recognize the

¹³ See *infra* Part II.

¹⁴ FED. R. CIV. P. 37.

direct applicability of the quasi-contractual model to the discovery cost allocation decision.

Because the costs incidental to discovery production are, morally and economically, properly attributable to the requesting party, allocation of discovery costs to the producing party effectively transforms discovery costs into a litigation subsidy, which requires a party to fund a portion of its opponent's case. Part I describes the presumption embedded in the current system that gives rise to this litigation subsidy. Part II explains the troubling democratic, economic, and constitutional implications of this hidden subsidy. Part III presents two sets of proposals for the reallocation of discovery costs in accordance with the first principles shaped earlier in the Article. The first set of proposals reconstructs the procedural system's model of cost allocation from scratch. The second set of proposals, by contrast, suggests how this Article's discussion of first principles could be incorporated into ongoing discovery disputes. Even though the current procedural system might preclude judges from embracing the full scope of the model presented here, the analysis can and should be incorporated into judicial evaluation of novel cost allocation issues.

I. DISCOVERY COST ALLOCATION UNDER THE FEDERAL RULES OF CIVIL PROCEDURE: THE CURRENT APPROACH

Most likely no one today would dispute the accepted understanding that, in the federal courts, expenditures made during the discovery process are presumed to lie where they fall. Consistent with the "American Rule"¹⁵ that each party must bear its own attorney's fees and legal costs, the party seeking information from its opponent must bear any costs associated with the drafting—and, where applicable, filing¹⁶—of its discovery requests. The producing party, however, must bear all of the costs associated with the response to the discovery request, including "interpreting the demand, gathering the informa-

¹⁵ The term "American Rule" is used to distinguish the American method of cost allocation from that of countries like England, in which the losing party customarily must reimburse its opponent's attorney's fees and legal costs. John F. Vargo, *The American Rule on Attorney Fee Allocation: The Injured Person's Access to Justice*, 42 AM. U. L. REV. 1567, 1569 (1993). The issue of discovery costs is generally irrelevant in the English system because it does not authorize broad discovery procedures. See generally Geoffrey C. Hazard, Jr., *From Whom No Secrets Are Hid*, 76 TEX. L. REV. 1665, 1677-82 (1998) (comparing American and English discovery procedures).

¹⁶ To be sure, the discovery process is designed to minimize judicial management and intervention. There are, however, some instances in which a party will need to file an official motion with the court, including requests for mental and physical exams, FED. R. CIV. P. 35, or for protective orders to curtail allegedly excessive discovery, *id.* 26(c).

tion, and formulating and delivering a response.”¹⁷ These collective expenses constitute the responding party’s discovery costs.

It is important to emphasize that discovery costs are conceptually, economically, and morally distinct from attorney’s fees and other costs a defendant incurs in association with the litigation process. A party—even a defendant—fully controls the extent of its expenditures on legal fees, and all benefits deriving from those expenditures, legally or strategically, inure to that party. A party thus has full decisionmaking power—within the scope of its resources—as to how to shape its strategic efforts most effectively. By contrast, the extent of a party’s discovery costs are determined not by the litigant himself but by the scope and content of the request filed by his opponent, and none of those expenditures benefits the producing party’s own case. To the contrary, discovery costs benefit the requesting party and actually impose both a financial and a legal detriment on the producing party.¹⁸ This dichotomy between discovery costs and attorney’s fees is reflected by statutes addressing cost allocation, which often allow victorious parties to recover reasonable legal fees from their opponents but do not contemplate reimbursement of discovery costs.¹⁹

At their inception, the rules that governed the discovery process utilized broad, expansive language clearly designed to facilitate the gathering of any information relevant to the litigation. At the time, there existed only limited means by which a party from whom discovery was sought could protect itself. The protective order provision gave the court broad authority to allocate discovery costs between the parties, should it so desire.²⁰ If the producing party believed that the request was too broad or burdensome, or that the information sought was irrelevant, it could file a motion for a protective order in which

¹⁷ Cooper, *supra* note 12, at 466; *see also* Oppenheimer Fund, Inc. v. Sanders, 437 U.S. 340, 358 (1978) (“[T]he presumption is that the responding party must bear the expense of complying with discovery requests . . .”). The sole exception to this general rule involves the deposition of expert witnesses, for which requesting parties are required to bear the costs. *See* FED. R. CIV. P. 26(b)(4)(C).

¹⁸ *See infra* Part II.B.

¹⁹ *See, e.g.*, Clayton Act § 4, 15 U.S.C. § 15 (2006). *See also infra* note 211 for other examples of fee-shifting statutes. Additionally, Federal Rule of Civil Procedure 37(a) allows the court to award “reasonable expenses incurred in making the motion, including attorney’s fees.” FED. R. CIV. P. 37(a)(5)(A).

²⁰ *See* FED. R. CIV. P. 26(c)(2) (“If a motion for a protective order is wholly or partly denied, the court may, on just terms, order that any party or person provide or permit discovery.”). Originally, the provision appeared as FED. R. CIV. P. 30(b) (1938).

the court could deny the request or limit its scope.²¹ One of those conceivable limitations would be the reallocation of costs. Alternatively, the party could respond to the discovery request by filing an objection on the grounds that the request is burdensome²² or irrelevant.²³ This would require the requesting party to seek a motion to compel discovery, thereby enabling the court to rule on the merits of the responding party's objection.²⁴

Starting in the 1980s, a variety of prophylactic devices was introduced into the Rules in an effort to reduce the costs and burdens of discovery.²⁵ Under one of those devices, the court must deny requests for information that is duplicative, that can be obtained in a less burdensome or expensive manner, or that could have been obtained earlier in the action.²⁶ Federal Rule 26(b)(2) also requires the court to limit discovery where it finds that "the burden or expense of the proposed discovery outweighs its likely benefit, considering the needs of the case, the amount in controversy, the parties' resources, the importance of the issues at stake in the action, and the importance of the discovery in resolving the issues."²⁷ The extent to which this device has actually improved the situation, however, is open to serious question. As both judges²⁸ and commentators²⁹ have noted, this propor-

21 This provision now appears in FED. R. CIV. P. 26(c)(1), but originally appeared as FED. R. CIV. P. 30(b) (1938).

22 *See id.* 26(c)(1) (allowing the court to issue an order to protect a person or party from "undue burden or expense").

23 *See id.* 26(b)(1) ("For good cause, the court may order discovery of any matter relevant to the subject matter involved in the action.").

24 *See id.* 37(a), (b).

25 *See supra* note 1 and accompanying text.

26 FED. R. CIV. P. 26(b)(2)(C)(i)-(ii).

27 *Id.* 26(b)(2)(C)(iii).

28 *See* Ronald J. Hedges, *A View from the Bench and the Trenches: A Critical Appraisal of Some Proposed Amendments to the Federal Rules of Civil Procedure*, 227 F.R.D. 123, 127 (2005) ("[T]he proportionality principle of Rule 26(b)(2) . . . is not being utilized by judges . . ."); Shira A. Scheindlin & Jeffrey Rabkin, *Electronic Discovery in Federal Civil Litigation: Is Rule 34 Up to the Task?*, 41 B.C. L. REV. 327, 349 (2000) (describing the proportionality requirement of Rule 26(b)(2) as "seldom-used"). Judge Rosenthal has suggested that the proportionality provisions might be biased against wealthy litigants, noting that

there is also a fear among producing parties that if they have a deep pocket, the issue gets lost before they even open their mouths, because they can afford it. If it is a \$100,000 case and it is going to cost \$200,000 in forensics to restore back-up tapes, that may cut against ordering production, but if you are focusing on the resources of the parties, that may decide the issue.

Panel Discussion, *Managing Electronic Discovery: Views from the Judges*, 76 FORDHAM L. REV. 1, 24 (2007) (quoting Hon. Lee H. Rosenthal, District Judge, United States District Court for the Southern District of Texas).

29 *See, e.g.,* Richard L. Marcus, *Discovery Containment Redux*, 39 B.C. L. REV. 747, 773

tionality requirement has not proven to be an effective limitation on the scope or costs of discovery.

Dissatisfaction with the proportionality provisions, combined with the rapidly increasing use of electronic discovery, led several courts to seek an intermediate solution to alleviate a portion of the financial burden imposed on producing parties. Rather than denying litigants (generally plaintiffs) the opportunity to uncover potentially valuable information, or ordering the opposing party (generally the defendant) to bear the heavy costs associated with the request, several courts shifted all or part of the discovery costs to the plaintiff. The leading case in the area, *Zubulake v. UBS Warburg LLC*,³⁰ contemplated cost shifting where electronic data was “relatively inaccessible,”³¹ on the basis of a seven-factor balancing test that determined how much, if any, of the cost of the request should be borne by the plaintiff.³² Three years after *Zubulake*, the Federal Rules were amended to include “Specific Limitations on Electronically Stored Information.”³³ The amendment orders courts to employ the proportionality analysis of Rule 26(b)(2) to decide whether to limit or shift costs in the context of electronic discovery when certain conditions are found to exist.³⁴ Within the broader landscape of discovery, however, the use of cost shifting remains extremely limited—if not all but nonexistent. To date, as far as can be determined, it has only been em-

(1998) [hereinafter Marcus, *Discovery Containment Redux*] (“[T]here is little evidence that a significant shift has actually occurred as a result of the amendment to the rule”); Richard L. Marcus, *Retooling American Discovery for the Twenty-First Century: Toward a New World Order?*, 7 TUL. J. INT’L & COMP. L. 153, 162–63 (1999) (describing proportionality provisions as “something of a dud”); Scott A. Moss, *Litigation Discovery Cannot Be Optimal but Could Be Better: The Economics of Improving Discovery Timing in a Digital Age*, 58 DUKE L.J. 889, 907 (2009) (“Although denying relevant discovery due to cost may be defensible pragmatically, it is an unsatisfying concession that litigation accuracy inevitably is limited due to the cost of finding and analyzing evidence needed for accurate verdicts or settlements.”).

³⁰ *Zubulake v. UBS Warburg LLC*, 217 F.R.D. 309 (S.D.N.Y. 2003).

³¹ *Id.* at 324.

³² *Id.* at 322–23 (creating a hierarchy of factors in which “[t]he extent to which the request is specifically tailored to discover relevant information” and “the availability of such information from other sources” were the most important, followed by three factors addressing cost issues—“the total cost of production compared to the amount in controversy,” “the total cost of production compared to the resources available to each party,” and “the relative ability of each party to control costs and its incentive to do so”—and concluding with “the importance of the litigation itself” and “the relative benefits of production” to the parties, which rarely come into play). The *Zubulake* balancing test modified an earlier eight-factor test proposed in *Rowe Entertainment, Inc. v. William Morris Agency, Inc.*, 205 F.R.D. 421, 429 (S.D.N.Y. 2002).

³³ FED. R. CIV. P. 26(b)(2)(B).

³⁴ *See id.*

ployed in disputes over electronic discovery,³⁵ and even in this context, courts' use of cost shifting has been restrained at best.³⁶ A number of courts have been reluctant to entertain the notion at all, particularly when the party seeking the cost shifting is a corporate defendant.³⁷ Courts that have been amenable to cost shifting often model their analyses after the multifactor test developed in *Zubulake*, rather than the proportionality approach of Rule 26(b)(2),³⁸ and even then generally shift only a portion of the discovery costs to the re-

³⁵ As one of us has previously shown, electronic discovery requests are, in certain instances, more expansive in nature and considerably more costly to fulfill than are more traditional requests, such as production of physical documents. See Redish, *supra* note 7, at 588–92 (dividing and explaining the differences between electronic and regular discovery in terms of three categories: volume, retrieval, and translation).

³⁶ Corinne L. Giacobbe, Note, *Allocating Discovery Costs in the Computer Age: Deciding Who Should Bear the Costs of Discovery of Electronically Stored Data*, 57 WASH. & LEE L. REV. 257, 293 (2000) (noting that the majority of cost-shifting requests are denied). Put another way, even if a producing party can show that a disproportionately broad request for electronic discovery would impose an undue financial burden, it would not be able to trigger the *Zubulake* balancing test unless the information were stored in a “relatively inaccessible” format, such as backup tapes. *Zubulake*, 217 F.R.D. at 324. One commentator argues that this restriction leads to a net reduction in the chances that costs will be shifted. See Jessica Lynn Repa, Comment, *Adjudicating Beyond the Scope of Ordinary Business: Why the Inaccessibility Test in Zubulake Unduly Stifles Cost-Shifting During Electronic Discovery*, 54 AM. U. L. REV. 257, 285 (2004).

³⁷ See, e.g., *United States v. Davey*, 543 F.2d 996, 1001 (2d Cir. 1976) (denying cost shifting because “[f]or this billion dollar company, . . . the cost of duplicating[,] . . . estimated at approximately \$1305, would be minimal, representing but a small outlay in comparison with the other amounts which the taxpayer expends annually in cooperating with the IRS”).

³⁸ See, e.g., *Quinby v. WestLB AG*, 245 F.R.D. 94, 106–10 (S.D.N.Y. 2006) (analyzing a dispute over the production of emails using the seven-factor *Zubulake* test). Some courts have also used the eight-factor balancing test established by *Rowe Entertainment*, which differed from *Zubulake* in that it did not establish a hierarchy among the factors. See *Rowe Entm't, Inc. v. William Morris Agency, Inc.*, 205 F.R.D. 421, 429 (S.D.N.Y. 2002). Judge Francis, who decided *Rowe Entertainment*, resisted using the hierarchy approach because he feared that “the factor at the top of the hierarchy will almost always wash out the other factors.” Panel Discussion, *supra* note 28, at 24 (quoting Hon. James C. Francis IV, District Judge, United States District Court for the Southern District of New York). For cases using the *Rowe Entertainment* analysis, see, for example, *Murphy Oil U.S.A., Inc. v. Fluor Daniel, Inc.*, 52 Fed. R. Serv. 3d (West) 168, 173 (E.D. La. 2002) (noting that *Rowe Entertainment* “provides sound guidance for resolution of these issues where the retrieval, production and review of e-mail from backup tapes is at issue”), and *In re Bristol-Meyers Squibb Sec. Litig.*, 205 F.R.D. 437, 443 (D.N.J. 2002) (following the *Rowe Entertainment* cost-shifting approach where the requesting party bears the costs of discovery). Courts' preference for *Zubulake* over Rule 26(b)(2) is perhaps not surprising, considering the ineffectiveness of the Rule in limiting nonelectronic discovery. See Moss, *supra* note 29, at 905 (stating that “proportionality is widely criticized as having been ineffective at convincing judges to rein in discovery excess”); see also Lee H. Rosenthal, *A Few Thoughts on Electronic Discovery After December 1, 2006*, 116 YALE L.J. POCKET PART 167 (2006), <http://thepocketpart.org/images/pdfs/82.pdf> (noting how difficult it is for judges to use the proportionality requirement to draw distinctions that provide predictable guidelines for limiting discovery).

requesting party.³⁹ Because of the complexity and unwieldiness of the balancing tests used to determine cost shifting, as well as the long-standing presumption against requiring a party to bear the costs of its own discovery requests,⁴⁰ cost shifting has proven more popular among scholars⁴¹ than among federal district court judges, though even in the scholarly context its success has at best been limited. In the discussion that follows, we show that the governing presumption against cost shifting—the likely reason that judges either categorically reject cost shifting or feel compelled to engage in elaborate balancing before showing a willingness to overcome this inertia and order the requesting party to bear some or all of the costs generated by its demands—contradicts foundational moral, economic, democratic, and (in certain instances) constitutional principles.

