Litigation Discovery Cannot Be Optimal But Could Be Better: The Economics of Improving Discovery Timing in a Digital Age

Scott A Moss
LITIGATION DISCOVERY CANNOT BE OPTIMAL BUT COULD BE BETTER:
THE ECONOMICS OF IMPROVING DISCOVERY TIMING IN A DIGITAL AGE
Scott A. Moss

ABSTRACT

Cases are won and lost in discovery, yet discovery draws surprisingly little academic attention. Most scholarship focuses on how much discovery to allow, not how courts decide discovery disputes – which, unlike trials, occur in most cases. Today, much evidence is “e-discovery” – imprudent emails or still-lingering “deleted” files – making costly discovery battles increasingly salient. But the e-discovery rules are not truly new, just a strengthening of old cost/benefit “proportionality” limits on discovery enacted when the spread of copiers similarly increased the amount of discovery. Proportionality limits are topic of broad consensus among civil procedure scholars as well as economists concerned that discovery excess yields extortionate settlements and results from parties ignoring costs they impose on others.

Contrary to the consensus, this Article deems proportionality rules impossible to apply effectively. By both failing to curb discovery excess and disallowing discovery that meritorious cases need, ineffective proportionality limits let bad cases predominate over good cases, just as the bad can drive out the good in product markets. This Article acknowledges proportionality’s flaws but rejects the consensus blaming bad rulemaking or judging. Rather, proportionality requires impossible comparisons: how can courts compare discovery cost to evidentiary value before the parties gather the evidence? Like other arguments that procedural rulings are never truly separate from case merits, this Article explains how discovery has more probative value in close cases than in the strongest and weakest cases. Yet case merits remain uncertain in discovery, when courts are not yet able to examine all the evidence.

In game theory terms, parties with discovery disputes cannot convey case merit credibly; courts have too little information, so low-merit parties can claim high merit, and courts act as if all cases warrant similar discovery. In this “pooling equilibrium,” ruling the same on all cases in the “pool,” regardless of merit, is courts’ best strategy but a sub-optimal one, yielding too much discovery in low-merit cases, too little in higher-merit cases. Thus, the quest for better discovery has disappointed not because of bad rules or cases, but because courts and parties are stuck in a pooling equilibrium. This is an information-timing circularity: optimal evidence-gathering requires merits analysis, which requires more evidence-gathering.

One answer is to defer close decisions on possibly useful but costly evidence until meritorious cases separate from the pool, turning a pooling equilibrium into a “separating equilibrium.” Summary judgment can be this separating point: cases going to trial after summary judgment likely have 50/50 odds – better than most cases. Costly evidence has more value in a 50/50 case, where the jury will struggle to reach a verdict, than a very weak (or very strong) case. Nobody yet has proposed solving the costly discovery dilemma with post-summary judgment discovery (summary judgment typically follows discovery), but high-cost evidence can be an exception: cases surviving summary judgment are the close calls warranting more fact-gathering, so some costly discovery commonly denied now should be allowed later, after summary judgment. While imperfect, this solution could improve the status quo, and imperfection is inevitable given the fundamental information timing problem that prevents accurate proportionality decisions.

Thus, the prevailing debate too narrowly focuses on discovery quantity and could benefit from focusing more on discovery timing. Interestingly, existing rules give courts the discretion to use this proposal, but a new rule could minimize the risk of courts misusing the proposal to deny more discovery. This Article concludes by noting how economic analysis of litigation must do more than prescribe cost-benefit comparisons; it must also consider the details and information timing of the litigation process.
LITIGATION DISCOVERY CANNOT BE OPTIMAL BUT COULD BE BETTER: THE ECONOMICS OF IMPROVING DISCOVERY TIMING IN A DIGITAL AGE

Scott A. Moss*

TABLE OF CONTENTS

I. Introduction: The Need to Move Beyond “Proportionality” Limits on Litigation Discovery .......................... 1

II. The Rise of Proportionality Limits on Litigation Discovery – and Their Disappointing Results ... 6
   A. What is Old is New Again: Proportionality Requirements as Solution to Tech-Driven Discovery Excess ................................................................. 6
   B. The E-Discovery Rules: Just More Disappointingly Conventional and Ineffective Proportionality Limits ........................................................................ 11

III. An Economic Analysis of Why Proportionality Limits, Though Popular, Cannot be Optimal .... 14
   A. The Consensus: Limit Discovery Based on Cost/Benefit Proportionality Principles .......... 14
   B. An Economic View of Discovery: Helping Factfinders Assess Case Value and Merit
      \((L \text{ and } p)\) – Which Makes Discovery Decisionmaking Circular ..................................... 16
      1. Size of Case \((L)\): Difficult to Determine in Many Cases ........................................... 17
      2. Probative Value of the Evidence Sought \((\Delta p)\): Difficult to Assess Before Fully Analyzing That and Other Evidence ......................................................................... 18
      3. Probability the Claim Has Merit \((p)\): Difficult for Court to Assess Before Seeing All the Evidence and the Parties’ Arguments about That Evidence ....................... 20
         a. Why Courts Do Not Consider Case Merit in Making Discovery Decisions:
            The Conventional Wisdom that Discovery is Unrelated to Case Merit ........................................ 20
         b. Why Courts Should Consider Case Merit: Optimal Discovery Depends on Whether a Case is a “Close Call” .................................................................................. 21
         c. Why Courts Cannot Easily Consider Case Merit: In Discovery, Cases Are in a “Pooling Equilibrium” in which Parties Cannot Signal, and Courts Cannot Assess, Case Merit Effectively ........................................................................ 24

IV. Solving the Pooling that Prevents Better Discovery Decisions: In Close Calls about Costly Discovery, Preserve the Evidence But Delay the Discovery until After Summary Judgment ..... 28
   A. Replacing The Pooling Equilibrium with a Separating Equilibrium as a Case Progresses .... 28
   B. Summary Judgment as the Key “Separating” Process that Allows Courts to Distinguish Cases by Level of Merit \((p)\) ..................................................................................................................... 30
   C. The Prescription: In Close Calls, Preserving the Evidence But Delaying the Discovery Until After Summary Judgment ................................................................................. 35
   D. The Devil is in the Details: Making Workable the Proposal for Post-Summary Judgment Revisiting of Discovery the Court Had Denied Earlier ......................................................... 38
      Concern #1: Summary Judgment May Not Clarify Case Merit ............................................ 38
      Concern #2: Courts Should Use Alternatives Such as Sampling and Cost-Shifting .......... 40
      Concern #3: Judges Might Excessively Deny Discovery and Grant Summary Judgment ..... 41
      Concern #4: Judicial Reluctance to Delay Trial to Reargue a Discovery Dispute ............ 46
      Concern #5: Loss of Evidence while Discovery is Delayed .................................................. 47

V. Conclusion ................................................................................................................................. 49