³⁹ See, e.g., *Zubulake v. UBS Warburg LLC*, 216 F.R.D. 280, 289 (S.D.N.Y. 2003) (ordering *Zubulake* to bear twenty-five percent of the costs associated with her request); *Williams v. E.I. du Pont de Nemours & Co.*, 119 F.R.D. 648, 651 (W.D. Ky. 1987) (ordering *du Pont* to pay to the plaintiff a “fair portion of the fees and expenses incurred in the past by the [plaintiff] for the work of the [plaintiff’s] expert in encoding the requested data” (internal quotation marks omitted)).

⁴⁰ *Zubulake* urges courts to resolve close cases in favor of the governing presumption against cost shifting. *Zubulake*, 217 F.R.D. at 320.

⁴¹ For scholarly arguments in favor of partial or total cost-shifting, see Cooter & Rubinfeld, *supra* note 7, at 455–56 (proposing a two-part cost-shifting rule that would “require the defendant to bear the costs of reasonable compliance up to a level deemed appropriate for this class of cases, beyond which the reasonable costs of complying with further discovery requests would shift to the plaintiff”), Hedges, *supra* note 28, at 129 (“Perhaps we should consider a new paradigm for costs. If discovery is sought that is relevant to a claim or defense, the producing party might bear the costs. If the requesting party can show good cause for ‘expanded’ discovery under Rule 26(b)(1), it might bear the costs. This would at least impose a measure of predictability and enable parties to consider the costs of e-discovery before requests are made and responded to.”), Redish, *supra* note 7, at 608–15 (proposing a conditional cost-shifting model for electronic discovery), and Panel Discussion, *supra* note 28, at 25 (quoting Hon. James C. Francis IV, District Judge, United States District Court for the Southern District of New York) (“With cost shifting, we can make decisions that are somewhat more nuanced in the face of this uncertainty and say, ‘Well, I think it may be forty percent likely that it is useful. In light of that, I am going to shift a portion of the costs.’ That, I think, is a more satisfying result than having to make a black-or-white, yes-or-no determination.”). For arguments to the contrary, see Cooper, *supra* note 12, at 469–70 (responding to Cooter and Rubinfeld’s proposal), and Moss, *supra* note 29, at 943 (“[E]ven with more cost shifting, courts still would face information-intensive decisions about which discovery is (1) sufficiently important that the requesting party should get it without paying production costs; (2) important enough that the requesting party could get it by paying production costs; or (3) sufficiently lacking in value that the requesting party cannot get it even if willing to pay for it. Such decisions remain intractable in many cases because (as is this Article’s primary diagnosis about the problem of costly discovery) courts often lack sufficient information about case value and evidentiary value to undertake accurate cost-benefit analyses on discovery disputes.”).

II. ALLOCATING DISCOVERY COSTS: A RETURN TO FIRST PRINCIPLES

A. *Discovery Costs as Quantum Meruit*

In order to explain our reconsideration of the first principles of discovery cost allocation, it is necessary at the outset to set aside more than seventy years of procedural history and its well-established presumptions about cost allocation, and view the behavior of the requesting and producing parties through the lens of foundational principles of moral and economic justice. Throughout the evolution of modern contract doctrine, courts have recognized that it is just to require a party to pay for services that it requested, even if the request itself is legally unenforceable purely as a matter of contract law.⁴² The theory of quantum meruit provided the conceptual foundation for the reimbursement of expenses incurred by one individual in providing a benefit to another person who had requested or was otherwise expecting that benefit.⁴³ After briefly describing the doctrine of quantum meruit, we liken the current allocation of discovery costs to quasi contracts, which allow a party that has conferred a benefit at its own expense to recover the fair value of its services under the theory of quantum meruit.

1. *Theory of Quantum Meruit*

Quantum meruit, which literally means “as much as he has deserved,”⁴⁴ is a legal action brought to recover compensation for work performed without existence of a prior formal, legally enforceable agreement between the parties.⁴⁵ The action originated in the seventeenth-century English common law courts.⁴⁶ Although quantum meruit actions were brought under the writ of assumpsit, which was generally used to enforce a contractually incurred debt,⁴⁷ the common law judges were often willing to infer the existence of a promise to pay

⁴² See Candace S. Kovacic, *A Proposal to Simplify Quantum Meruit Litigation*, 35 AM. U. L. REV. 547, 551 (1986).

⁴³ See Judy Beckner Sloan, *Quantum Meruit: Residual Equity in Law*, 42 DEPAUL L. REV. 399, 402 (1992). Actually, the theory of quantum meruit extends well beyond mere reimbursement to include compensation. However, for present purposes, we need not focus on the compensation issue.

⁴⁴ BLACK'S LAW DICTIONARY 1361 (9th ed. 2009).

⁴⁵ Sloan, *supra* note 43, at 402.

⁴⁶ *Id.* at 423.

⁴⁷ J. H. BAKER, AN INTRODUCTION TO ENGLISH LEGAL HISTORY 282–83 (2d ed. 1979).

from the circumstances of the case.⁴⁸ Quantum meruit originally enabled recovery for contracts implied in fact, in which both parties agreed or expected “that the performing party would be paid the reasonable value, as opposed to a specified rate, of the services rendered.”⁴⁹ In other words, courts enforced contracts implied in fact because the conduct of the parties clearly indicated that the performing party acted pursuant to an implicit consensual agreement that bore some, though not all, of the elements of a legally cognizable contract.

Over time, courts also allowed recovery in quantum meruit for contracts implied in law, or “quasi contracts.”⁵⁰ Unlike a contract implied in fact, a quasi contract bore virtually no resemblance to a legally enforceable contract.⁵¹ There was no “meeting of the minds” between the parties prior to the conferring of the benefit, nor any indication of an express bargain.⁵² The defendant’s formal acceptance of the benefit was not required and, indeed, was deemed irrelevant in cases in which the defendant requested the performed services.⁵³ Only three essential elements needed to be present to support recovery in quantum meruit: (1) the defendant must have received a benefit,⁵⁴ (2) at the plaintiff’s expense, (3) which would be unjust for the defendant to retain without compensation and which was not gratuitously rendered.⁵⁵ Based on these three elements, the court would infer a completely fictitious promise on the basis of which it granted recovery for the injured party.⁵⁶ Although quasi contracts have technically been subsumed within the theories of restitution and unjust

⁴⁸ *Id.* at 305. For a detailed history of the evolution of quantum meruit, see generally Sloan, *supra* note 43, at 414–25.

⁴⁹ Sloan, *supra* note 43, at 426.

⁵⁰ *See id.* at 427.

⁵¹ *Id.* at 407.

⁵² *Id.*; see also 1 GEORGE E. PALMER, *THE LAW OF RESTITUTION* § 1.1, at 3 (1978).

⁵³ Kovacic, *supra* note 42, at 563.

⁵⁴ “Legal benefit” is a term of art referring to “something that could serve or benefit in some way the [receiving party’s] interests.” Peter Benson, *The Unity of Contract Law*, in *THE THEORY OF CONTRACT LAW: NEW ESSAYS* 118, 158 (Peter Benson ed., 2001). It includes, but is not limited to, economic gain or loss. Legal benefits can have “any content whatsoever so long as they reasonably appear to be the interests of a given party, taken as an individual distinct from and independent of the other party.” *Id.*

⁵⁵ Kovacic, *supra* note 42, at 563; see also FREDERIC CAMPBELL WOODWARD, *THE LAW OF QUASI CONTRACTS* §§ 7, 9 (1913).

⁵⁶ See 1 PALMER, *supra* note 52, § 1.1, at 3; see also Emily Sherwin, *Restitution and Equity: An Analysis of the Principle of Unjust Enrichment*, 79 *TEX. L. REV.* 2083, 2094 (2001) (“Thus, not only were the formal legal requirements for these ‘quasi-contract’ claims much simpler than the requirements for other actions, but such requirements as there were, were acknowledged to be fictitious.”).

enrichment,⁵⁷ modern courts still rely on the formal legal fiction of an implied promise to evaluate actions brought for recovery in quantum meruit.⁵⁸

Although it superficially conforms to the rigid constructs of contract law, quantum meruit—particularly as it was used to enforce quasi contracts—relies on a fundamentally different rationale in order to compensate an injured party.⁵⁹ Rather than demanding that a checklist of factors (offer, acceptance, and consideration) be met in order to provide a remedy to an injured party, quantum meruit appeals to foundational principles of economic justice—specifically, that he who confers a benefit on another should be compensated for it regardless whether or not he formally bargained to be paid.⁶⁰ As a legal doctrine, quantum meruit allows the court to enforce “the moral tenet that one person should not obtain unfair advantage at another’s expense.”⁶¹

The principle that actions in quantum meruit turn on concerns of fundamental fairness dates back to the Roman jurist Pomponius, who wrote that “this by nature is equitable, that no one be made richer by another’s loss.”⁶² In eighteenth-century England, Lord Mansfield famously explained quasi contracts in terms of principles of fundamental morality: “If the defendant be under an obligation, from the ties of

⁵⁷ See RESTATEMENT OF RESTITUTION pt. I, intro. note (1937) (explaining that the rules of restitution articulated in the *Restatement* are applicable to “situations in which, before the development of modern procedure, an action of general assumpsit could have been brought to secure restitution,” including “the subject of quasi contracts”).

⁵⁸ For recent examples of courts discussing quasi contracts under the theory of unjust enrichment, see, for example, *Sullivan v. Leor Energy, LLC*, 600 F.3d 542, 550 (5th Cir. 2010), *Boring v. Google Inc.*, 362 F. App’x 273, 281–82 (3d Cir. 2010), and *Boccardi Capital Sys., Inc. v. D.E. Shaw Laminar Portfolios, L.L.C.*, 355 F. App’x 516, 519–20 (2d Cir. 2009). See also Sherwin, *supra* note 56, at 2086 (noting that much of restitution law developed from quasi-contractual remedies).

⁵⁹ See CHARLES FRIED, *CONTRACT AS PROMISE* 26 (1981) (noting that promise (which forms the basis of reliance and expectation interests in contract law) and restitution (which forms the basis of actions in quantum meruit) are “distinct principles”); see also *id.* at 125 (explaining how a breach of a contract releases the nonbreaching party from its contractual obligations but does not release him from the “quite different obligation to pay for tangible benefits that he chooses to retain”).

⁶⁰ George K. Gardner, *An Inquiry into the Principles of the Law of Contracts*, 46 HARV. L. REV. 1, 20–21 (1932); see also Kovacic, *supra* note 42, at 551.

⁶¹ Sherwin, *supra* note 56, at 2104; see also Kovacic, *supra* note 42, at 550–51 (noting that an action in quantum meruit is premised on the basic moral principle that it is just to force a defendant to pay for requested services even if the agreement containing the request is legally unenforceable).

⁶² JOHN P. DAWSON, *UNJUST ENRICHMENT: A COMPARATIVE ANALYSIS* 3 (1951) (internal quotation marks omitted).

natural justice, to refund; the law implies a debt, and gives this action, founded in the equity of the plaintiff's case, as it were upon a contract ('quasi ex contractu,' as the Roman law expresses it)."⁶³ Although most courts have not been quite so explicit about the moral underpinnings of quantum meruit, they have willingly used quasi contract,⁶⁴ as well as the modern theory of unjust enrichment, as an ex post rationalization for the legal enforcement of this broader moral principle of economic justice.⁶⁵ Scholars continue to conceptualize quantum meruit, as well as the more general theory of unjust enrichment, as a "morally inspired" principle of law⁶⁶ that appeals to a "primitive intuition of fairness" operating independently of the parties involved in the transaction.⁶⁷

The measure of recovery in a quantum meruit action accords with its core purpose—namely, the restoration of both parties to the positions they occupied before the plaintiff conferred a benefit on the defendant.⁶⁸ As such, recovery in quantum meruit is to be measured, in theory, not by the loss to the injured party but rather by the amount that the party receiving the benefit actually gained from the services.⁶⁹ In many contexts, however, that benefit is either not readily quantifiable or exceeds the loss incurred by the performing party. In those cases, most courts award the reasonable market value of the plaintiff's services, which can be viewed as the defendant's gain in requesting those services.⁷⁰ If the actual costs incurred by the plaintiff coincide with the reasonable market value, some courts simply award the plaintiff's costs.⁷¹ In essence, in actions for recovery in quantum meruit, courts have the discretion to shape the recovery in the way that most fully compensates the performing party for the service that he pro-

⁶³ *Moses v. Macferlan*, (1760) 97 Eng. Rep. 676 (K.B.) 678.

⁶⁴ One scholar aptly noted that the "rigidity of Anglo-American legal thought requires some tangible basis for enforcement of an implied-in-law contract." Sloan, *supra* note 43, at 408.

⁶⁵ *See id.*

⁶⁶ Sherwin, *supra* note 56, at 2106.

⁶⁷ FRIED, *supra* note 59, at 25.

⁶⁸ RESTATEMENT OF RESTITUTION § 1 cmt. a (1937).

⁶⁹ *Id.* § 1 cmt. e; Sloan, *supra* note 43, at 427–28.

⁷⁰ *See* RESTATEMENT (SECOND) OF CONTRACTS § 371(a) (1981) (describing the amount recoverable in an action in restitution (or unjust enrichment) as "the reasonable value to the other party of what he received in terms of what it would have cost him to obtain it from a person in the claimant's position"); Sloan, *supra* note 43, at 428 ("If the receiving party actually requested that the services be performed, the benefit he has gained can be appropriately measured by the fair market value of the services."); *see also* WOODWARD, *supra* note 55, § 8; Kovacic, *supra* note 42, at 556–57.