*Associate Professor, University of Colorado Law School (scott.moss@colorado.edu). The author thanks the following: for reviewing and editing drafts of this Article, Professors Robert Bone, Roberto Corrada, Martha Minow, Paul Ohm, Nantiya Ruan, Michael Waggoner, and Phil Weiser; for offering feedback on presentations of this Article, the faculty, attorneys, and judges who attended conferences and colloquia at the University of Colorado Law School, the Benjamin N. Cardozo School of Law, the Boston University Law School, the Chapman Law School, the Colorado Employment Law Faculty (CELF) working group, the Second Annual Labor/Employment Law Colloquium, and the 2007 Midwest Law and Economics Association (MLEA) Meeting; and for their excellent work as research assistants, University of Colorado Law School students Tiffany Howard and Lori Mayall.
I. INTRODUCTION: THE NEED TO MOVE BEYOND “PROPORTIONALITY” LIMITS ON LITIGATION DISCOVERY

As all litigators know, “discovery . . . is the battleground where civil suits are won and lost.”1 “It is in discovery that the facts are developed,” sometimes leading to a dismissal of the case on summary judgment2 but other times to “vital evidence” that yields a trial win or a “favorable settlement.”3 Discovery’s cost attests to its importance: in the federal case, discovery is half of all litigation costs; in the most expensive 5% of cases, discovery amounts to 90% of litigation cost and totals 32% of the amount in controversy.4 Yet on no other topic is there more disconnect between academy and bar; discovery controversies often are “not something that law professors pay a lot of attention to, but lawyers do.”5 While discovery is the topic of numerous bar publications and conferences,6 it draws far fewer academic articles than hotter civil rights issues,7 even though discovery is key to civil rights litigation. Most of the limited academic writing focuses on how much discovery is too much for parties to request, at least without paying more of the production cost.8 Academic attention rarely focuses on how courts decide discovery disputes, which, unlike trials, occur in most lawsuits, frustrating judges and parties alike. More such focus is critical; if cases are won and lost in discovery, then they are really won and lost in discovery decisions.

2 Steinfield & Bertsche, supra note 1, at 107.
3 Green & Francis, supra note 1, at 238.
6 For example, the Practicing Law Institute’s annual conference, Electronic Discovery and Retention Guidance for Corporate Counsel, yields a lengthy symposium issue. See 766 PLI/LIT. 1-425 (2007) (17 articles); 747 PLI/LIT. 1-403 (13 articles); 733 PLI/LIT. 1-408 (2005) (14 articles).
7 An ExpressO search of top-50 law school specialty reviews found thirteen on women’s/gender issues, seven on constitutional/civil rights, and only one on litigation.
8 See infra Parts II(B), III(A) (discussing proposals to limit discovery and cost-shifting proposals).
The difficulty of deciding discovery disputes took on a new salience once computerization brought litigation into a “brave new world” of costly “e-discovery” where much evidence (perhaps 90% of corporate data) is digital, not paper, and where the best evidence can be email sent in unguarded moments or still-lingering “deleted” incriminating files. In 2006, the Federal Rules of Civil Procedure clarified that “electronically stored information” (“ESI”) is discoverable unless it is “not reasonably accessible because of undue burden or cost.” The “reasonably accessible” rule reflects worry about e-discovery’s “enormous costs . . . becoming the single most expensive facet of litigation.” The amount of digital data is staggering: businesses exchange 2.5 trillion e-mails annually, 2 million at a typical company. The problem is not just quantity, but accessibility, as digital data “can be expensive or virtually impossible to recover,” especially over time, due to “outmoded storage media and software, and dispersion of information.” As one judge noted, “[i]t is hard to overstate the importance and . . . anxiety generated by electronic discovery, . . . not just in the world of big business, [but] organizations generally, large data producers.”

Yet this brave new world may not be so new. Applying old litigation rules “to new technology presents additional challenges, . . . [but] this is not the first time someone has argued that the discovery rules are no longer suitable for . . . contemporary discovery.” After all, when the federal rules were enacted, “[t]he


13 For the details of these amendments, see infra note 65.


15 Bronte, supra note 9, at 55 (2007); Christopher S. Rugaber, E-Documents Subject to Stricter Storage, SOUTH FLORIDA SUN-SENTINEL, 3D (Dec. 2, 2006) (noting that companies pay e-discovery consultants over $1 billion a year and must have lawyers review email and “things more difficult to track, like digital photos . . . [on] cell phones and information on removable memory cards”).

16 Kaplan, Electronic Discovery in the 21st Century, supra note 10, at 67 (citations omitted).

17 William R. Maguire, Current Developments In Federal Civil Discovery Practice: Setting Reasonable Limits In The Digital Era, 754 PLI/Lit 169, 175 (2007).


19 Richard L. Marcus, Confronting the Future: Coping with Discovery of Electronic Material, 64 L.
photocopy machine did not exist,” and like all new technologies, it offered a yin and a yang. As with the discovery-multiplying effect of computers, “the volume of paper created grew enormously when photocopies became ubiquitous,” but also, as one court marveled in now-quaint terms, “the modern convenience of photocopying lessens . . . burdensome transportation” of documents, just as computers can make document exchange as easy as emailing attachments.

Because e-discovery is just a new instance of an old problem – technology facilitating more discovery, increasing existing controversy over discovery costs – just as photocopiers led to rules limiting discovery based on cost/benefit “proportionality,” computers have led to similar limits. With proportionality their guiding principle, the new e-discovery rules mandate a balancing test: courts must consider whether there is “good cause” for e-discovery of high “burden or cost.” Such proportionality limits long have been prescribed by law-and-economics scholars – not just those in the free-market, litigation-skeptical “Chicago school” of economics who blame discovery excess for runaway tort liability and extortionate settlements, but also less ideological economists who advocate tougher cost/benefit “proportionality” limits requiring parties to pay more of the cost of demanding heavy discovery. With contrary views largely relegated to the periphery of the economic debate, discovery restrictions are a topic of unusual consensus among civil procedure scholars, economists, and the federal judges who repeatedly pass rules to rein in discovery.

Dissenting from the consensus, this Article believes “proportionality” limits cannot be implemented effectively. Sometimes they do too little, failing to curb discovery excess or allowing costly discovery on meritless claims; other times they do too much, disallowing discovery meritorious cases need. These opposing errors do not average out any more than it is on average comfortable to have one foot frozen and one foot on fire. Where product quality is hard to determine, bad

---

20 Id. at 266.
21 Id. at 266.
26 See infra Part III(A).
products drive good ones out of the market.\textsuperscript{27} Analogously, if bad cases get too much discovery and good cases too little, court dockets will have more bad cases and fewer good cases; \textit{i.e.}, sub-optimal discovery skews the mix of cases.

After noting how proportionality rules have proven disappointing at regulating discovery, and how the e-discovery rules seem headed for a similar fate, this Article parts company with those blaming bad rulemaking or judges too timid to curb discovery.\textsuperscript{28} Rather, the problem is that proportionality rules ask the impossible: that judges decide when the cost of discovery is proportional to some measure of “value” that includes both the \textit{evidence’s value} to jury deliberation and the \textit{case’s value} to the parties and society. This presents a fundamental information timing problem: discovery disputes occur before parties marshal all the evidence, so how can courts measure the value of particular evidence, much less the merits of the case? Like other arguments that litigation procedure rulings cannot be as separate from case merits as commonly assumed,\textsuperscript{29} this Article notes how discovery has the most probative value in “close call” cases than in the strongest and weakest cases (where more evidence is less likely to affect case outcome). Case merits, though critical to discovery decisions, typically remain hidden in a cloud of uncertainty during discovery, because the court is not yet able to sift fully through the evidence and arguments.