⁷¹ RESTATEMENT OF RESTITUTION § 1 cmt. e (1937); Kovacic, *supra* note 42, at 556–57.

vided, while simultaneously limiting recovery so that it does not exceed the performing party's out-of-pocket costs.⁷²

2. *Reconceptualizing Discovery as Quantum Meruit*

During the discovery process, the parties to the litigation function in a manner strikingly similar to parties engaged in a quasi-contractual relationship. Initially, the requesting party formally solicits its opponent's labor by invoking one of the discovery devices in order to obtain a particular piece of information. Absent a protective order from the court,⁷³ the producing party is bound to comply with the request, and is thus legally required to incur the costs associated with it.⁷⁴ In fact, the producing party incurs two distinct types of discovery costs. First, and most obviously, the producing party must absorb all the direct costs of locating and organizing the relevant information. In many instances, these costs will no doubt be substantial, particularly when the requesting party seeks production of electronically stored information that must first be restored or reformatted by the producing party.⁷⁵ The producing party also incurs opportunity costs when its employees are forced to participate in the production of the requested information, whether by locating particular documents, participating in a deposition, or answering written questions.⁷⁶ Second, and perhaps even more importantly, the fruits of the producing party's labor are then handed over to its opponent for use *against* the producing party during the lawsuit. Not only does discovery impose immediate financial costs on the producing party, but it also increases the likelihood that the producing party's opponent (whether plaintiff or defendant) will ultimately win the lawsuit.⁷⁷

⁷² See RESTATEMENT OF RESTITUTION ch. 8, topic 2, intro. note (1937) (explaining that if the defendant is not at fault, "he is not required to pay for losses in excess of benefit received by him").

⁷³ See FED. R. CIV. P. 26(c).

⁷⁴ See *supra* note 14 and accompanying text.

⁷⁵ See, e.g., *Semsroth v. City of Wichita*, 239 F.R.D. 630, 632–33 (D. Kan. 2006) (requesting restoration of e-mails from the municipality's backup servers, which would require the purchase of expensive, but otherwise superfluous, software and would require the city's IT personnel to devote substantial time to a manual search of the files); *Zubulake v. UBS Warburg LLC*, 217 F.R.D. 309, 311–12 (S.D.N.Y. 2003) (seeking restoration of e-mails from the defendant's electronic backup tapes, which would cost the defendant an estimated \$175,000).

⁷⁶ See Judith A. McKenna & Elizabeth C. Wiggins, *Empirical Research on Civil Discovery*, 39 B.C. L. REV. 785, 799 (1998).

⁷⁷ See Cooter & Rubinfeld, *supra* note 7, at 450 (noting that additional discovery benefits the plaintiff by adding value to his claim).

The party requesting the information correspondingly receives two levels of benefit from the producing party's efforts. Initially, the requesting party saves itself the financial expense of having to uncover the same information in a less efficient (and likely more costly) manner. Of course, it is also conceivable that the requesting party would not have been able to obtain the equivalent information at all absent the court-enforced discovery devices.⁷⁸ In that case, the benefit to the requesting party, though impossible to quantify in terms of costs avoided, is arguably even more substantial, because without the producing party's labor, the requesting party would not have access to the information at all. Second, just as the producing party suffers an intangible detriment when it is forced to turn over information that potentially benefits its opponent's case, the requesting party correspondingly benefits from the receipt of relevant information.

The third and final element of a claim recoverable in quantum meruit is unjust enrichment of the party receiving the benefit, which necessarily requires that the services not be performed gratuitously.⁷⁹ To be clear, the requirement of unjust enrichment does not require any wrongdoing on the part of the party receiving the services.⁸⁰ Rather, it refers to the retention of a nongratuitous benefit that was conferred at the expense of another party.⁸¹ The enrichment is unjust only to the extent that it has not been paid for by the party benefitted. In the context of discovery, the producing party's efforts can hardly be characterized as a voluntary effort to perform a favor for or provide a gift to the requesting party, because (1) the responding party, for the most part, acts only in response to a formal request by the requesting party,⁸² and (2) the responding party is subject to Rule 37 sanctions if it fails to respond to a discovery request.⁸³ Indeed, some of these potential sanctions could have such a detrimental effect on the produc-

⁷⁸ See Marcus, *Discovery Containment Redux*, *supra* note 29, at 751 (noting that employment discrimination law is premised on the availability of discovery to order production of otherwise confidential information).

⁷⁹ WOODWARD, *supra* note 55, §§ 7, 9.

⁸⁰ See Lionel Smith, *Restitution: The Heart of Corrective Justice*, 79 TEX. L. REV. 2115, 2115 (2001) (noting that one of the central features of the law of restitution for unjust enrichment is liability without fault).

⁸¹ *Id.* at 2118 (arguing that the theory of corrective justice—one of the underpinnings of restitution—"requires that pretransaction holdings be respected").

⁸² This is not the case when the discovery is produced as part of automatic disclosure. See FED. R. CIV. P. 26(a)(1).

⁸³ *Id.* 37.

ing party's case that they leave a party with little choice but to comply with its opponent's demands.⁸⁴

Absent compensation to the party that produced the information, the enrichment of the requesting party is indeed unjust. As Lon Fuller and William Perdue aptly explained, if the Aristotelian tradition of regarding the purpose of justice as the "maintenance of an equilibrium of goods among members of society" is followed, the restitution interest presents a strong claim for judicial intervention—indeed, twice as strong as does the reliance interest.⁸⁵ The defendant not only causes a loss to the plaintiff but appropriates a gain to himself, making the resulting discrepancy between the parties "not one unit but two."⁸⁶ In the case of discovery costs, the resulting discrepancy between the parties is not *two* units, but *four*. As noted above, the producing party not only bears the financial costs of complying with its opponent's request, but it also suffers the additional detriment of having the fruits of its labor used against it in the ongoing litigation. Essentially, the producing party suffers on two distinct levels as a result of its efforts.⁸⁷ The requesting party, in turn, receives two distinct benefits as a result of the producing party's work: it obtains the immediate benefit of receiving the specific information it requested, as well as a simultaneous detriment to its opponent. If, as Fuller and Perdue argued, restitution in general presents the strongest moral case for judicial intervention,⁸⁸ discovery costs present an even stronger case for

⁸⁴ Rule 37 authorizes dismissal of a plaintiff's action, default judgment against the disobedient party, and establishment or preclusion of the matter at issue. *See id.* 37(b)(2)(A). If a party disobeys a court order compelling discovery, it can also be held in contempt of court. *Id.* 37(b)(2)(A)(vii).

⁸⁵ L. L. Fuller & William R. Perdue, Jr., *The Reliance Interest in Contract Damages*, 46 *YALE L.J.* 52, 56 (1936). In a claim for recovery of a reliance interest, the plaintiff has relied on an agreement with the defendant to his detriment but has not conferred any benefit upon the defendant. *See id.*

⁸⁶ *Id.*

⁸⁷ While compiling the information requested by its opponent, the producing party might take some actions that have an incidental benefit to its own case, such as the redaction of privileged information. Even in a case where discovery costs are driven largely by the costs of redaction, however, our central argument still holds: The producing party is only undertaking this labor because it has been requested to do so by its opponent. The requested information, regardless of the interstitial processes needed to prepare it, will ultimately create a benefit to the requesting party. As such, the requesting party should still bear the costs of redaction, even though the producing party might reap an incidental benefit from it. Moreover, it might be a misnomer to describe redaction as an "incidental benefit" to the producing party. More accurately, redaction ensures that the producing party is harmed less by the production of information than it otherwise might have been. Redaction is not so much a benefit that the producing party retains for itself so much as it is a protective mechanism.

⁸⁸ Fuller & Perdue, *supra* note 85, at 56.

recovery in quantum meruit, as the multiple levels of benefit and detriment move the parties even farther away from the Aristotelian equilibrium.

Under the theory of quantum meruit, the producing party would be entitled—at the very least—to reimbursement of the costs it incurred in fulfilling its opponent's discovery request.⁸⁹ Although that recovery would not fully capture the benefit received by the requesting party, it would at least restore the producing party to the same financial position that it occupied before it undertook its labor on behalf of its opponent.⁹⁰ To reduce the incentive for the producing party to inflate its expenses, however, the court could choose to award the producing party the reasonable market value of its services if the requesting party could demonstrate that the actual costs allegedly incurred by the producing party were unreasonable.⁹¹ Reimbursement would remedy the unjust enrichment of the requesting party because the requesting party would have paid for the benefit it received as a result of the performance of services it acquired through its use of the court-enforced discovery devices. To be sure, one could make an argument for going even further and forcing the requesting party to compensate the producing party for the additional benefit that the requested information will have to the requesting party's case.⁹² But this Article need not delve into that issue, as it does not affect its main point—that the moral underpinnings of quantum meruit dictate that each party should at the very least be required to reimburse the discovery costs of its opponent.

*B. The Implications of the Quantum Meruit Perspective:
Discovery Costs as a Litigation Subsidy*

As shown in the previous Section, our current procedural system allocates discovery costs without regard to whether the producing party suffers an unrecoverable financial detriment as a result of its

⁸⁹ In cases in which the benefits to the defendant and the detriments to the plaintiff do not directly coincide, courts often limit quantum meruit recovery to the costs incurred by the plaintiff. RESTATEMENT OF RESTITUTION § 1 cmt. e (1937).

⁹⁰ See RESTATEMENT (SECOND) OF CONTRACTS § 344(c) (1981) (noting that the purpose of the restitution remedy is the restoration to the plaintiff of any benefit that it has conferred upon the defendant).

⁹¹ Indeed, allowing the requesting party to challenge the reasonableness of the producing party's costs would reduce the producing party's incentive to inflate its costs. See *infra* Part II.B.2 for a discussion of this and other economic implications of cost allocation.

⁹² One could admittedly argue, however, that although the fruits of each discovery request would benefit the requesting party by advancing its case, such a benefit would be highly speculative and difficult to quantify.

efforts and, more importantly, without regard to whether the requesting party unjustly enriches itself when it takes and uses the information provided by its opponent. Were the well-established premises of unjust enrichment to be applied to the discovery context, then (as has been argued throughout) those costs would properly be deemed those of the requesting party, not those of the responding party. The systemic choice to allow discovery costs to lie where they fall, without regard to the intrinsic injustice of doing so, transforms these costs into a *de facto* litigation subsidy. Because each party bears the costs of producing the information that will be used against it by its opponent, each party effectively subsidizes that portion of its opponent's case.

Viewing discovery costs as a litigation subsidy reveals fundamental problems with our system's current method of cost allocation from the perspectives of democratic, economic, and constitutional theory. First, creation of a hidden litigation subsidy undermines the fundamental political goals of legislative transparency and democratic accountability that are central to any viable theory of representative democracy. Second, the subsidization of discovery costs by the requesting party's opponent undermines the efficiency of our procedural system by giving rise to a significant economic externality in the conduct of the discovery process. Finally, forced provision of a litigation subsidy by a defendant prior to any judicial evaluation of the underlying merits of the claim brought against it violates that party's constitutional right to procedural due process. Although the constitutional implications for plaintiffs are not nearly as severe, here, too, constitutional difficulties arise.

1. Implications for Democratic Theory

Our system's failure to openly conceptualize discovery costs as a litigation subsidy, both at the time of the adoption of the original Federal Rules and throughout their subsequent revision and modification, has caused substantial misperception in the minds of the populace by creating a hidden subsidy. In effect, a law has been adopted in a manner that effectively prevents free and open debate in a democratic society on an issue of arguably great public importance.

At first blush, one might reasonably question this reasoning on several grounds. Initially, it might be argued that the Federal Rules are promulgated through a fundamentally undemocratic procedure. In the Rules Enabling Act,⁹³ Congress delegated "the power to pre-

⁹³ Rules Enabling Act, 28 U.S.C. §§ 2071–2077 (2006).

scribe general rules of practice and procedure” to the Supreme Court—the one branch of the federal government unaccountable to the electorate, even indirectly.⁹⁴ Although Congress retained the power to revise or reject any proposed rules through legislative action,⁹⁵ the Act quite clearly set legislative inertia in favor of the Court’s rules. Unless proposed rules are rejected, modified, or deferred, they automatically become law, provided that Congress has had at least seven months to consider them.⁹⁶

Congress has had a mixed history of involvement (and, indeed, interest) in the rulemaking process. The original Federal Rules of Civil Procedure automatically passed into law without a congressional vote.⁹⁷ In fact, for the thirty-five years after Congress passed the Rules Enabling Act, it acquiesced in all of the proposed Rules promulgated by the Supreme Court.⁹⁸ In later years, however, Congress took a more active role in the drafting of the Federal Rules, even going so far as to rewrite proposed rules of evidence defining privileges.⁹⁹ In 1993, Congress came close to enacting a bill to eliminate proposed mandatory disclosure provisions in the Federal Rules of Civil Procedure.¹⁰⁰ Additionally, Congress amended the Rules Enabling Act in 1988 in an attempt to bring about more transparency and participation during the initial stages of the rulemaking process.¹⁰¹ The fact remains, however, that given the way the rulemaking process is structured, the inertia is clearly in favor of the rule choices made by the unaccountable Supreme Court. In light of this fact, it could be argued that democratic theory is largely irrelevant to rule choices. If that is true, one could perhaps conclude that an argument grounded in electorate confusion is incoherent.

The largely undemocratic nature of the rulemaking process should be fully acknowledged. But that does not automatically mean that the potential congressional role in the process should be completely ignored. When the rules make clear, on their face, how they

⁹⁴ *Id.* § 2072(a).

⁹⁵ Paul D. Carrington, *Learning from the Rule 26 Brouhaha: Our Courts Need Real Friends*, 156 F.R.D. 295, 297 (1994).

⁹⁶ 28 U.S.C. §§ 2074, 2075. For a detailed description of the rulemaking process, see *The Rulemaking Process*, USCOURTS.GOV, <http://www.uscourts.gov/RulesandPolicies/FederalRulemaking/RulemakingProcess.aspx> (last visited Feb. 20, 2011).

⁹⁷ Carrington, *supra* note 95, at 297; Subrin, *supra* note 8, at 729.

⁹⁸ Leslie M. Kelleher, *Taking “Substantive Rights” (in the Rules Enabling Act) More Seriously*, 74 NOTRE DAME L. REV. 47, 55 (1998).

⁹⁹ *Id.*

¹⁰⁰ *Id.* at 57.

¹⁰¹ Carrington, *supra* note 95, at 300.

shape the legal topography, both the electorate and members of Congress may make their own decisions as to whether alteration or rejection is called for; the democratic safety valve is always available. When, however, the rules fail to make clear how they are, in reality, shaping the procedural landscape, the effectiveness of this democratic safety valve is undermined, if not destroyed.