This Article applies economics and game theory to analyze courts’ decisions on litigation discovery disputes. Due to the information costs (including time) of assessing case merit during the discovery stage, courts often cannot tell which plaintiffs’ braggadocio is cheap talk and which reflects actual case merit. Consequently, courts are forced to ignore parties’ arguments about case merit and adjudicate discovery disputes as if all cases in the “pool” (all cases with discovery disputes) warrant discovery to the same degree. In game theory terms: courts’ discovery rulings should base on claims’ merit, but merit cannot be communicated effectively (because those with lesser merit can get away with claiming greater merit); as a result, those rulings must base on the average value of all cases in the pool. In this “pooling equilibrium,” the best available strategy for courts is to rule the same regardless of case merit – even though these rulings are sub-optimal in the sense of yielding too much discovery in low-merit cases and too little discovery in high-merit cases.

Under this analysis, the quest for better discovery limits has disappointed not

\textsuperscript{27} See infra Part III(B)(3)(c).
\textsuperscript{28} See id.
\textsuperscript{29} See, \textit{e.g.}, Robert G. Bone & David S. Evans, \textit{Class Certification and the Substantive Merits}, 51 DUKE L. J. 1251 (2002) (so arguing as to courts deciding on certifying cases as class actions).
because of bad decision-making or bad rule-making, but because courts and parties are stuck in a pooling equilibrium. It is a fundamental information timing problem inherent in the discovery stage of litigation: optimal evidence-gathering decisions require more merits analysis, but merits analyses require more evidence-gathering. As in the folk song, “There’s a Hole in the Bucket” – the hole is fixable only with a machine requiring water poured from the broken bucket\textsuperscript{30} – the problem is a classic circularity; the problem prevents the solution.

One solution to the information problem of a pooling equilibrium is to defer close decisions on potentially useful but costly evidence, until case merit is clearer; \textit{i.e.}, defer such decisions until meritorious cases distinguish themselves, turning a “pooling equilibrium” into a “separating equilibrium.” Fortunately, litigation has such a point: after summary judgment. A case going to trial, having survived summary judgment, has a reasonable probability of merit; economic theory predicts (and some data show) that cases going to trial have 50/50 odds. Additional evidence like costly digital data has more probative value in a 50/50 case than a very weak or strong case; in the 50/50 case, the jury likely faces a close call and so will find more evidence helpful. This proposal aims not to deny or postpone more discovery, but to allow more discovery, after summary judgment, of helpful but costly evidence courts now deem non-discoverable.

Accordingly, much of the scholarly debate on discovery misses the mark by focusing on \textit{how much} costly discovery is warranted, such as with numerical and proportionality limits.\textsuperscript{31} This Article suggests focusing on \textit{when} in litigation to allow such discovery. To date, nobody has suggested solving the dilemma of costly discovery with post-summary judgment discovery, which ordinarily is a counterintuitive idea;\textsuperscript{32} summary judgment typically is after all discovery. Unusually costly evidence should be an exception: surviving summary judgment means a case likely is the sort of close call warranting more fact-gathering, so courts should allow truly costly discovery, such as heavy e-discovery that they commonly disallow now, only once a case survives summary judgment.\textsuperscript{33}

\textsuperscript{30} Harry Belafonte and Odetta are probably the most famous pair to have sung this duet. \textsc{Harry Belafonte, A Man And His Music} (RCA 1990) (track #5). The author first heard it sung by muppets on \textsc{The Muppet Show} in the 1970s but, sadly, cannot find a video clip or proper citation.

\textsuperscript{31} See Marcus, \textit{Confronting the Future}, supra note 19, at 256 (“The retrenchment effort . . . [has] two themes . . . : the principle of proportionality and quantitative limits.”). For discussion of rules applying proportionality and numerical limits, see infra notes 39 and 107, respectively.

\textsuperscript{32} The exception is that courts allow pre-discovery limited-scope summary judgment motions in certain cases, such as those limited to governmental immunity defenses. See infra note 168.

\textsuperscript{33} This Article does not discuss \textit{all} discovery reforms, just one problem that cannot be fully fixed (the impossibility of optimal discovery decisions) and a partial fix well-targeted to that information
Interestingly, no rule change is required to implement this Article’s proposal that courts revisit denials of burdensome discovery if a case survives summary judgment. Existing rules give courts broad case management authority, including as to the extent and sequence of discovery and summary judgment. Thus, this proposal could not only improve litigation discovery, but also provide a welcome answer to courts’ riddle of how to rule on “proportionality” without circular, premature case merit evaluations. A new rule would be advisable, though, to minimize the risk of courts misusing the proposal to deny discovery excessively.

This Article concludes with a brief broader point about economic analysis of law. Fitting into a line of scholarship analyzing litigation as a series of points in time at which information emerges, this Article suggests that if economic analysis of litigation is to provide accurate diagnoses and useful recommendations, it must do more than just prescribe cost-benefit comparisons; instead, it must take account of the timing-and-stages nature of litigation, such as by delving into the details of discovery, pre-litigation settlement, and other litigation events that perhaps have drawn too little academic attention to date.

II. THE RISE OF PROPORTIONALITY LIMITS ON LITIGATION DISCOVERY – AND THEIR DISAPPOINTING RESULTS

A. WHAT IS OLD IS NEW AGAIN: PROPORTIONALITY REQUIREMENTS AS SOLUTION TO TECH-DRIVEN DISCOVERY EXCESS

The most surprising aspect of the new rules about the newest discovery phenomenon – e-discovery – is how un-new those rules are. This is not the first time new rules have targeted discovery excess that new technology facilitated. The federal rules long have prescribed “liberal discovery”\(^34\) that the producing party must pay for itself\(^35\) and a broad relevance standard (discoverable evidence need not be admissible, only “reasonably calculated to lead to the discovery of admissible evidence”\(^36\)). These principles arose in an older, pre-photocopier era,


\[^{35}\text{See Oppenheimer Fund, Inc. v. Sanders, 437 U.S. 340, 358 (1978) (“the presumption is that the responding party must bear the expense of complying with discovery requests”)}\]

\[^{36}\text{Fed. R. Civ. P. 26(b)(1); see Oppenheimer Fund v. Sanders, 437 U.S. 340, 351 (1978) (noting that discovery relevance “has been construed broadly to encompass any matter that bears on, or that reasonably could lead to other matter that could bear on, any issue that is or may be in the case”).}\]
when typical discovery was just on-premises review of original evidence. 37

Like the modern computer revolution, the older photocopying revolution facilitated the discovery overkill we still see – massive document demands and “paper dump” responses – prompting a powerful counter-movement at the highest levels of the legal establishment, including the Attorney General and an ABA “Special Committee on Abuse of Discovery.” 38 The federal judges on the Judicial Conference ultimately enacted the “proportionality” rule that discovery shall be limited by the court... [if the burden or expense of the proposed discovery outweighs its likely benefit, taking into account the needs of the case, the amount in controversy, the parties’ resources, the importance of the issues at stake in the litigation, and the importance of the proposed discovery in resolving the issues]. 39

The sense of urgency behind “proportionality” limits has not dissipated; the Supreme Court in 1998 “signal[led] the importance of the proportionality concept in some cases by quoting portions of Rule 26(b)(2) and observing that... ‘[it] vests the trial judge with broad discretion to tailor discovery narrowly.” 40

The timing of the earlier proportionality movement is striking: coming in the 1980s and 1990s, it shortly preceded the widespread adoption in the 1990s and 2000s of email, networked computers, and the internet. To be sure, though computers were less ubiquitous before the 1990s, major companies 41 and

37 See, e.g., Harris v. Sunset Oil Co., 2 F.R.D. 93, 93 (W.D. Wash. 1941) (ordering documents produced “at [parties’] respective places of business during reasonable office hours”); Compagnie Continentale D’Importation v. Pacific Argentine Brazil Line, 1 F.R.D. 388, 389 (S.D.N.Y. 1940) (“Document inspection should be held at defendant’s convenience... at its place of business.”); Gielow v. Warner Bros. Pictures, 26 F. Supp. 425, 426 (S.D.N.Y. 1938) (ordering party “to exhibit the documents... but at the office of the complainant; not to be removed therefrom”).


41 See, e.g., Dunn v. Midwestern Indemn., 88 F.R.D. 191, 193 (S.D. Ohio 1980) (granting plaintiffs, who claimed widespread “pattern or practice” of discrimination, a search that could take thousands of hours, of defendant’s “computer systems, including access to and information about... equipment, raw data, programs, data management systems, and the by-products of their analysis”).
government entities long have computerized their data. But courts often ignored cost-based objections to computer discovery, blaming the computer-using parties for “a system of record-keeping which conceals rather than discloses relevant records, or makes it unduly difficult to identify or locate them”; by blaming parties who use computers, courts justified un receptiveness “to defendants’ impossibility contentions insofar as they are grounded in the peculiar manner in which defendants maintain their computer systems."

When computer discovery was a rarity, courts could ignore its cost. Once it was common, courts faced the reality of tactical demands for costly searches. As in the 1970s, technology raised discovery cost and volume, inspiring efforts to limit discovery. The Manual for Complex Litigation in 1995 “advocated a cost-benefit approach” to stop “fishing expeditions,” proposing “condition[ing] particular discovery on payment of its costs by the party seeking it.”

Yet not all courts have followed the Manual for Complex Litigation, and one later case quickly became the leading word on e-discovery. In 2003, Zubulake v. UBS Warburg set out a multi-factor test, essentially a detailed cost-benefit analysis comparing the costs and benefits of discovering the disputed evidence, with cost evaluated in light of the parties’ resources, the amount in controversy, and the “relative ability of each party to control costs.” Zubulake never fully harmonized e-discovery law and practice, however. Not qualifying as “final” orders, discovery rulings ordinarily are non-appealable, so “few trial

---

42 United States v. Greenlee, 380 F. Supp. 652, 658 (E.D. Pa. 1974) (denying criminal defendant a requested weeks-long search of Internal Revenue Service computers that would have created risks of security breaches, privacy violations, and “serious interruption of the operations of the IRS”).


44 Pulver, supra note 12, at 1386.


46 See, e.g., In re Brand Name Prescription Drugs Antitrust Litig., 94 C 897, 1995 U.S. Dist. LEXIS 8281, *2-3 (N.D. Ill. June 13, 1995) (granting class action plaintiffs’ motion to compel defendant to produce computer-stored electronic mail at defendant’s own expense, estimated at $50,000-$70,000, and expressly rejecting alternative suggested by the Manual of Complex Litigation).


48 Factors focusing on the benefit of discovering the benefit include the following: “The extent to which the request is specifically tailored to discover relevant information”; “The availability of such information from other sources”; “The importance of the issues at stake in the litigation”; and “The relative benefits to the parties of obtaining the information.” Zubulake I, 217 F.R.D. at 322.

49 Zubulake I, 217 F.R.D. at 322.

50 28 U.S.C.A. § 1291 (authorizing appeal only of “final decisions” in cases); see Church of
court decisions regarding the scope and logistics of discovery . . . [reach] the appellate level.”[^51] Because *Zubulake* and virtually all e-discovery opinions are nonbinding district court precedent, *Zubulake*, though “widely regarded as the leading case authority” on e-discovery,[^52] has not drawn universal adherence.[^53]

Also helpful are the e-discovery reform proposals of a non-governmental group, *The Sedona Conference Working Group on Best Practices for Electronic Document Production.*[^54] Sedona distinguished data types as presumptively discoverable or not, such as “active data” (ordinary files) and backup/deleted “legacy data” accessible only at high cost.[^55] It also proposed requiring certain best practices for e-discovery, such as that requests “make as clear as possible what electronic documents and data” are sought.[^56] *Zubulake* cited Sedona’s early work,[^57] and Sedona’s updated report notes that courts have cited its earlier work,


[^53]: See, e.g., Multitechnology Servs. v. Verizon Southwest, Civ. 02-CV-702-Y, 2004 WL 1553480, at *1 (N.D. Tex. Jul 12, 2004) (ordering cost-shifting to requesting party even though *Zubulake* “weighs against shifting any expense” because “*Zubulake* is a district court opinion without binding authority”); Panel Discussion, supra note 18, at 24 (comments by the Hon. James C. Francis IV) (discussing use of multi-factor tests in e-discovery disputes: “[I]t depends on whether you adopt Judge Scheindlin’s view [in *Zubulake*] of a hierarchy or whether you think . . . [the] factors will probably play out differently in different cases. I am resistant to the hierarchy approach because my fear is that the factor at the top of the hierarchy will almost always wash out the other[s].”).


[^55]: “Active data is typically stored on local hard drives, networked servers, and distributed devices or offline archival sources from which information can be accessed without a special restoration effort.” Id. at 6-7. “Information stored solely for disaster recovery purposes, ‘legacy’ data retained in obsolete systems, and deleted or fragmentary information that can be restored only through extraordinary efforts” ordinarily are “unduly burdensome and costly to access.” Id. at 5.

[^56]: Id. at 5.

[^57]: E.g., *Zubulake I*, 217 F.R.D. at 320 n.61, 321 n.67;}
“helping provide de facto ‘national standards.’” Yet courts have not adopted
the Sedona principles wholesale; one topic they addressed heavily, preservation
of data, was “relegated . . . to evolving case law” by the e-discovery amendment
authors’ “refusal to enact preservation standards.”

Moreover, Zubulake and Sedona, while thoughtful and useful guides to e-
discovery, promise limited impact. They do not aim for a paradigm shift, instead
relying on status quo methods to limit discovery. Like Zubulake, Sedona
accepted as its touchstone the old “‘proportionality’ standard” of assessing costs
“in light of the nature of the litigation and the amount in controversy.” Sedona
rejected other ideas, like Texas’s mandatory cost-shifting (making requesting
parties pay for e-discovery not “reasonably available” in the “ordinary course” of
business) in favor of a mere permissive suggestion that cost “‘may’ (instead of
‘should’)” be shifted. Zubulake likewise said courts “should only consider
cost-shifting” for “relatively inaccessible [data], such as in backup tapes” costly
to recover, and “close calls should be resolved in favor of the presumption”
against cost-shifting, with cost-shifting denied if evidence is sufficiently useful.