It might also be argued that even if the democratic concern were assumed to be accurate in theory, it would be unrealistic to believe that the concern is relevant to the discovery cost allocation question. That question, the argument would proceed, involves a highly technical issue in civil procedure—one that is of questionable interest to law students, much less to the public at large. But it would be wrong to dismiss the political salience of issues of civil procedure in the current political climate. Indeed, issues of so-called tort reform—many of which directly affect or involve procedure—are of enormous importance to significant portions of the electorate.¹⁰² This is demonstrated by the fact that a variety of bills has been introduced in Congress to alter current procedural requirements and, as previously noted, on occasion have actually come close to passage.¹⁰³

Discovery cost allocation is not explicitly addressed by any Federal Rule of Civil Procedure. Rather, its treatment is embedded in the interstices of Federal Rules 26(c) and 26(b)(2), which provide the court with “broad authority to control the cost of discovery.”¹⁰⁴ Rule 26(c)(2), for example, allows courts to wholly or partially deny a protective order, and to compel discovery “on just terms.”¹⁰⁵ Rule 26(b)(2)(C) gives courts the greater power to limit discovery, which can plausibly be interpreted to include the lesser power of allowing discovery while shifting all or a portion of the costs to the requesting

¹⁰² See STANFORD UNIV. & THE ROBERT WOOD JOHNSON FOUND., THE ASSOCIATED PRESS POLL 23 (2009), available at http://surveys.ap.org/data/gfk/ap-stanford-rwj%20healthcare%20topline%20final_nov18%20edits.pdf (finding that fifty-four percent of Americans favor “[m]aking it harder to sue doctors and hospitals for making mistakes taking care of patients”).

¹⁰³ Kelleher, *supra* note 98, at 57–58.

¹⁰⁴ MANUAL FOR COMPLEX LITIGATION (THIRD) § 21.433 (1995); see also FED. R. CIV. P. 26 advisory committee’s note to 1970 amendment; *Seema Gems, Inc. v. Shelgem, Ltd.*, No. 93 Civ. 4473 (KMW), 1994 WL 86381, at *2 (S.D.N.Y. Mar. 16, 1994) (“[T]he court has considerable discretion to allocate the costs of discovery, pursuant to Rule 26(c)(2) of the Federal Rules of Civil Procedure.”); *Bills v. Kennecott Corp.*, 108 F.R.D. 459, 462 (D. Utah 1985) (“Rule 26(c) commonly has been interpreted to grant courts the power to shift the financial burden of discovery where, in the court’s discretion, such a shifting is warranted.”); *Fine v. Grossman*, No. 80 Civ. 4771-CSH, 1982 U.S. Dist. LEXIS 11838, at *6 (S.D.N.Y. Mar. 17, 1982) (“[W]here the burden is heavy, where a segregation and analysis of a great mass of material is necessary[,] . . . some, and perhaps the greatest share[,] of that burden and effort should fall on [the requesting party].”).

¹⁰⁵ FED. R. CIV. P. 26(c)(2).

party.¹⁰⁶ Rule 26(b)(2)(B), enacted in 2006, gives the court the power to “specify conditions” for electronic discovery.¹⁰⁷ To be sure, it requires some interpretive legwork to infer anything about cost allocation from the text of these provisions, since none of them comes close to stating explicitly the general principle that, absent judicial intervention, discovery costs lie where they fall.¹⁰⁸ Indeed, a legislator unfamiliar with the litigation process might not be able to infer any standard for cost allocation from the text of the Rules. As a result, it is not surprising that Congress has never debated the merits of various methods of cost allocation.¹⁰⁹

The general lack of legislative transparency surrounding cost allocation issues is certainly problematic in and of itself. When one recognizes that the established version of discovery cost allocation effectively operates as a litigation subsidy, however, the magnitude of the problem increases exponentially. By definition, a subsidy employs the legal power of the government to take resources from one entity and redistribute them to another in a manner that conflicts with the ordinary processes of the market economy and precepts of private property.¹¹⁰ Because the Supreme Court never explicitly conceptualized discovery costs as a subsidy, however, neither the Court nor Congress ever so much as debated, let alone sanctioned, this use of governmental power to force one litigant to use its own resources to fund a potentially important part of its opponent’s case. Even if one were to concede, for purposes of argument, that the politically unaccountable Supreme Court can properly promulgate the Federal Rules of Civil Procedure,¹¹¹ the rulemaking process does provide Congress the authority to overrule any Federal Rule—an authority that it has

¹⁰⁶ *Id.* 26(b)(2)(C).

¹⁰⁷ *Id.* 26(b)(2)(B).

¹⁰⁸ *See supra* note 3 and accompanying text.

¹⁰⁹ *See Subrin, supra* note 8, at 692 (“What is perhaps most surprising in the [twenty-year] public debate among those who most vigorously fought for and against the Enabling Act is the insignificance of discovery issues.”).

¹¹⁰ The Department of Commerce defines a subsidy as follows:

In a market economy, scarce resources are channeled to their most profitable and efficient uses by the market forces of supply and demand. We believe a subsidy (or bounty or grant) is definitionally any action that distorts or subverts the market process and results in a misallocation of resources, encouraging inefficient production and lessening world wealth.

Carbon Steel Wire Rod from Poland, 49 Fed. Reg. 19,374, 19,375 (Dep’t of Commerce May 7, 1984) (final determination).

¹¹¹ Although this discussion falls outside the scope of this Article and there exists no directly supporting precedent for the position, one of us has previously questioned the constitutionality of the Rules Enabling Act. *See* Martin H. Redish & Uma M. Amuluru, *The Supreme*

chosen to exercise in the past.¹¹² The failure to recognize the embedded subsidy prevented members of Congress from considering the ramifications for their constituents, as either potential litigants or consumers, stemming from the chosen method of cost allocation. As a result, both voters and their representatives have effectively been denied the opportunity to make a considered choice as to whether litigation should be subsidized at all, much less by the litigation opponent of the discovering party. Nor have voters had a meaningful opportunity to hold their representatives accountable for the subsidies that they have heretofore chosen to provide.

2. *Implications for the Efficiency of the Discovery Process*

The previous Subsection demonstrates that our current model of discovery cost allocation gives rise to serious issues of democratic accountability. Our current method of cost allocation has also created a perverse set of economic incentives that have profoundly affected discovery practice in negative ways. The subsidization of discovery costs creates two distinct and equally problematic incentives—one of which arises out of the scope of the subsidy, and one of which arises out of the source of the subsidy. Using the model of cost avoidance first developed by Judge Guido Calabresi, this Subsection explains why the current method of cost allocation—the full subsidization of discovery costs by the requesting party’s opponent—fuels practices of excessive and abusive discovery. It next argues, under Calabresi’s theory of the “cheapest cost avoider,” that allocating discovery costs to the requesting party would significantly reduce the problematic incentives created by the litigation subsidy by removing the troubling externalities that currently dominate the process, thereby increasing the efficiency of the process.

a. *Primary Cost Reduction and the Cheapest Cost Avoider*

A number of years ago, then-Professor Calabresi proposed a three-tiered model of cost allocation designed to reduce the overall societal costs brought about by a problematic activity.¹¹³ The first tier

Court, the Rules Enabling Act, and the Politicization of the Federal Rules: Constitutional and Statutory Implications, 90 MINN. L. REV. 1303, 1304–07, 1319–31 (2006).

¹¹² See *supra* notes 99–101 and accompanying text.

¹¹³ See GUIDO CALABRESI, *THE COSTS OF ACCIDENTS* 26–31 (1970). Calabresi’s initial model was focused on avoidance of the costs of accidents. *Id.* Subsequently, both Calabresi and other scholars expanded the theory into other areas of substantive law. For recent examples of the application of these principles in different legal contexts, see Shmuel I. Becher, *A “Fair Contracts” Approval Mechanism: Reconciling Consumer Contracts and Conventional Contract*

of the model, “primary cost reduction,” seeks to reduce the overall costs incurred by the system by imposing financial responsibility on the actor in the best position to determine how the accidents can most cheaply be avoided.¹¹⁴ “Secondary cost reduction,” by contrast, is concerned with the proper allocation of costs once they have been incurred, with the goal of spreading these costs in order to cushion their impact.¹¹⁵ Finally, the third tier, tertiary cost reduction, tells us to question whether the measures taken to reduce costs (of either the primary or secondary variety) actually cost more than they save.¹¹⁶ The general aim of Calabresi’s model, regardless of the particular context in which it is applied, is to determine the optimal combination of primary, secondary, and tertiary cost reduction, so that costs are reduced within the limits of economic efficiency.¹¹⁷

In most cases, Calabresi theorized, primary cost reduction can be most effectively accomplished through use of an individualized, market-based approach. His theory of “general deterrence” is premised on the notion that certain behavior, although imposing systemic costs on society, either cannot or should not be subjected to a blanket prohibition.¹¹⁸ Instead, he argued, society can reduce the overall costs created by the harmful behavior by imposing the associated costs on

Law, 42 U. MICH. J.L. REFORM 747, 783 (2009) (noting that sellers are the lowest cost avoiders in consumer contracts), Timothy R. Holbrook, *Equivalency and Patent Law’s Possession Paradox*, 23 HARV. J.L. & TECH. 1, 29 (2009) (questioning the notion that patentees are the lowest cost avoiders on questions of equivalency), Gerald Korngold, *Resolving the Intergenerational Conflicts of Real Property Law: Preserving Free Markets and Personal Autonomy for Future Generations*, 56 AM. U. L. REV. 1525, 1563–64 (2007) (arguing that, in a transaction for real property, the grantor is the lowest cost avoider), Scott A. Moss & Peter H. Huang, *How the New Economics Can Improve Employment Discrimination Law, and How Economics Can Survive the Demise of the “Rational Actor,”* 51 WM. & MARY L. REV. 183, 246 (2009) (noting that the complaint requirement for employment discrimination plaintiffs appears to assume that the victim is the lowest cost avoider because he knows the most about the harassment), Ariel Porat, *A Comparative Fault Defense in Contract Law*, 107 MICH. L. REV. 1397, 1411–12 (2009) (discussing how to encourage low-cost avoidance of overreliance on contracts), and Wilson L. White, Comment, *Attorney-Client Privilege as a Patent Sword and Shield: The Role of the Adverse Inference Rule in the Efficiency of the Patent System*, 84 N.C. L. REV. 1049, 1077 (2006) (arguing that the public, and not the patentee, is the lowest cost avoider in the context of patent infringement).

¹¹⁴ See CALABRESI, *supra* note 113, at 68–72.

¹¹⁵ *Id.* at 39.

¹¹⁶ *Id.* at 28.

¹¹⁷ *Id.* at 29.

¹¹⁸ *Id.* at 68–69. This element of general deterrence is different from the other potential method of primary cost avoidance: “specific deterrence.” Specific deterrence involves a blanket prohibition of specific acts or activities thought to incur costs. All decisions pertaining to these prohibitions should be made collectively through the political process, wherein all the benefits and costs of the activity would be evaluated, and a collective, transparent decision would be made regarding how much of each activity would be allowed and how much of each activity

the people who wish to engage in that activity, thereby appropriately reducing its attractiveness to market participants.¹¹⁹ In this manner, the cost is to be imposed on the party best able to determine the most cost-effective means of reducing or avoiding the harm in question. To work effectively, however, the scheme requires the initial allocation of costs to the actors who were in a position to avoid the costs most cheaply.¹²⁰ Put another way, costs are to be allocated to the party that is in “the best position to make the cost-benefit analysis between accident costs and accident avoidance costs and to act on that decision once it is made.”¹²¹ Generally speaking, the cheapest cost avoider is the individual with the maximum degree of internalization of the costs among the participants in the transaction,¹²² after taking into account any externalization due to transfers¹²³ or inadequate knowledge.¹²⁴ The theory of general deterrence proceeds from the postulate “that individuals know what is best for themselves,” as long as all efforts, consistent with the interests of economic efficiency, have been taken to limit those costs and their occurrence, and the participants are fully aware of an activity’s true costs.¹²⁵

Calabresi’s market-based approach recognizes the possibility that some individuals will still choose to engage in cost-generating behavior, regardless of the price established by the market.¹²⁶ Thus, any complete model of cost allocation must not only explain how to reduce systemic costs overall but also how to distribute the costs that are still inevitably incurred. Calabresi refers to the latter goal as secon-

would be performed. *Id.* at 95–96. Put another way, activities would be openly judged to be desirable or undesirable and would be penalized or subsidized accordingly. *Id.* at 96.

¹¹⁹ *Id.* at 26.

¹²⁰ *See id.* at 72.

¹²¹ Guido Calabresi & Jon T. Hirschoff, *Toward a Test for Strict Liability in Torts*, 81 *YALE L.J.* 1055, 1060 (1972) (emphasis omitted).

¹²² CALABRESI, *supra* note 113, at 144.

¹²³ Externalization after transfers requires a comparison of the cost-avoidance potential of not only the initial cost bearers, but also the avoidance potential of those who would actually bear the costs after transfers. *Id.* at 147–48. For example, if the costs at issue would be fully subsidized by the government or an insurance company, one must take the cost-avoidance potential of those entities, and not the initial cost bearer, into account when determining which entity is the cheapest cost avoider.

¹²⁴ Externalization due to inadequate knowledge comes into play when the potential cheapest cost avoider cannot make an accurate *ex ante* estimate of the cost of engaging in an activity. *Id.* at 148. If inadequate knowledge prevents the individual from making such an assessment, putting the cost on him would not affect his behavior and would thus have as little effect on primary cost reduction as scattering the cost would. *Id.*

¹²⁵ *Id.* at 72.

¹²⁶ *See id.* at 18–19.

secondary cost reduction¹²⁷ and offers two potential methods for accomplishing it. The first is the “loss spreading method,” which seeks to minimize the net impact of the cost on payers by distributing it among the largest possible group of people.¹²⁸ The second is the “deep pocket notion,” which posits that “secondary losses can be reduced most by placing them on the categories of people least likely to suffer substantial social or economic dislocations as a result of bearing them, usually thought to be the wealthy.”¹²⁹ Theoretically, partial spreading can reduce secondary costs more effectively than total spreading if the right people are forced to pay.¹³⁰ Calabresi emphasizes, however, that secondary reduction cannot be the only aim of a cost allocation model because neither loss spreading nor deep pockets do anything to discourage the behavior that actually generates the costs.¹³¹ Thus, the transfer of financial responsibility to an effective model must include mechanisms that further all three types of cost reduction.

b. Why the Litigation Subsidy Fails to Reduce the Primary and Secondary Costs of Discovery

The subsidization of discovery costs, as currently structured, fails to optimally reduce either the primary or the secondary costs of discovery. As noted above, two distinct aspects of the litigation subsidy create problematic economic incentives. First, full subsidization¹³² renders discovery costs a complete externality for the requesting party, preventing any incentive to reduce the amount of discovery sought. Second, subsidization by the requesting party’s opponent also creates a perverse incentive by allowing the requesting party to use the cost of discovery as a weapon to force its opponent to settle the case before the costs are incurred, or, at the very least, to burden the opponent by increasing his costs unnecessarily. Moreover, this incentive possibly gives rise to an inefficient increase in discovery. This Subsection addresses each of these incentives in turn and concludes by arguing that because the requesting party is the cheapest cost avoider,

¹²⁷ *Id.* at 28.

¹²⁸ *Id.* at 39–40.

¹²⁹ *Id.* at 40.

¹³⁰ *Id.* at 40–41.

¹³¹ *Id.* at 43.

¹³² As noted in Part I, some courts have ordered cost shifting in disputes involving e-discovery. *See supra* notes 30–32, 38 and accompanying text. These cases, however, are exceptions to the general rule. For purposes of argument, this Part assumes that insofar as the producing party has disputed a discovery request, its motion to limit discovery or shift cost has been denied by the court.

forcing it to bear all costs associated with its discovery requests reduces primary costs and eliminates problematic incentives.

i. Full Subsidization of Discovery Costs

In his discussion of primary cost reduction, Calabresi points out that a complete subsidy is not the optimal way to advance the goal of primary cost avoidance.¹³³ Indeed, the effectiveness of his model of general deterrence depends on the allocation of costs to the cheapest cost avoider, a concept that is defined in part as the party with the greatest degree of internalization of costs.¹³⁴ A full subsidy prevents the identification of the cheapest cost avoider because it removes all internalization of the costs associated with the problematic activity. Instead, it reduces costs to a complete externality for the requesting party, thereby removing any incentives to avoid harmful use of the behavior in question.

Calabresi's warning has certainly proven true in the context of discovery costs, in which full subsidization has arguably led to an overall increase in costs.¹³⁵ Aside from the comparatively minimal costs of drafting their discovery requests and considering the responses, litigants bear none of the costs of producing the information that they demand.¹³⁶ Subsidization—through allocation of the total costs to the responding party—renders discovery costs a complete externality¹³⁷ and removes all incentives for litigants to limit the scope of their requests.¹³⁸ In fact, because a party bears none of the costs associated with the fulfillment of its requests, a rational party “will request information that increases the expected value of her legal claim by a little, even though compliance costs the other party a lot.”¹³⁹ Judge Frank Easterbrook noted that our system's current model of cost allocation leads “[l]awyers practicing in good faith [to] engage in extensive discovery; anything less is foolish.”¹⁴⁰

¹³³ See CALABRESI, *supra* note 113, at 37.

¹³⁴ *Id.* at 144.

¹³⁵ See Bruce L. Hay, *Civil Discovery: Its Effects and Optimal Scope*, 23 J. LEGAL STUD. 481, 500 (1994) (“Discovery is costly, and many of its costs are externalized by the requesting party.”).

¹³⁶ See *id.*

¹³⁷ An externality is “a cost or benefit resulting from a decisionmaker's activity that does not accrue to the decisionmaker and is thus ‘external’ to his decisionmaking process.” Note, *Discovery Abuse Under the Federal Rules: Causes and Cures*, 92 YALE L.J. 352, 352 n.5 (1982).

¹³⁸ Hay, *supra* note 135, at 500–01.

¹³⁹ Robert D. Cooter & Daniel L. Rubinfeld, *Reforming the New Discovery Rules*, 84 GEO. L.J. 61, 65 (1995).

¹⁴⁰ Frank H. Easterbrook, *Discovery as Abuse*, 69 B.U. L. REV. 635, 641 (1989).

The externalization of discovery costs, accomplished through the de facto hidden litigation subsidy caused by our current model of cost allocation, incentivizes what can most appropriately be called excessive discovery. Excessive discovery, it should be emphasized, is by no means identical to abusive discovery. The latter, as explained in more detail in the next Subsection, is discovery designed primarily to burden, harass, or intimidate a litigation opponent.¹⁴¹ In contrast, excessive discovery refers to a situation in which there exists a small potential to produce useful information, but where that small likelihood is outweighed, under a rational economic analysis, by the costs to which the process would give rise.¹⁴² Thus, like abusive discovery costs, it “includes costs that are not—on an objective basis—necessary to the fair and accurate adjudication of the case.”¹⁴³ In other words, an attorney engaging in excessive discovery might serve his opponent with a broad, costly discovery request simply to cover all his bases and ensure that he has not missed the proverbial something that might help his client’s case. For a paradigmatic example of excessive discovery, one need look no further than the seminal case of *Hickman v. Taylor*,¹⁴⁴ in which the plaintiff’s attorney served its opponent with a set of thirty-nine interrogatories.¹⁴⁵ One of these interrogatories asked “whether any statements of the members of the crews of the Tugs ‘J.M. Taylor’ and ‘Philadelphia’ or of any other vessel were taken in connection with the towing of the car float and the sinking of the Tug ‘John M. Taylor.’”¹⁴⁶ The interrogatory immediately following requested that the defendant “[a]ttach hereto exact copies of all such statements if in writing, and if oral, set forth in detail the exact provisions of any such oral statements or reports.”¹⁴⁷ The plaintiff’s attorney admitted that he only requested the oral statements “to help prepare himself to examine witnesses and to make sure that he [had] overlooked nothing.”¹⁴⁸

The rise of electronic discovery has likely increased the frequency of excessive discovery, as computers have enabled parties to create, transmit, and store information “in quantities that would have been

¹⁴¹ See *infra* notes 154–56 and accompanying text.

¹⁴² See Redish, *supra* note 7, at 602–03.

¹⁴³ *Id.* at 602; see also Note, *supra* note 137, at 353.

¹⁴⁴ *Hickman v. Taylor*, 329 U.S. 495 (1947).

¹⁴⁵ *Id.* at 498.

¹⁴⁶ *Id.*

¹⁴⁷ *Id.* at 499 (internal quotation marks omitted).

¹⁴⁸ *Id.* at 513.

unthinkable twenty-five years ago.”¹⁴⁹ More recent examples of excessive discovery include one plaintiff’s request for the restoration of 2500 backup tapes (containing seven years of e-mail) at an estimated cost of \$1 million,¹⁵⁰ and another plaintiff’s request to restore and search five years’ worth of e-mails written by nineteen different employees of the defendant.¹⁵¹ The proposed search included 170 different search terms, many of which were commonly used industry terms that could be found in almost any work-related communication.¹⁵² Even if these requests were intended solely to uncover information relevant to the case, the fact remains that the subsidization of the requesting party’s discovery by its opponent removes all incentive for the requesting party to tailor its request narrowly in order to impose the least possible burden on the producing party. Were the costs of the discovery to be imposed on the requesting party, it would not automatically follow that the discovery would not take place. Rather, the decision as to whether the discovery would take place would be made by the party who has to incur the cost of the process.

ii. Subsidization by the Producing Party and Incentives for Discovery Abuse

The economic problems caused by the subsidization of a party’s discovery are significantly exacerbated by the source of that subsidy. Because a party’s opponent is the source of the litigation subsidy, the requesting party also has a perverse incentive to make the request as broad and expensive as possible in order to impose costs on its opponent. In other words, as one of us has previously written, “the bigger the expense to be borne by the opponent, the bigger the incentive to make the request.”¹⁵³ Subsidization by the requesting party’s opponent has led to the phenomenon commonly referred to as “discovery abuse,” which is defined as a discovery request “justified by the costs it imposes on one’s adversary rather than by the gains to the requester derived from the contribution the information will make to the accu-

¹⁴⁹ Scheindlin & Rabkin, *supra* note 28, at 366; *see also* Redish, *supra* note 7, at 588–92 (arguing that the volume and difficulty of retrieval and translation necessitates special treatment for electronic discovery under the Federal Rules of Civil Procedure).

¹⁵⁰ Major Tours, Inc. v. Colorel, No. 05-3091 (JBS/JS), 2009 WL 3446761, at *1 (D.N.J. Oct. 20, 2009).

¹⁵¹ Quinby v. WestLB AG, 245 F.R.D. 94, 99 (S.D.N.Y. 2006).

¹⁵² *Id.* Still other terms were words that might be found in any message, such as “go,” “her,” “okay,” and “she.” *Id.*

¹⁵³ Redish, *supra* note 7, at 603; *see also* Note, *supra* note 137, at 352–53.

racy of the judicial process.”¹⁵⁴ Abusive requests are “motivated by goals other than the exchange of information fairly related to the issues in dispute”¹⁵⁵ and are generally designed to “force favorable settlements by driving up the other party’s discovery costs beyond the case’s value, calculated in terms of the likelihood of a favorable outcome, the value of such an outcome, and the cost of litigating the case to conclusion.”¹⁵⁶ Although excessive discovery might be defended on the grounds that, at the very least, it is motivated on some level by a legitimate desire to uncover information that will contribute to the rendering of an accurate decision—regardless of the likelihood of that actually happening—abusive discovery has no redeeming systemic value.

To be sure, the Federal Rules authorize judges to sanction parties for discovery abuse. Rule 26(g) states that by signing a discovery request, an attorney or party certifies that it is “not interposed for any improper purpose, such as to harass, cause unnecessary delay, or needlessly increase the cost of litigation”¹⁵⁷ and that it is “neither unreasonable nor unduly burdensome or expensive, considering the needs of the case, prior discovery in the case, the amount in controversy, and the importance of the issues at stake in the action.”¹⁵⁸ The Rule further authorizes the court to impose an “appropriate sanction” on any party who violates it, which can include the “reasonable expenses, including attorney’s fees, caused by the violation.”¹⁵⁹ In practice, however, these standards have not proven effective as a means of either identifying or limiting discovery abuse. A judicial officer supervising discovery cannot make an accurate *ex ante* determination of the productivity of a given request “because the nature of the requester’s claim and the contents of the files (or head) of the adverse party are unknown.”¹⁶⁰ As one scholar recently pointed out, there exists a profound information-timing problem during discovery; the Federal Rules require judges to conduct cost-benefit analyses of discovery requests at a point in the litigation at which they cannot possibly have

¹⁵⁴ Easterbrook, *supra* note 140, at 637–38.

¹⁵⁵ Earl C. Dudley, Jr., *Discovery Abuse Revisited: Some Specific Proposals to Amend the Federal Rules of Civil Procedure*, 26 U.S.F. L. REV. 189, 193 (1992).

¹⁵⁶ *Id.* at 194; *see also* Hay, *supra* note 135, at 501 (“The plaintiff may be tempted to use discovery strategically, as a means of imposing costs on the defendant in the hopes of extracting a settlement unrelated to the merit of her claim.”).

¹⁵⁷ FED. R. CIV. P. 26(g)(1)(B)(ii).

¹⁵⁸ *Id.* 26(g)(1)(B)(iii).

¹⁵⁹ *Id.* 26(g)(3).

¹⁶⁰ Easterbrook, *supra* note 140, at 638.

access to all the information necessary to make an accurate assessment.¹⁶¹ As a result of this inability to distinguish abusive requests from legitimate requests that might nevertheless result in “dry holes,” judges are often hesitant to deny discovery or to shift the costs to the requesting party.¹⁶² The existing sanctions therefore do little to curb the problematic and often perverse economic incentives caused by the decision to force a party to subsidize its opponent’s discovery costs.

iii. Allocating Discovery Costs to the Cheapest Cost Avoider

The first step in aligning discovery cost allocation under Judge Calabresi’s model, particularly the goal of primary cost reduction, would be to require identification of the cheapest cost avoider. As explained above, the cheapest cost avoider is the party best positioned to make an accurate ex ante analysis of the costs and benefits of the targeted behavior and to act on that decision after it has been made.¹⁶³ In the context of discovery costs, the requesting party is, for the most part, unambiguously the cheapest cost avoider. It is, of course, the discovering party who makes the initial decision of whether and what to request. Although a court may theoretically intervene to limit that request, at best the tertiary costs created by judicial intervention are substantial, and at worst the process is of little use.¹⁶⁴ Absent internalization of cost, the requesting party has no meaningful incentive to curb his request in the interests of economic efficiency.

It is true that a litigant might not always have perfect knowledge of the exact evidence he will need in order to prevail. But forcing a litigant to bear the reasonable costs of his discovery requests will force him and his attorney to make a thorough ex ante evaluation of what information they definitely need and what information will provide only a marginal benefit. Moreover, nothing prevents attorneys from tackling discovery in stages or filing narrower discovery requests to determine whether a more expansive request is warranted. In fact, at least one court ordered sampling of expensive electronic discovery in order to determine whether the benefits of an expansive request by the plaintiff outweighed the costs to the defendant.¹⁶⁵

¹⁶¹ Moss, *supra* note 29, at 912–18 (explaining the information-timing problems inherent in each element of the balancing test prescribed by Rule 26(b)(2)).

¹⁶² William H. Wagener, Note, *Modeling the Effect of One-Way Fee Shifting on Discovery Abuse in Private Antitrust Litigation*, 78 N.Y.U. L. REV. 1887, 1897 (2003).

¹⁶³ Calabresi & Hirschhoff, *supra* note 121, at 1060.

¹⁶⁴ See *supra* notes 28–29 and accompanying text.

¹⁶⁵ See, e.g., *Zubulake v. UBS Warburg LLC*, 217 F.R.D. 309, 323–24 (S.D.N.Y. 2003) (or-

Concededly, a counterargument can be fashioned. One can make a plausible response that it is the producing party who is in the best position to reduce costs because it can exercise control over the methods it uses to respond to its opponent's request. For example, the producing party decides whether to outsource document production to contract attorneys or whether to hire a law firm to do it. Although this may be true to a point, the responding party's role in controlling costs is limited because it is entirely reactionary. The responding party has no power to limit the primary driver of its discovery costs, which is undoubtedly the scope of the initial request. It is true that if discovery costs were allocated to the requesting party, bill padding would be a legitimate concern, as the producing party would possess some of the same perverse incentives that currently lead to discovery abuse. But there are at least two ways in which these pathological practices could be curtailed. First, parties could be expected to set out parameters for anticipated discovery costs during the discovery conference, and at least some of the dangers could be detected and avoided at that point. In addition, the court (quite probably through a magistrate judge) could exercise control over these disputes by limiting the producing party's recovery to the objectively reasonable market value of the work.¹⁶⁶ At the very least, this analysis is much less burdensome and unwieldy, and more amenable to objective scrutiny, than the seven- and eight-factor balancing tests currently being employed by a few courts to adjudicate e-discovery cost-shifting disputes.¹⁶⁷ Concerns over a responding party's bill padding, then, are not sufficient justification for abandoning a model of cost allocation that would eliminate the starkly perverse economic incentives that currently plague the system.¹⁶⁸

3. *Constitutional Implications*

The procedural system's failure to acknowledge its current method of discovery cost allocation as a litigation subsidy has also

dering UBS to restore e-mails from five of its ninety-four backup tapes in order to obtain a factual basis to support the cost-shifting analysis).