The 2006 Federal Rules amendments followed at least the spirit of
Zubulake and Sedona in stressing cost-benefit proportionality limits: when
digital or electronically stored information is “not reasonably accessible because
of undue burden or cost, . . . the court may nonetheless order discovery . . . if the

58 Sedona 2007 Update, supra note 54, at 2 (citations omitted).
59 Id. at 2.
60 Id. at 5.
61 Tex. R. Civ. P. 196.4.
63 Zubulake I, 217 F.R.D. at 324.
64 Id. at 322-323 (noting factors such as the “extent to which the request is specifically tailored . . .
[and] availability of such information from other sources”); see also AAB Joint Venture v. United
States, 75 Fed.Cl. 432, 444 (Fed. Cl. 2007) (“Defendant shall bear the costs of restoration of the
initial sample of back-up tapes and screening the sample to identify responsive documents. The
parties will then have an opportunity to argue . . . [whether] additional restoration of back-up tapes
is likely to lead to production of relevant evidence and consequently who should bear the cost.”).
65 The main e-discovery rules are Fed. R. Civ. P. 26(a)(1)(B) (requiring initial disclosures to give
locations of documents and “electronically stored information” (“ESI”)), 34(a) (deeming ESI part
of “document” demands and allowing ESI testing or sampling), 34(b) (allowing parties to “specify
the form . . . in which ESI is to be produced,” with default rule that ESI be produced as “ordinarily
maintained” or in form “reasonably usable”), and 45 (allowing ESI discovery from non-parties).
66 See Panel Discussion, supra note 18, at 10 (federal judge Lee H. Rosenthal explaining that it is
“clear that the key is proportionality” in the new e-discovery rule, because its “good cause
determination must be based on the proportionality limits that [already] have been in the rules”).
requesting party shows good cause, considering the limitations of Rule 26(b)(2)(C).\(^{67}\) the proportionality rule. The Advisory Committee Notes to these new rules built upon Zubulake and other case law by prescribing an essentially similar cost-benefit analysis instructing courts to look to various factors relevant to the likely benefit and cost of a disputed discovery request.\(^{68}\)

**B. THE E-DISCOVERY RULES: JUST MORE DISAPPOINTINGLY CONVENTIONAL AND INEFFECTIVE PROPORTIONALITY LIMITS**

With the 2006 e-discovery rules following earlier proposals by the Manual for Complex Litigation, Zubulake and other district courts, the Sedona Group, and academic writings,\(^{69}\) the rules evolved in a highly decentralized fashion. Decentralization, though chaotic, often has the virtue of yielding a flourish of creative, varied independent efforts. Oddly, the major proposals to limit tech-inspired discovery proliferation, from the 1970s to date, have been similar, even derivative, cost/benefit “proportionality” rules, with perhaps a mild, permissive suggestion of rare cost-shifting for the most costly discovery.\(^{70}\)

But the main problem with proportionality limits on discovery is not that they are old news. It is that such limits never have worked terribly well, and appear unlikely to work well as to e-discovery. Although proportionality limits have gained momentum for decades, led by powerful forces in the judiciary and the bar, they are widely criticized as ineffective. “Whatever the theoretical possibilities,” the proportionality rule “created only a ripple in the caselaw,” the leading civil procedure treatise notes; “no radical shift has occurred.”\(^{71}\)

Even if the proportionality rule did not yield any major shift, courts do deny


\(^{68}\) See Advisory Cmte. Note to Rule 26(b)(2) (2006) (“[S]pecificity of the discovery request[]; [I]nformation available from . . . more easily accessed sources[]; The failure to produce relevant information that seems likely to have existed but is no longer available from more easily accessed sources[]; [L]ikelihood of finding relevant, responsive information that cannot be obtained from . . . more easily accessed sources[]; Predictions as to the importance and usefulness of the further information[]; The importance of the issues . . . in the litigation[]; and] parties’ resources.”).

\(^{69}\) See, e.g., Martin H. Redish, Electronic Discovery And The Litigation Matrix, 51 DUKE L.J. 561 (2001) (proposing more cost-shifting to parties requesting costly e-discovery); Pulver, supra note 12 (same); Horning, supra note 23 (proposing that courts more often require production of digital rather than paper data, as well as that courts more often allow requesting parties to require that producing parties put that data into specific forms and facilitate interpretation of complex data).

\(^{70}\) See supra notes 38-40 and accompanying text.

discovery for cost reasons. Certainly, some plaintiffs do get broad discovery, such as in *Kozlowski v. Sears, Roebuck*,\(^72\) in which the court allowed plaintiff discovery of all prior complaints of similar defective products, despite defendants’ claim that the discovery would require a burdensome search of all its product records.\(^73\) The court explained that “most courts have held that the existence and nature of other complaints in product liability cases is a proper subject for pretrial discovery,” and it would have been unfair “[t]o allow a defendant whose business generates massive records to frustrate discovery by creating an inadequate filing system, and then claiming undue burden.”\(^74\)

But however persuasive *Kozlowski’s* logic, its permissiveness with costly discovery is the exception, not the rule. Courts often disallow discovery on relevant matters for reasons of cost, burden, and inconvenience, such as in decisions denying plaintiffs’ requests for discovery on similar instances of misconduct in claims of discrimination\(^75\) or other wrongdoing,\(^76\) as well as in

\(^73\) *Id.* at 76.
\(^74\) *Id.*
\(^75\) See, e.g., Sallis v. Univ. of Minnesota, 408 F.3d 470, 477-78 (8th Cir. 2005) (upholding, in claims of racially discriminatory failure to promote and hostile work environment, denial of discovery of discrimination complaints at all university departments, despite plaintiff’s argument that complaints “were contained in an easily accessible, central database, and he experienced discrimination at the hands of other [university] departments”; deeming request “unduly burdensome, particularly . . . [because] the discovery deadline is rapidly approaching[:] . . . discovery must be limited, in both its temporal and geographical reach, so as to ameliorate the burdensomeness,” to just complaints in plaintiff’s department for one year); Balderston v. Fairbanks Morse Engine Div., 328 F.3d 309, 319-20 (7th Cir. 2003) (upholding, in “pattern or practice” claim of “systematic elimination of older employees” in reductions in force and companywide reorganization, denial of discovery of statistical data, including companywide personnel records of all employees terminated over nine years, because of courts’ “substantial discretion to curtail the expense and intrusiveness of discovery in limiting . . . broad discovery of personnel files”); Equal Emp. Opportunity Comm’n v. D.C. Public Sch., 217 F.R.D. 12, 13, 15 (D.D.C. 2003) (denying, in claim of age discrimination in termination of teacher in reduction in force, request for “teaching disciplines of each teacher . . . [in the] academic year” when plaintiff was terminated, where data was “perhaps retrievable only from a search of every personnel file,” which “would be oppressive”); Lee v. Executive Airlines, 31 F. Supp. 2d 1355, 1356 (S.D. Fla. 1998) (denying race discrimination plaintiff’s requests as to all employees “disciplined but not terminated for . . . time card infractions” (plaintiff’s alleged offense) over five years, where discovery “would require extensive searches of files outside of the locations” plaintiff worked); Aramburu v. Boeing Co., 885 F. Supp. 1434, 1444 (D. Kan. 1995) (denying race/disability discrimination plaintiff discovery that would take 240 hours to procure 1,500 personnel files, because “plaintiff’s need for the information” was “disproportionate” to the burden).

\(^76\) See, e.g., Kowalski v. Stewart, 220 F.R.D. 599, 602(D. Ariz. 2004) (denying motion to compel discovery, in prisoner claim of denial of court-ordered medical care, due to burden of “photocopying, organizing, and taking adequate measures to ensure prisoner confidentiality for the
decisions denying defendants’ efforts to discover information a government agency used to enact a disputed regulation. As to e-discovery in particular, case law under the new rules has been limited and mixed, but given the level of angst about costly e-discovery underlying the new rules, those rules seem likely to disappoint their advocates. Whatever the mix of decisions allowing and disallowing costly e-discovery, however, a more fundamental problem remains

77 See, e.g., United States v. Duke Energy Corp., 214 F.R.D. 392, 393 (M.D.N.C. 2003) (denying defendant, an energy company challenging the validity of a clean air regulation the Environmental Protection Agency (“EPA”) enforced against it, discovery from another federal agency, the Department of Energy, whose “personnel may have been present when some decisions were made by the EPA to determine its present interpretation of the Clean Air Act,” and “statements and positions taken by an EPA employee are relevant to the [statutory] interpretation underlying the regulation, but the “burden to the [defendants] far outweighs the relevance”); Wyoming v. Dep’t of Agric., 208 F.R.D. 449, 454 (D.D.C. 2002) (denying plaintiff, in suit against federal agency for violating rules on issuing regulations, an order for non-party witnesses to produce documents regarding those regulations, where court saw request as expensive and unduly burdensome).

78 Some cases allow costly e-discovery where justified by high case stakes. See, e.g., PSEG Power NY v. Alberici Constr., No. 1:05-CV-657, 2007 WL 2687670, at *1, *10 (N.D.N.Y. Sept. 7, 2007) (in $4.4 million construction contract claim, ordering plaintiff, at cost of $40,000-$200,000, “to produce all electronically stored emails, numbering approximately 3,000, conjunctively with their corresponding attachments as ‘married’ documents”). Other cases allow costly e-discovery despite modest case stakes when the information appears valuable. For example, in W.E. Aubuchon Co. v. BeneFirst, LLC, 245 F.R.D. 38 (D. Mass. 2007), a claim that an employee benefit administrator breached its fiduciary duty, the court found that the data sought – thousands of employee claims stored electronically as un-indexed images – was “not readily accessible” where it could cost $80,000 and 4,000 hours to recover 34,000 requested claim forms and medical bills, though plaintiff then narrowed its request to 3,000 claims. Id. at 44. Yet even though the discovery was burdensome and the “importance of the issues at stake” was low, the court allowed the discovery because the information was “clearly an integral part of the litigation . . . not only to BeneFirst’s culpability, but also to the amount of damages.” Id.

79 Courts also have denied seemingly high-relevance e-discovery, see, e.g., Wiginton v. CB Richard Ellis, Inc., 229 F.R.D. 568, 577 (N.D. Ill. 2004) (requiring class action harassment plaintiffs to pay 75% of $249,000 e-mail search for known pornographic and other harassing emails), and have denied costly discovery even in high-stakes litigation where the request seemed insufficiently essential, see, e.g., Best Buy Stores v. Developers Diversified Realty Corp., No. 05-2310, 2007 WL 4230806, *2-4 (D. Minn. Nov. 29, 2007) (in claim that landlords breached contracts and fiduciary duties (causing actual damages of $800,000, “enhanced damages” for fraud, and “long-term economic impact” on parties’ relationship), denying defendants’ request for plaintiff’s database on its other landlords’ lease charges, where data was not in searchable format and required restoration from original sources, costing $124,000 plus $27,823 per month, and defendant failed to
with both the original proportionality rule and the e-discovery rules.

Although denying relevant discovery due to cost may be defensible pragmatically, it is an unsatisfying concession of limited litigation accuracy due to the cost of finding and analyzing evidence needed for accurate verdicts or settlements. Less accuracy is troubling not only morally but economically. Failing to impose liability on the guilty, because evidence of guilt is too costly, yields insufficient deterrence of misconduct and insufficient assurance that parties internalize costs (e.g., of pollution) they impose on others. Conversely, imposing liability on the innocent, because exculpatory evidence was too costly, yields ill-targeted deterrence of innocent activity; imposing pollution liability on a non-polluting business just disincentivizes that socially useful commerce. 80

Thus, proportionality rules can be criticized equally for allowing entirely opposite errors, both “false negatives” (failing to detect and halt discovery abuse) and “false positives” (finding “disproportionate” some costly discovery that actually is justified by high evidentiary value and case merit). Erroneous pro-plaintiff rulings unjustifiedly increase litigation costs and pressure defendants to settle unmeritorious cases; conversely, erroneous pro-defendant rulings deny plaintiffs the ability to press meritorious claims successfully. 81

If the new e-discovery rules look to yield the sort of uninspiring results we saw after the original proportionality rule that so much of the legal establishment demanded, the question is: how can it be that powerful forces attempting to respond to a hugely costly phenomenon have proven so impotent for so long?

III.
AN ECONOMIC ANALYSIS AN ECONOMIC ANALYSIS OF WHY PROPORTIONALITY LIMITS, THOUGH POPULAR, CANNOT BE OPTIMAL

A. THE CONSENSUS: LIMIT DISCOVERY BASED ON COST/BENEFIT PROPORTIONALITY PRINCIPLES

For a field featuring so much controversy, discovery has featured an odd

---


81 See Bone and Evans, supra note 29, at 1287, so noting as to rulings on class action certification:

Judges make mistakes. They grant certification when it should be denied, and they deny certification when it should be granted. . . . A erroneous grant creates unnecessary administrative and litigation costs and . . . unjustified settlements. An erroneous denial adds to plaintiffs’ litigation costs and can make it harder for plaintiffs to recover.
degree of consensus among analysts in disparate fields. Free-market, litigation-skeptical “Chicago school” economists like Richard Epstein blame “underregulated” discovery for the “inexorable expansion of [tort] liability.”\textsuperscript{82} Frank Easterbrook, fretting about “impositional (excessive, abusive)” discovery that induces settlement by imposing high costs on defendants,\textsuperscript{83} advocates “limit[ing] discovery to matters admissible at trial”\textsuperscript{84} – a drastic change from allowing any discovery “relevant to the claim or defense of any party,” including material that is inadmissible but “reasonably calculated to lead to . . . admissible evidence.”\textsuperscript{85} Surprisingly, many non-Chicago-school economists have similar qualms about discovery. Robert Cooter and Daniel Rubinfeld propose that after a “reasonable” amount of discovery in a case, the cost of responding to discovery requests shift to the requesting party.\textsuperscript{86} Though disagreeing with that proposal, Bruce Hay notes how heavy discovery, by scaring defendants into settling early, can counter-productively lead to less, not more, disclosure of illegality.\textsuperscript{87}

There of course are dissenting voices criticizing the drive to narrow discovery,\textsuperscript{88} and there have been empirical findings that discovery excess may be confined to exceptional cases.\textsuperscript{89} But those voices have been on the periphery of the consensus in favor of more proportionality-based limits on discovery – a consensus spanning (as discussed above) civil proceduralists, economists, and the judges behind the past few decades of rules imposing new discovery limits.