¹⁶⁶ This would accord with the recovery traditionally awarded in actions brought under the theory of quantum meruit. See *supra* note 49 and accompanying text.

¹⁶⁷ See *supra* note 32 and accompanying text (describing the *Zubulake* balancing test).

¹⁶⁸ Moreover, Calabresi's model requires the incorporation of tertiary cost-reduction mechanisms, which monitor the effectiveness of any measures taken to reduce primary or secondary costs. CALABRESI, *supra* note 113, at 28. Thus, under this model, incidents of bill padding would be monitored, and the system would be adjusted if these disputes did, in fact, create a problematic amount of satellite litigation.

masked the constitutional problems to which use of that method gives rise. Our current system requires the producing party, in the first instance, to bear the financial burden of fulfilling any discovery requests made by its opponent. Rule 37 provides the court with broad power to enforce the litigation subsidy by means of a variety of sanctions that can, in most situations, include imprisonment for contempt of court.¹⁶⁹ This Subsection demonstrates that the forced subsidization of its opponent's discovery costs gives rise to serious concerns about the procedural due process rights of the producing party.

The Due Process Clause of the Fifth Amendment provides that “[n]o person shall . . . be deprived of life, liberty, or property, without due process of law.”¹⁷⁰ The Supreme Court has consistently recognized money as a “property” interest protected by the Due Process Clause, extending constitutional protection to bank accounts,¹⁷¹ wages,¹⁷² welfare payments,¹⁷³ and disability benefits.¹⁷⁴ Any deprivation of a constitutionally recognized “life, liberty, or property” interest triggers the aggrieved party's right to be afforded due process of law.¹⁷⁵ Although the type and amount of process that is due varies according to the circumstances of the case, the Supreme Court has established three essential principles of due process: “notice and op-

¹⁶⁹ FED. R. CIV. P. 37(b).

¹⁷⁰ U.S. CONST. amend. V.

¹⁷¹ *N. Ga. Finishing, Inc. v. Di-Chem, Inc.*, 419 U.S. 601, 606 (1975) (describing a bank account as “surely a form of property”).

¹⁷² *Sniadach v. Family Fin. Corp. of Bay View*, 395 U.S. 337, 340 (1969) (“A prejudgment garnishment of the Wisconsin type is a taking which may impose tremendous hardship on wage earners with families to support.”).

¹⁷³ *Goldberg v. Kelly*, 397 U.S. 254, 262 & n.8 (1970) (“It may be realistic today to regard welfare entitlements as more like ‘property’ than a ‘gratuity.’”).

¹⁷⁴ *Mathews v. Eldridge*, 424 U.S. 319, 332 (1976) (“[T]he interest of an individual in continued receipt of [disability] benefits is a statutorily created ‘property’ interest protected by the Fifth Amendment.”).

¹⁷⁵ The Court most clearly articulated this operational principle in *Cleveland Board of Education v. Loudermill*, 470 U.S. 532, 541 (1985), in which it stated that “the Due Process Clause provides that certain substantive rights—life, liberty, and property—cannot be deprived except pursuant to constitutionally adequate procedures.” *Accord Zinermon v. Burch*, 494 U.S. 113, 125 (1990) (“In procedural due process claims, the deprivation by state action of a constitutionally protected interest in ‘life, liberty, or property’ is not in itself unconstitutional; what is unconstitutional is the deprivation of such an interest *without due process of law.*”); *Eldridge*, 424 U.S. at 332–33 (proceeding directly from the recognition of a “property” interest to a determination of what procedures are required); *Bd. of Regents v. Roth*, 408 U.S. 564, 569–70 (1972) (“When protected interests are implicated, the right to some kind of prior hearing is paramount.”); *Bell v. Burson*, 402 U.S. 535, 542 (1971) (“While [m]any controversies have raged about . . . the Due Process Clause, it is fundamental that . . . when a State seeks to terminate [a protected] interest[,] . . . it must afford ‘notice and opportunity for hearing appropriate to the nature of the case’ *before* the termination becomes effective.” (emphasis added) (citation omitted)).

portunity for hearing appropriate to the nature of the case”¹⁷⁶ that must be conducted by a neutral adjudicator.¹⁷⁷ The scope and timing of the constitutionally mandated notice and hearing are currently determined by application of a three-part balancing test established in *Mathews v. Eldridge*, which does not mandate any specific procedures.¹⁷⁸ Because the constitutional implications of the accepted presumption in the discovery cost allocation context vary for the defendant and the plaintiff, the next two Subsections address the constitutional implications for each side.

a. Defendant

The mere filing of a lawsuit by a plaintiff, in and of itself, does not establish a defendant’s liability. A complaint is nothing more than a series of unproven allegations made by an interested party. To impose the nonreimbursable costs of plaintiff’s discovery on the defendant on the basis of nothing more than the plaintiff’s unilateral allegation of liability surely takes defendant’s property without due process. The judicial process has imposed a financial burden on the defendant without even a preliminary judicial finding of wrongdoing. In a long line

¹⁷⁶ *Loudermill*, 470 U.S. at 542 (internal quotation marks omitted); *see also Eldridge*, 424 U.S. at 333 (“This Court consistently has held that some form of hearing is required before an individual is finally deprived of a property interest.”); *Fuentes v. Shevin*, 407 U.S. 67, 97 (1972) (“[D]ue process is afforded only by the kinds of ‘notice’ and ‘hearing’ that are aimed at establishing the validity . . . of the underlying claim against the alleged debtor *before* he can be deprived of his property.” (emphasis added) (internal quotation marks omitted)); *Armstrong v. Manzo*, 380 U.S. 545, 550 (1965) (holding that lack of notice “violated the most rudimentary demands of due process of law”); *Baldwin v. Hale*, 68 U.S. (1 Wall.) 223, 233 (1863) (“Common justice requires that no man shall be condemned in his person or property without notice and an opportunity to make his defence.”); Trevor W. Morrison, *Suspension and the Extrajudicial Constitution*, 107 COLUM. L. REV. 1533, 1611 (2007) (describing the three core requirements of notice, hearing, and a neutral adjudicator as a “reasonably concrete operative proposition”); Edward L. Rubin, *Due Process and the Administrative State*, 72 CALIF. L. REV. 1044, 1173 (1984) (arguing that a hearing is a necessary component of due process of law and that it should be provided before a deprivation of liberty except when “the purpose of institutionalizing the individual was to remove an immediate, demonstrable danger”).

¹⁷⁷ *See, e.g., Hamdi v. Rumsfeld*, 542 U.S. 507, 537–38 (2004) (holding that an individual detained by the government was entitled to a neutral adjudicator as a matter of due process of law); *Concrete Pipe & Prods. of Cal., Inc. v. Constr. Laborers Pension Trust*, 508 U.S. 602, 617–18 (1993) (“[O]ne is entitled as a matter of due process of law to . . . a neutral and detached adjudicator.”); *Tumey v. Ohio*, 273 U.S. 510, 522 (1927) (“That officers acting in a judicial or quasi-judicial capacity are disqualified by their interest in the controversy to be decided is, of course, the general rule.”). Because any procedures implicating the Federal Rules regarding discovery cost allocation would be conducted before an Article III or federal magistrate judge, this Article assumes for purposes of argument that the requirement of a neutral adjudicator is satisfied.

¹⁷⁸ *Eldridge*, 424 U.S. at 334–35.

of cases, the Supreme Court has held that taking a defendant's property on the basis of unilateral, judicially untested allegations made by an adversary violates the Due Process Clause because the procedure fails to provide defendants with an adequate hearing.¹⁷⁹ Those decisions appear to be directly applicable to the discovery cost allocation context.

It is true that both Federal Rule 12(b)(6), which allows dismissal for "failure to state a claim upon which relief can be granted,"¹⁸⁰ and Federal Rule 12(c), which enables a party to obtain judgment on the pleadings,¹⁸¹ provide the defendant with an opportunity for a certain kind of judicial hearing prior to the discovery process. But the mere availability of some form of a hearing is not necessarily sufficient to satisfy the constitutional guarantee of due process of law. In evaluating a motion to dismiss, the court proceeds on the assumption that all of the plaintiff's factual allegations are true¹⁸² and determines only whether those factual allegations, when assumed to be factually true, "allow[] the court to draw the reasonable inference that the defendant is liable for the misconduct alleged."¹⁸³ Adjudication of a 12(b)(6) or 12(c) motion is not intended to serve as an evaluation of the actual merits of the claim, as neither party, at that point in the process, has had the opportunity to uncover or present evidence to support its case. The evaluation of the sufficiency of a pleading is merely intended to screen out the complaints of those plaintiffs armed with nothing more than an "unadorned, the defendant-unlawfully-harmed-me accusation."¹⁸⁴ Even if the court suspects that the plaintiff is ultimately unlikely to prevail after a determination of the facts, it is required to allow him to proceed to discovery so long as his complaint contains sufficiently detailed factual allegations to suggest that discovery will enable him to uncover evidence in support of his claim.¹⁸⁵

At no point during an evaluation of the pleadings on either a 12(b)(6) or 12(c) motion is the defendant allowed to offer his own

¹⁷⁹ See *Fuentes*, 407 U.S. at 82–83 (discussing Supreme Court precedent). The Court proceeded to note that, as a practical matter, the constitutional requirements "may not even test that much" because "an uneducated, uninformed consumer" might not be aware of the procedures enabling him to challenge the seizure of his property. *Id.* at 83 n.13.

¹⁸⁰ FED. R. CIV. P. 12(b)(6).

¹⁸¹ *Id.* 12(c).

¹⁸² *Ashcroft v. Iqbal*, 129 S. Ct. 1937, 1949 (2009) ("To survive a motion to dismiss, a complaint must contain sufficient factual matter, accepted as true, to 'state a claim to relief that is plausible on its face.'" (quoting *Bell Atl. Corp. v. Twombly*, 550 U.S. 544, 570 (2007))).

¹⁸³ *Id.*

¹⁸⁴ *Id.*

¹⁸⁵ See *Twombly*, 550 U.S. at 556.

explanation for the allegedly unlawful conduct or to offer his own version of the events in question.¹⁸⁶ Under Rule 8(b), a party, in its answer, may only “state in short and plain terms its defenses to each claim asserted against it” and “admit or deny the allegations asserted against it by an opposing party.”¹⁸⁷ Even his ability to challenge the veracity of the plaintiff’s assertions is somewhat limited. Rule 11 does allow the court to impose sanctions for factual allegations under limited circumstances.¹⁸⁸ But that provision requires merely that factual allegations are based on an “inquiry reasonable under the circumstances” and makes clear that a plaintiff need not have firsthand knowledge of the truth of the allegation at the pleading stage.¹⁸⁹ Rather, he need only certify that his factual contentions “will likely have evidentiary support after a reasonable opportunity for further investigation or discovery.”¹⁹⁰ Rule 11 recognizes that although a plaintiff may have good reason to believe that his allegation is true, he might need the court-enforced discovery devices to compile or confirm his evidentiary basis.¹⁹¹ Although the plaintiff retains his duty to conduct an appropriate investigation into the allegations made on information and belief, the Rule, as it is currently written, grants a plaintiff substantial latitude to make factual allegations of which he has no personal knowledge.¹⁹²

In *Connecticut v. Doehr*,¹⁹³ the Supreme Court held that the one-sided evaluation of the factual sufficiency of a complaint did not constitute due process of law because it failed to adequately reduce the risk of an erroneous deprivation of property, as required by the *Ma-*

¹⁸⁶ This Article assumes, for purposes of argument, that the defendant has not elected to bring a counterclaim under Federal Rule of Civil Procedure 13(b). Even if he did file his own suit against the plaintiff, however, the points made in Part II.B.3.b, *infra*, would still apply.

¹⁸⁷ FED. R. CIV. P. 8(b)(1). This might be a bit overstated. Even if the defendant does have a (limited) opportunity to present his side of the story in the response, however, the court still must assume that all of the plaintiff’s allegations are true for purposes of deciding on the motion to dismiss.

¹⁸⁸ *Id.* 11(b)(3), (c). It should be noted, however, that the plaintiff has a twenty-one-day safe harbor in which he can withdraw any overreaching or false allegations without being subject to sanctions. *Id.* 11(c)(2).

¹⁸⁹ *Id.* 11(b).

¹⁹⁰ *Id.* 11(b)(3).

¹⁹¹ *Id.* 11(b)(3) advisory committee’s note to 1993 amendment.

¹⁹² See *N. Ga. Finishing, Inc. v. Di-Chem, Inc.*, 419 U.S. 601, 607 (1975) (holding that procedural due process was not satisfied when seizure was authorized based on a writ issuable on the affidavit of the creditor or his attorney, the latter of whom need not have any personal knowledge of the facts).

¹⁹³ *Connecticut v. Doehr*, 501 U.S. 1 (1991).

thews balancing test.¹⁹⁴ In *Doehr*, the Court refused to allow the lower court to authorize attachment of real property based on the allegations of a complaint filed in a separate lawsuit:

Permitting a court to authorize attachment merely because the plaintiff believes the defendant is liable, or because the plaintiff can make out a facially valid complaint, would permit the deprivation of the defendant's property when the claim would fail to convince a jury, when it rested on factual allegations that were sufficient to state a cause of action but which the defendant would dispute The potential for unwarranted attachment in these situations is self-evident and too great to satisfy the requirements of due process¹⁹⁵

The Court also made clear that the sine qua non of a due process hearing is the ability of the judge to make a "realistic assessment concerning the likelihood of an action's success."¹⁹⁶ The evaluation of a complaint, however, occurs before the parties have had the opportunity to gather or present information in support of their claims. Even putting aside the Supreme Court's repeated statements that an evaluation of the pleadings should not reflect a judge's perception of the underlying factual merits at this early stage of the litigation,¹⁹⁷ a judge simply does not have the information necessary to evaluate accurately the likelihood that the plaintiff will prevail. Thus, the adjudication of a motion to dismiss is not a constitutionally adequate hearing, and the financial burdens of the litigation subsidy brought about by the prevailing discovery cost allocation presumption are levied on the defendant in violation of his right to procedural due process.

In reaching this conclusion, it is important to distinguish the burdens caused by the forced subsidy from the normal costs incurred by a defendant in preparing his own case after a complaint is filed. Unlike the costs incurred by a defendant in mounting his own case, the costs involved in responding to a plaintiff's discovery requests are a financial benefit that the defendant is required—at the risk of severe sanctions—to provide to the plaintiff on the basis of nothing more than the unilateral filing of the plaintiff's complaint.

¹⁹⁴ See *id.* at 13–14.

¹⁹⁵ *Id.*

¹⁹⁶ *Id.* at 14.

¹⁹⁷ See, e.g., *id.* at 13–14.

b. Plaintiff

There is no doubt that a defendant can make a much stronger constitutional argument challenging the imposition of discovery costs upon it than can a plaintiff.¹⁹⁸ Still, it is at least arguable that even the plaintiff's rights to procedural due process are violated by the litigation subsidy that the current discovery cost allocation presumption requires it to provide the defendant. First, unlike a complaint (which requires at least some level of factual detail), the defendant's response need not contain any factual allegations that reveal how it plans to rebut the plaintiff's claim.¹⁹⁹ Of course, the implications for a plaintiff who is unwittingly roped into expensive discovery are different from the implications for a similarly situated defendant. The plaintiff, unlike his opponent, can choose to abandon his case at any time if he can no longer afford the costs of litigation. This response does not, however, provide a complete answer to the constitutional problem posed by the potentially inadequate notice.

It is well established that a plaintiff's chose in action constitutes a protected property interest.²⁰⁰ Like the defendant, then, the plaintiff is constitutionally entitled to a hearing before his property can be taken. Although a plaintiff cannot seek pre-discovery judicial intervention by invoking Rule 12(b)(6),²⁰¹ once the defendant files his answer, a plaintiff could conceivably file a motion for judgment on the pleadings under Rule 12(c).²⁰² Similar to the court's analysis of a motion to dismiss, however, the court must assume that all of the factual allegations in the pleadings are true and should grant the plaintiff's motion only if, based solely on the pleadings, "the moving party is entitled to judgment as a matter of law."²⁰³ Like the motion to dismiss for failure to state a claim, a motion for judgment on the pleadings does not involve a judicial analysis of the substantive merits of the

¹⁹⁸ Indeed, a plaintiff might be strategically unwilling to even make such an argument because, generally speaking, plaintiffs are the primary beneficiaries of the litigation subsidy.

¹⁹⁹ If the defendant plans to make an affirmative defense, he must include it in the responsive pleading under Federal Rule of Civil Procedure 8(c)(1). The Rule does not, however, require the defendant to provide any details showing a factual basis for that defense.

²⁰⁰ See, e.g., *Phillips Petrol. Co. v. Shutts*, 472 U.S. 797, 807 (1985); *Mullane v. Cent. Hanover Bank & Trust Co.*, 339 U.S. 306, 313 (1950).

²⁰¹ FED. R. CIV. P. 12(b)(6). This Article assumes, for the sake of simplicity, that the defendant has not filed a counterclaim or crossclaim under Federal Rule of Civil Procedure 13. If the defendant did so, the plaintiff's rights regarding discovery pertaining to that claim would mirror the discussion in Part II.B.3.a.

²⁰² *Id.* 12(c).

²⁰³ *Burns Int'l Sec. Servs., Inc. v. United Plant Guard Workers Local 537*, 47 F.3d 14, 16 (2d Cir. 1994) (per curiam).

underlying claim. Indeed, such a determination would be impossible at this early stage of the litigation because neither party has yet had the opportunity to use the court-enforced discovery devices to gather evidence in support of its claims. Thus, as is true for the defendant, the prediscovery hearing provided by the court is not sufficient to satisfy the plaintiff's right to procedural due process.

In assessing the due process implications for the plaintiff of the cost allocation presumption, it is also important not to ignore the practical realities of the situation. Armed with the presumption that the responding party must bear the costs of discovery, a defendant can easily seek to employ the discovery process to throw artificial obstacles in the plaintiff's path. Although procedural devices designed to deter abusive discovery exist in the Federal Rules, such devices are inconsistent at best in achieving this goal.²⁰⁴ Thus, imposing the costs of defendants' discovery requests on plaintiffs at least threatens to undermine plaintiffs' abilities to vindicate their substantive rights.

III. ALTERING THE DISCOVERY COST ALLOCATION PRESUMPTION: THE OPTIONS

The preceding Subsections established three foundational principles for building a morally, democratically, and economically sound model of discovery cost allocation. First, under the foundational moral principle of economic justice, a producing party should be compensated for any benefit that it confers on its opponent. Thus, allocating discovery costs to anyone other than the requesting party effectively creates a litigation subsidy. Second, any subsidization of litigation—to the extent one concludes it should be used—should either be explicitly adopted in the Federal Rules, so that Congress may make a transparent decision as to whether it wishes to alter that presumption, or made by Congress in the first instance. Third, when viewed from the perspective of economic efficiency, the primary goal of a cost allocation system is the *ex ante* deterrence of unnecessary cost-generating behavior. This result is best accomplished by allocating costs to the party in the best position to determine whether the costs are in fact efficient. Finally, imposing costs of an opponent's discovery on either party prior to a judicial evaluation of the substantive merits of the claim gives rise to a substantial constitutional problem.

²⁰⁴ See *supra* notes 25–29 and accompanying text.

This Part incorporates these principles into constructive proposals for the allocation of discovery costs. First, it offers a purely hypothetical proposal describing how discovery costs should be allocated if one could magically return to the time of the Federal Rules' original promulgation and thus begin by writing on a blank procedural slate. Second, it explains how the three foundational principles already established can reshape judicial attitudes toward discovery cost allocation, even under the current procedural framework.

A. *Reconstructing Discovery Cost Allocation Based on Moral, Economic, and Democratic First Principles: Why Choose to Subsidize Discovery Costs?*

As already noted, the litigation subsidy imposed in the discovery cost allocation context has always been unstated, unexplained, and untested. This Section discusses how one should resolve the discovery cost allocation question were one to consider the question openly without reference to the last seventy years of judicial practice.

From its inception, the discovery process was designed to increase the accuracy of adjudication. Theoretically, the more information produced during discovery, the more accurately the factfinder will be able to assess the merits of the case.²⁰⁵ Enhanced accuracy of the legal process would likely increase compliance with the law because potential wrongdoers should logically be deterred from breaking the law, due to the recognition that they will probably be held legally accountable for their actions.²⁰⁶ Thus, one might fear that, absent subsidization of discovery costs, certain areas of substantive law might be underenforced. As one of us has previously noted, absent subsidization of discovery costs, “[m]any laws enacted for the purpose of assisting those in a vulnerable political or economic position—for example, civil rights laws—could be undermined.”²⁰⁷ Society might also choose to subsidize litigation that furthers the public interest or even a private interest that extends beyond those of the actual parties to a given suit.²⁰⁸ If society in general stands to benefit from the outcome of a

²⁰⁵ See *Hickman v. Taylor*, 329 U.S. 495, 507 (1947) (“Mutual knowledge of all the relevant facts gathered by both parties is essential to proper litigation.”).

²⁰⁶ Hay, *supra* note 135, at 502; see also Albert W. Alschuler, *Mediation with a Mugger: The Shortage of Adjudicative Services and the Need for a Two-Tier Trial System in Civil Cases*, 99 HARV. L. REV. 1808, 1813 (1986) (noting that, in the context of court costs, subsidization “furthers a general perception that wrongdoing will prove unprofitable”).

²⁰⁷ Redish, *supra* note 7, at 607.

²⁰⁸ See Thomas D. Rowe, Jr., *The Legal Theory of Attorney Fee Shifting: A Critical Overview*, 1982 DUKE L.J. 651, 662.

lawsuit, it is morally sound to disperse the costs of litigation among all the beneficiaries.²⁰⁹

With these considerations in mind, we return to the question whether a litigation subsidy should ever be adopted and, if so, under what circumstances. There are three potential answers to this question: (1) subsidize all such costs, as in the current system; (2) subsidize none of the costs, thereby requiring all parties to pay their own way by reimbursing the responding party for costs incurred; or (3) subsidize some of the costs, either within particular areas of substantive law or for particular litigants who are unable to pay. The first option merits little additional discussion, as the underlying problems identified earlier (the moral implications of the unjust enrichment of the requesting party, the problematic economic incentives, and the procedural due process violation) would still apply. At the very least, however, from the perspective of democratic theory, formal and open ratification of the litigation subsidy would constitute a marginal improvement over the status quo because the issue would have been transparently considered and legally adopted, and voters would therefore have the opportunity to hold their elected representatives accountable for their choice.

The second, “pay your own way,” option, while it most fully addresses both the quantum meruit and economic incentive problems that have been identified, might be politically and socially infeasible. This system, while at least superficially fair to both parties, might have the effect of preventing the financially disadvantaged from enforcing their substantive rights. Rather than facilitating the enforcement of underlying substantive law, such a cost allocation system conceivably could completely frustrate it, particularly with respect to types of cases brought by comparatively poor litigants.²¹⁰

²⁰⁹ Antitrust and environmental litigation are two examples of such suits.

²¹⁰ See, e.g., Redish, *supra* note 7, at 607 (“Unlimited cost-shifting could significantly exacerbate preexisting economic disparities between litigants, and, where the costs involved are prohibitive, seriously threaten achievement of the goals underlying the governing substantive law by effectively precluding private enforcement.”); see also, e.g., Alschuler, *supra* note 206, at 1812–17; Panel Discussion, *supra* note 28, at 24 (quoting Hon. James C. Francis IV, District Judge, United States District Court for the Southern District of New York) (“I think there is a fear among requesting parties that they will be shut out of discovery to the extent that there is cost shifting because they simply cannot afford to bear a portion of those costs.”). *But see* Easterbrook, *supra* note 140, at 646 (arguing that concerns that cost shifting would freeze out litigants of modest means is overblown because “[i]mpositional discovery is practiced in big-stakes cases between substantial litigants, represented by the most costly legal talent. This problem should be tackled, with the difficulties of impoverished and middle-class litigants carved off for different treatment if need be.”).

The third includes two potential middleground options under which a subset of discovery costs would be subsidized. The first, which is foreshadowed above, would subsidize discovery costs for certain areas of substantive law that the legislature deems to be of particular social value or fears would be undermined by forcing parties to pay their own way. The second would subsidize discovery costs for litigants otherwise unable to afford to proceed with the suit.

1. *Subsidization of Particular Areas of Substantive Law*

Congress is already well accustomed to singling out particular areas of substantive law for special treatment with respect to the allocation of legal costs. Many statutes contain fee-shifting arrangements designed to encourage individuals to file claims,²¹¹ and some of the motivation behind that legislation might overlap with the rationales for subsidization of discovery costs. For example, two of the major purposes of fee-shifting legislation include promoting public interest litigation and compensating the prevailing plaintiff.²¹² Congress could certainly use these statutes as starting points for debates regarding the appropriate targets for discovery cost subsidization. It is crucial to note, however, that the fee-shifting arrangements within these statutes generally compensate only the plaintiff for his legal fees, and then only if he is victorious.²¹³ Both plaintiffs whose claims ultimately fail and defendants receive nothing. Essentially, fee-shifting laws are designed to encourage the filing of *meritorious* claims, meaning claims that are more likely than not to succeed at trial. These caveats are crucial because they give rise to two additional permutations of a potential discovery cost subsidy.

On the one hand, Congress might simply extend the fee-shifting statutes to discovery costs and subsidize only the costs incurred by the plaintiff. But, as noted above, *both* parties are forced to incur discov-

²¹¹ See, e.g., Clayton Act § 4(a), 15 U.S.C. § 15(a) (2006). Other statutes that are mandatory in terms of awarding attorney's fees include the Fair Labor Standards Act, 29 U.S.C. § 216(b) (2006), the Truth in Lending Act, 15 U.S.C. § 1640(a)(3) (2006), and the Merchant Marine Act of 1936, 46 U.S.C. § 58106(c) (2006). Under the Civil Rights Attorney's Fees Awards Act of 1976, 42 U.S.C. § 1988(b) (2006), the prevailing party is entitled to attorney's fees at the discretion of the court, but the Supreme Court has construed the statute as intending to award attorney's fees absent exceptional circumstances. See *Newman v. Piggie Park Enters.*, 390 U.S. 400, 402 (1968). Other statutes authorize courts to award attorney's fees in "exceptional cases." See, e.g., 35 U.S.C. § 285 (2006) (allowing the prevailing parties in patent infringement suits to be awarded reasonable attorney's fees). See generally *Alyeska Pipeline Serv. Co. v. Wilderness Soc'y*, 421 U.S. 240, 260 n.33 (1975) (listing fee-shifting statutes).

²¹² Vargo, *supra* note 15, at 1588.

²¹³ See *supra* note 211.

ery costs before any judicial analysis of the merits of the underlying claim. From a moral perspective, it seems wrong to force a defendant to bear (potentially substantial) costs because a plaintiff made a unilateral, autonomous decision to sue him. One could also argue, however, that in order to survive a motion to dismiss, a plaintiff must provide sufficient factual detail in his complaint to allow the judge to make a “reasonable inference that the defendant is liable” for the alleged unlawful conduct.²¹⁴ Even if this threshold screening requirement were deemed sufficient to remove any moral misgivings caused by imposing costs on the defendant while simultaneously subsidizing the plaintiff, however, it is still constitutionally inadequate for purposes of procedural due process.²¹⁵

On the other hand, Congress might choose to subsidize both plaintiffs and defendants in particular types of lawsuits. It might be politically unseemly, however, to enact a subsidy that benefits both individual plaintiffs and the corporate defendants who allegedly wronged them. This seems particularly true in high-visibility areas of public interest litigation, where the effects of the corporation’s wrongful conduct extend beyond the parties to the individual case.²¹⁶ Moreover, fully subsidizing discovery costs for a particular area of substantive law would not eliminate the perverse economic incentives created by our current system of cost allocation that have resulted in substantial inefficiencies.²¹⁷ Because discovery costs would remain an externality for the subsidized party,²¹⁸ that party would retain the incentive to draft overly broad requests. Thus, even if Congress attached the subsidy to particular laws, it might still have to enact some type of income-based requirement.

2. *Subsidization Based on Financial Need*

Subsidization based on the litigant’s “ability to pay” is the second of the two middleground options. If this alternative were chosen, it would probably make sense to have the subsidization come from the government, rather than from the litigation opponent.

To be sure, this option is significantly more complex than the other three, as the government would first have to define “ability to

²¹⁴ *Ashcroft v. Iqbal*, 129 S. Ct. 1937, 1949 (2009).

²¹⁵ And one could imagine that defendants would be more willing to challenge the constitutionality of the requirement if they receive no benefit from it.