Proportionality limits can be optimal, though, only if courts can perform the needed cost-benefit analyses passably well. They cannot, as the next subpart


\textsuperscript{84} Easterbrook, \textit{supra} note 83, at 644.

\textsuperscript{85} Fed. R. Civ. P. 26(b)(1).


\textsuperscript{89} See Willging, \textit{supra} note 4 (noting that discovery cost is unusually high in costliest 5% of cases).
discusses, which means that discovery limits are doomed to be sub-optimal.

B. AN ECONOMIC VIEW OF DISCOVERY: HELPING FACTFINDERS ASSESS CASE VALUE AND MERIT \((L \text{ and } p)\) – WHICH MAKES DISCOVERY DECISIONMAKING CIRCULAR

The purpose of broad discovery, in economic terms, is well established: “a full exchange of the information . . . enabl[es] each party to form a more accurate, and generally therefore a more convergent, estimate of the likely [case] outcome.”\(^90\) The case law uses similar logic to justify broad discovery.\(^91\) This Article focuses not only on how discovery helps parties assess cases, but also on a matter more rarely analyzed in economic terms: how judges decide discovery disputes. Such judicial decisionmaking draws little academic attention, probably because almost all discovery decisions are unappealable district court decisions\(^92\) and because only experienced litigators recognize that if cases are won and lost in discovery, they really are won and lost \textit{in litigating discovery motions} – dueling motions to compel discovery\(^93\) and motions for protective orders.\(^94\)

A court must undertake a cost-benefit analysis to decide, as the rules require, whether the value of particular discovery is proportional to its cost (both dollar cost and nonpecuniary burdens). It must assess the cost of finding and producing the evidence, and it must compare that cost to the benefit of having that evidence. Assessing cost often is feasible; parties litigating discovery regularly detail the time and dollar costs of producing disputed evidence.\(^95\)

Assessing the \textit{benefit} of particular discovery is the tricky part. As discussed below, for truly accurate judicial decision-making, a court must consider not only the probative value of the particular evidence and the size of the case (as the rules command), but also – contrary to the conventional wisdom on discovery decisionmaking – the likelihood that the case is meritorious, \textit{i.e.}, the odds the plaintiff will prevail at trial. In economic terms, the court’s proportionality determination bases on the following three variables:


\(^{91}\) “Mutual knowledge of all the relevant facts gathered by both parties is essential to proper litigation.” Hickman v. Taylor, 329 U.S. 495, 507 (1947).

\(^{92}\) See supra note 50.

\(^{93}\) See Fed. R. Civ. P. 37(a).

\(^{94}\) See Fed. R. Civ. P. 26(c).

\(^{95}\) See supra notes 75-79 (collecting cases in which parties detailed, and argued before the court about, the costs and burdens of responding to discovery requests).
L, the size of the case, typically the amount in controversy but also possibly the value of nonmonetary relief;

p, the probability that the case is meritorious; and

Δp, the difference the disputed evidence makes to p, the estimated likelihood the case is meritorious (the probative value of the evidence).

Yet, as discussed below, each of these three variables can be difficult or impossible for courts to assess during discovery – a diagnosis of pessimism about courts’ ability to make accurate proportionality decisions on discovery disputes.

1. SIZE OF CASE (L): DIFFICULT TO DETERMINE IN MANY CASES

Under the proportionality rule, in assessing “the likely benefit” of discovery, courts should “take[e] into account . . . the amount in controversy . . . [and] importance of the issues at stake.”96 This makes economic sense; discovery offers less benefit in low-value cases. More evidence-gathering expense is justified in a case that might result in millions of dollars changing hands; it is harder to justify similar evidence-gathering expense in a small-claims dispute.

“Amount in controversy” (dollar value) and “importance of the issues” (nonmonetary value), however, can be uncertain until trial; plaintiffs often couple strong claims for modest damages with weak claims for high damages.97 Worse, case value can be subjective; what is the value of an injunction stopping seal-clubbing, sex harassment, or other illegality?98 Certainly, courts do assign damages sums for complex, nonpecuniary harms like torture.99 Still, it remains wholly subjective whether evidence that might help win an injunction stopping seal-clubbing is “proportional” to a month-long, million-dollar data search.

With case value often subjective, one problem with proportionality is “finding principled criteria for differentiating between various types of cases”; “[w]here . . . are judges expected to find the criteria and analytical structure for making such judgments” as whether more discovery is warranted on a high-dollar contract claim than on a low-dollar discrimination claim for mainly

---

97 See Davis v. Ross, 107 F.R.D. 326, 327, 330(S.D.N.Y. 1985) (“[P]lausible claims for punitive damages can easily be made in many actions . . . . [T]he amount of damages will always be in issue; plaintiff seeks one million dollars in compensatory damages, and evidence must be introduced to demonstrate that the award should be more than nominal.”).
98 See, e.g., Lyons v. Mobil Oil, 554 F. Supp. 199, 201 (D. Conn. 1982) (holding that prevailing party wins attorney fees unless it obtains only nominal damages because “[i]njunctive relief is an important part of the [statutory] scheme . . . regardless of whether . . . damages are awarded”).
99 Hilao v. Marcos, 103 F.3d 767 (9th Cir. 1996) ( awarding damages in human rights class action).
These problems can be mitigated with “rules-of-thumb” as to “the amount of discovery normally permissible in certain types of cases,” e.g.:

[A] search for discriminatory intent in a civil rights case may be seen as involving constitutional values . . . [and] broader discovery . . . than in a personal injury or commercial case. On the other hand, judicial experience indicating that in certain civil rights cases, . . . further discovery is unlikely to shed additional light . . . , might lead a judge to place limits . . . .

Although “patterns of appropriate discovery . . . may emerge which can normally be followed unless the particular facts warrant otherwise, . . . [o]bviously these judgments will not be easy.” Nor will they be optimal: such rules of thumb would treat widely varied cases of the same subject matter similarly.

This difficulty estimating $L$ is not the main topic of this Article’s analysis, but the partial solution, assuming the same average value for all cases in the same “pool,” will return in the next subpart as a similar imperfect solution to the problem of estimating $p$, the likelihood that the plaintiff will prevail.

2. **Probative Value of the Evidence Sought ($\Delta p$): Difficult to Assess Before Fully Analyzing That and Other Evidence**

The most important consideration in a discovery dispute is the probative value of the evidence – $\Delta p$, the difference ($\Delta$) the evidence makes to the estimated probability ($p$) the case is meritorious. The proportionality rule asks courts to assess the “likely benefit” of discovery, taking into account “the needs of the case . . . and the importance of the proposed discovery in resolving the issues,” which are aspects of probative value. “[C]omplexity” of the issues, another proportionality factor, also goes to probative value: in a simple case (e.g., “did she sign the contract?”), more evidence has little probative value; more evidence is most useful in cases about technical matters, hidden intent, etc.