²¹⁶ Products liability, antitrust, and environmental litigation are potential examples.

²¹⁷ *See supra* Part II.B.2.

²¹⁸ The only potential difference is that that the government would be paying the discovery costs, not the requesting party.

pay” and then would have to evaluate each litigant to see if he satisfies the criteria. Finally, the court would have to apportion the costs according to the government’s findings. Nonetheless, these questions are not insurmountable, and this option, if correctly implemented, would enable underfunded litigants to enforce their rights while realigning the parties’ economic incentives and ensuring that neither party is unjustly enriched by the work of its opponent.

The threshold question for this option is obvious: how would the government determine an individual’s “ability to pay”? The most obvious starting point would be to assess the financial situation of the litigant himself. Income, as reported by one’s federal tax return, would likely be insufficient on its own, as it only reports what an individual takes in.²¹⁹ Perhaps the most feasible option would involve a system similar to the Free Application for Federal Student Aid (“FAFSA”), which takes account of all of an applicant’s relevant financial data, including income, savings, investments, and property.²²⁰ It then calculates, according to a predetermined formula, an individual contribution representing the amount of money that the applicant can be reasonably expected to contribute toward his education.²²¹ The larger point here is that the government already employs analytical systems and formulas that could be easily adapted for use in determining whether an individual qualifies for a litigation subsidy. It would therefore likely not require substantial time or money in order to put one into place.

One might argue that a substantial risk of manipulation inheres in any system that relies on self-reporting of financial information. There are two answers to this argument. First, the government can minimize this risk by creating a program in which an individual will always have to pay some portion of his discovery costs, albeit an adjusted amount that reflects his particular financial situation.²²² As

²¹⁹ A tax return, for example, does not report the net value of any savings or investments the individual has; it reports only the gains or losses that he incurs from them.

²²⁰ See FAFSA, <http://www.fafsa.ed.gov> (last updated Jan. 30, 2011).

²²¹ See *id.*

²²² The alternative is to create a firm baseline for benefits over which a person would not be eligible at all and under which a person would have all of his costs subsidized. Such a restriction would incentivize people who fall above the cutoff to conceal information in order to satisfy the eligibility requirement. There is less incentive to manipulate, however, when one is only talking about the degree of benefit.

Another potential counterargument is that administrative barriers might lead to incomplete utilization of these benefits—meaning a poor, uneducated person who suffered some sort of civil rights violation might not be able to locate, much less file, a claim for a litigation subsidy. See Janet Currie, *The Take-Up of Social Benefits*, in ALAN. J. AUERBAC ET AL., PUBLIC POLICY AND

long as the individual is forced to bear a portion of his own costs, he retains an incentive not to make excessive discovery requests. Such a proportional-pay requirement also keeps the subsidy aligned with the moral considerations identified earlier because the requesting party would not receive the fruits of the producing party's labor for free. Second, the government can make clear that if a party is caught falsifying any information in its subsidy application, it will be immediately subject to potential criminal penalties.

Although the litigant's personal finances provide a logical starting point for determining whether he can afford to pay his own discovery costs, under certain circumstances it may also make sense to consider the resources of his attorneys as well. This is particularly true in noninjunctive class actions, where it is accepted practice for plaintiff class attorneys to fund their clients' suits. Class discovery costs could thus be deemed a cost of doing business for the attorneys. If the plaintiff is successful, the firm might include discovery costs in the fee it assesses from the client's award. Such a scenario would encourage plaintiffs' firms to seek out and litigate potentially meritorious claims (i.e., claims that are more likely than not to result in monetary compensation for the client), while discouraging the filing of frivolous claims for which the firm might ultimately be left bearing the discovery costs.

There are, no doubt, longstanding legal doctrines that limit attorneys' abilities to make financial investments in their clients' cases. Strictly construed, however, these doctrines are irrelevant to the issue whether an attorney can legally fund his client's discovery requests. One example, champerty, is defined as an "agreement to divide litigation proceeds between the owner of the litigated claim and a party unrelated to the lawsuit who supports or helps enforce the claim."²²³ The doctrine's underlying policy is to discourage excessive, unnecessary, or speculative litigation, which is often associated with third parties seeking profit for themselves, rather than redress for their clients, through suits.²²⁴ The allocation of discovery costs, however, has nothing to do with the proceeds of the case. The attorney would simply bear the costs of his own labor on behalf of his client, as he can indis-

THE INCOME DISTRIBUTION 80, 87-111 (2006) (arguing that administrative barriers are a significant factor in explaining incomplete utilization of means-tested benefits). He might not even know what a litigation subsidy is.

²²³ BLACK'S LAW DICTIONARY 262 (9th ed. 2009).

²²⁴ See PETER H. SCHUCK, *AGENT ORANGE ON TRIAL: MASS TOXIC DISASTERS IN THE COURTS* 202 (1987).

putably do under a contingent agreement. If his client prevails, he may recover his costs out of the award. This element of risk for the plaintiff's attorney will advance, rather than undermine, the policy goals that animate the champerty doctrine, as rational attorneys will only be willing to fund cases that are likely to result in reimbursement of their costs.²²⁵ It also bears mention that champerty, although still technically recognized by most states, is enforced with varying degrees of zeal across different jurisdictions.²²⁶ One federal court recently remarked that champerty is "no longer part of the *argot* of lawyers and courts in this country as it once was."²²⁷ Indeed, many courts have held that there is no longer a cause of action for champerty or its predecessor, maintenance.²²⁸

3. *How Would the Subsidy Be Funded?*

The final question to be addressed by Congress would be how to fund the subsidy. This Subsection offers two preliminary suggestions on that issue. The first is a public litigation fund, created by Congress through the legislative process, and most likely administered by a governmental agency. All citizens would pay into the fund through a federal tax. Creation of a public litigation fund would foster Judge Calabresi's goal of reducing secondary costs in discovery cost allocation by spreading these costs over as many payers as possible.²²⁹ Cala-

²²⁵ Likewise, barratry, defined as "[v]exatious incitement to litigation, esp. by soliciting potential legal clients," BLACK'S LAW DICTIONARY 170 (9th ed. 2009), is not relevant here.

²²⁶ See Paul Bond, Comment, *Making Champerty Work: An Invitation to State Action*, 150 U. PA. L. REV. 1297, 1300–02, 1333–41 (2002). New York courts, for example, generally hesitate to deem an action champertous as a matter of law. *Bluebird Partners v. First Fid. Bank*, 731 N.E.2d 581, 586 (N.Y. 2000) (citing *Sprung v. Jaffe*, 147 N.E.2d 6, 8–9 (N.Y. 1957)); see also *Weigel Broad. Co. v. Topel*, No. 83 C 7921, 1985 WL 2360, at *6–7 (N.D. Ill. Aug. 21, 1985) (holding that, under Illinois law, a cause of action for maintenance exists, but explaining that it is so rare in modern times that the law on the subject is neither settled nor clear). Delaware and Minnesota, by contrast, still enforce the doctrine vigorously. See *Hall v. State*, 655 A.2d 827, 830 (Del. Super. Ct. 1994); *Johnson v. Wright*, 682 N.W.2d 671, 680 (Minn. Ct. App. 2004).

²²⁷ *In re Rite Aid Corp. Sec. Litig.*, 146 F. Supp. 2d 706, 718 (E.D. Pa. 2001).

²²⁸ "Maintenance" is defined as "an officious intermeddling in a suit which in no way belongs to the intermeddler by maintaining or assisting either party to the action, with money or otherwise, to prosecute or defend it." *Hall*, 655 A.2d at 829 (internal quotation marks omitted). Several state courts have declined to recognize causes of action under either doctrine. See *Hardick v. Homol*, 795 So. 2d 1107, 1108 (Fla. Dist. Ct. App. 2001) (holding that modern causes of action for abuse of process supersede champerty and maintenance); see also *PSI Metals, Inc. v. Firemen's Ins. Co.*, 839 F.2d 42, 43 (2d Cir. 1988) (applying New York law); *Sec. Underground Storage, Inc. v. Anderson*, 347 F.2d 964, 969 (10th Cir. 1965) (applying Kansas law); *McCullar v. Credit Bureau Sys., Inc.*, 832 S.W.2d 886, 887 (Ky. 1992); *Saladini v. Righellis*, 687 N.E.2d 1224, 1226 (Mass. 1997); *Osprey, Inc. v. Cabana Ltd. P'ship*, 532 S.E.2d 269, 277–78 (S.C. 2000).

²²⁹ See CALABRESI, *supra* note 113, at 39.

bresi posited that society should allocate costs in the way that produces the least societal harm, either through the deep pocket notion of allocating costs to the wealthiest portions of society, or the loss spreading method of distributing the costs as broadly as possible.²³⁰ If Congress chooses to use the deep pocket model, it could create a graduated tax that imposes the highest rate on the wealthiest members of society. Although such a burden would rest heaviest on those most able to afford it, it also arguably creates a bit of a moral quandary, as the group being taxed most heavily would be the least likely to benefit from the fund. By contrast, if Congress chose the loss spreading model, it could enact a tax with a flat rate across all income brackets. Although the wealthiest members of society would still be the largest contributors to the litigation fund, at least they would not be burdened at a higher rate than the rest of society. A third option, which would also fall under the loss spreading model, would be a regressive tax, which either removes or reduces the tax rates for incomes in excess of a specified amount. Technically speaking, a regressive tax might be the fairest, as it would impose the greatest proportional burden on the individuals who would be most likely to utilize the fund.²³¹ On the other hand, a regressive tax would likely not produce as much revenue as a flat-rate tax, because the latter would tax wealthier individuals based on the entirety of their income.

If a litigation fund were ultimately found not to be politically feasible (which, at this point in time, may be likely), Congress might choose to require the producing party to bear the costs of its underfunded opponent, much as it does in our current cost allocation system. Although such legislation would not alter the status quo as a practical matter, at the very least, the determination of who should bear the costs of discovery would be made openly and transparently through resort to the formal democratic process. However, as noted previously, such an approach may give rise to constitutional concerns, grounded in the dictates of procedural due process,²³² as well as to all the perverse economic incentives which we have described above.²³³

²³⁰ *Id.* at 39–40.

²³¹ The government could set the cap for the tax at roughly the point at which individuals would cease to qualify for a litigation subsidy. To be sure, it would be impossible to determine that cutoff for any given case because some types of claims almost inevitably require burdensome and expensive discovery. Nevertheless, for purposes of the tax, the cutoff point need not be exact. The government could simply choose to use either the average or median discovery costs calculated across all cases.

²³² See discussion *supra* Part II.B.3.

²³³ See discussion *supra* Part II.B.2.

B. Using First Principles to Guide Current Discovery Disputes

Although the proposals examined in the prior Section have merit as a theoretical matter, they likely would be infeasible in the current political climate. That does not mean, however, that no practical benefits can derive from recognition of the proper model of discovery cost allocation and rejection of the hidden litigation subsidy that has been present in the prevailing model for more than seventy years. The first principles articulated in our argument can, and indeed should, be used to resolve two current issues in discovery cost allocation. At the very least, our argument shows that “cost *shifting*,” as it is currently defined, is a complete misnomer. Foundational principles of economic justice—freed from the baggage of the last seventy years of poorly thought-out practice—clearly dictate that a party who requests that a service be performed should compensate the performing party in an amount commensurate to the benefit received. A judge who requires a party to bear the costs of its discovery requests is not shifting costs. Rather, he is upholding the time-honored moral principle that it is unjust to allow one person to enrich himself at the expense of another. The true cost shifting occurs when courts upend this moral and economic equilibrium by forcing one party to subsidize part of its opponent’s case. Moreover, this covert subsidy creates a host of problematic economic incentives by removing any mechanism for primary cost reduction and by encouraging litigants to use the subsidy as a weapon to force their opponents to settle (or abandon) the case.

Hampered by unwieldy balancing tests that are, as an *ex ante* matter, weighted against cost shifting and constrained by the governing presumption that parties should bear their opponents’ discovery costs, courts remain reluctant to shift costs in all but the most egregious cases.²³⁴ Our analysis suggests that courts should take precisely the opposite perspective. Because ordering the requesting party to bear its own discovery costs accords with fundamental moral and economic principles, courts should employ a presumption in favor of relieving the reimbursing party for some, if not all, of its costs. If appropriate, the court may adjust the proportion of costs to be reimbursed by the requesting party to correspond with its individual resources.

The latter consideration implicates our second nonlegislative suggestion: incorporation of the plaintiff’s resources into the court’s anal-

²³⁴ Giacobbe, *supra* note 36, at 269 (noting that the majority of cost-shifting requests are denied).

ysis of the requesting party's financial resources, at least in the class action context. As noted above, most states have eliminated any legal barriers to allowing an attorney to contribute to the costs of his case.²³⁵ Courts have consistently reasoned that corporate defendants are well equipped to absorb discovery costs as part of their general costs of doing business.²³⁶ But that expectation has yet to be applied in a neutral manner to all litigants. Both the *Zubulake* and *Rowe Entertainment* tests include the resources of the requesting party as one of the factors in the cost-shifting calculus,²³⁷ yet none of the courts applying these tests have incorporated the resources of the plaintiffs' attorneys into their estimation of the parties' abilities to pay for discovery.²³⁸ At the very least, attorneys who stand to benefit if their client wins or settles its case (by virtue of contingent fee contracts) should be forced to absorb some of the costs of that case as one of their own costs of doing business.²³⁹

CONCLUSION

For over seventy years, the presumption that discovery costs lie where they fall has largely gone unchallenged by both courts and scholars. An evaluation of fundamental moral, economic, democratic, and constitutional principles reveals that our current method of discovery cost allocation rests on a theoretical foundation that is at best shaky, and at worst completely illusory.

As a practical matter, it might be too late in the game to completely overhaul the litigation system in order to incorporate insights from the underlying theoretical framework that we propose. But when entertaining novel cost allocation issues, particularly in the context of electronic discovery or big cases, courts can use our theory to overcome their lingering reluctance to upset the dominant—though foundationally suspect—paradigm of discovery cost allocation.

²³⁵ See *supra* notes 226–28 and accompanying text.

²³⁶ See, e.g., *Daewoo Elecs. Co. v. United States*, 650 F. Supp. 1003, 1006 (Ct. Int'l Trade 1986) (“[The] normal and reasonable translation of electronic data into a form usable by the discovering party should be the ordinary and foreseeable burden of a respondent in the absence of a showing of extraordinary hardship.”).

²³⁷ See *supra* note 32.

²³⁸ See *supra* note 38 (listing cases applying the *Zubulake* and *Rowe Entertainment* tests).

²³⁹ The same result would not necessarily apply to attorneys on either side of the case whose compensation is not affected by the outcome of the litigation.