---


101 Id. at 279.

102 Id.


104 See infra note 115.

105 The rules and cases on expert witness admissibility expressly rely on this logic. See Andrews v. Metro N. Commuter R.R. Co., 882 F.2d 705, 708 (2d Cir. 1989) (“For an expert’s testimony to be admissible . . . , it must be directed to . . . scientific, technical, or specialized knowledge and not to lay matters which a jury is capable of understanding and deciding without the expert’s help.”).

106 Employment discrimination cases are the paradigmatic example. See supra note 1.
Yet it may be difficult for courts to discern the probative value of evidence before discovery of that evidence. Nobody knows in advance what a witness will say in a deposition, making it difficult to assess the probative value of going beyond the ten-deposition limit.\textsuperscript{107} The same holds for searching computers or voluminous paper files: the party opposing discovery will have to make its “proportionality” cost-benefit argument \textit{before} the search, so the court will not see the fruits of the discovery before having to rule on its “likely benefit.”\textsuperscript{108}

“In the absence of any information about [the] evidence,” Richard Posner noted in discussing how parties anticipate opposing evidence, the only option is to “assume that such evidence . . . is of average helpfulness.”\textsuperscript{109} For example, when a court must decide whether a data search for similar stock trades is worth the cost, all it knows is whether these kinds of searches, in these kinds of cases, are usually fruitful for plaintiffs (or defendants).\textsuperscript{110} As when assessing the size ($L$) of cases with subjective value, courts assessing probative value have little other than the Sherman and Kinnard idea that “patterns of appropriate discovery in certain cases may emerge” based on “rules-of-thumb for determining the amount of discovery normally permissible in certain [case] types.”\textsuperscript{111} In some “civil rights cases,” for example, “further discovery is unlikely to shed additional light” on the issues\textsuperscript{112} – or perhaps such cases’ complex intent questions require more extensive evidence-gathering.\textsuperscript{113} The tension between these views of civil rights cases shows how indeterminate and imperfect such rules of thumb can be.

Thus, even the most obvious proportionality factor – probative value – can be difficult for courts to analyze, especially if they cannot see the evidence (e.g., what will happen at a deposition) before ruling. As discussed below, it is just as hard for courts to analyze $p$, the probability that a case has merit – and, disturbingly, most courts do not even see merit as relevant to discovery decisions.

\textsuperscript{107} Fed. R. Civ. P. 30(a)(2)(A)(i) (requiring “leave of court” for more than ten depositions), 30(d)(1) (providing that “a deposition is limited to 1 day of 7 hours” absent leave of court).

\textsuperscript{108} See Bone, Economics of Civil Procedure, \textit{supra} note 90, at 229 n.36 (noting that to expand discovery past presumptive limits, courts must assess the value of greater discovery, which “is bound to be difficult in the absence of precise knowledge of what the discovery will reveal”).

\textsuperscript{109} Posner, Economic Analysis of Law, \textit{supra} note 90, at 571.

\textsuperscript{110} The plaintiff, the party with the burden of proving its opponent’s misconduct, usually is the one seeking more discovery. Defendants may seek extensive discovery to prove misdeeds by plaintiffs, \textit{see, e.g., supra note 77} (citing cases); such a defendant is in a position akin to a plaintiff, seeking evidence to prove its opponent’s misdeeds, which is why this Article takes as its paradigmatic examples plaintiffs seeking evidence to prove allegations of misconduct that defendants deny.

\textsuperscript{111} Sherman & Kinnard, \textit{supra} note 100, at 279.

\textsuperscript{112} Id.

\textsuperscript{113} \textit{See supra} note 1 (discussing Hollander v. American Cyanamid Co.).
3. **Probability the Claim Has Merit (p): Difficult for Court to Assess Before Seeing All the Evidence and the Parties’ Arguments about That Evidence**

a. **Why Courts Do Not Consider Case Merit in Making Discovery Decisions: The Conventional Wisdom that Discovery Is Unrelated to Case Merit**

It is received wisdom that proportionality rulings do not depend on the probability the case is meritorious (p), as distinct from case size (amount in controversy and importance of issues, L), or issue complexity (which goes to probative value, Ap).\(^\text{114}\) The rule and its Advisory Committee note detail various factors in the “likely benefit” of requested discovery, none of which relates to the probability the case has merit; all the factors relate to case size, probative value, the parties’ resources, and whether the evidence is available elsewhere.\(^\text{115}\)

Courts rarely say *anything* about case merits in deciding discovery disputes; when they do, it almost always is *disclaiming* any consideration of case merits in deciding discovery disputes, as in this classic passage from a case decided soon after the 1938 enactment of the Federal Rules of Civil Procedure:

[D]efendant is really arguing . . . that the issue raised by the plaintiff is irrelevant, not that the interrogatories are irrelevant to the issue. . . . Whether or not the plaintiff is right is immaterial at this stage. . . . [T]o ask the Court to decide the whole case on answers to interrogatories involves a misconception of the office of discovery procedure.\(^\text{116}\)

\(^\text{114}\) *See supra* Parts III(B)(1) (as to \(L\)) and III(B)(2) (as to \(Ap\)).

\(^\text{115}\) Fed. R. Civ. P. 26(b)(2)(C)(iii) (assessing “likely benefit” by “the needs of the case, the amount in controversy, the parties’ resources, the importance of the issues at stake in the litigation, and the importance of the proposed discovery in resolving the issues”); Advisory Cmte. Note to R. 26 (listing similar factors and noting that “cases in public policy spheres, such as employment practices, free speech, and other matters, may have importance far beyond the monetary amount”).

There are rare exceptions to the rule that courts do not consider case merit in discovery decisions, but typically when the disputed discovery implicates such a weighty public interest that it must be disallowed absent a sufficient showing of merit. In this light, the exceptions show not that courts do consider the merits on discovery motions, but instead that in rare situations, a public interest creates a sort of privilege against disclosure that only sufficient case merit can overcome. Sporadic exceptions—that-prove-the-rule aside, the rule against considering case merit on discovery motions is quite well-established and consistently followed.

**b. Why Courts Should Consider Case Merit: Optimal Discovery Depends on Whether a Case is a “Close Call”**

Whether or not the conventional wisdom is an accurate statement of how courts actually decide discovery disputes, it is dead wrong as to how they should decide them. Accurate cost-benefit analysis of the value of evidence is impossible without considering case merits, because the benefit of evidence (helping a plaintiff impossible without considering case merits, because the benefit of evidence is enough merit that the factfinder is permitted, but not compelled, to rule for the plaintiff. In the lowest-odds cases, additional evidence has little value because it is unlikely to affect the outcome, which is why parties can move to dismiss before discovery – to avoid discovery where, given the lack of merit, “[n]o

---

117 Discovery decisions considering case merits tend to feature a high public interest that discovery would jeopardize, e.g., a public interest in newsgathering harmed by discovery from journalists, see Apel v. Murphy, 70 F.R.D. 651, 654 (D.R.I. 1976), or in avoiding publicity that would discourage reports of air accidents, see United Air Lines, Inc. v. United States, 26 F.R.D. 213 (D. Del. 1960).

118 See supra note 110 (discussing how same analysis applies to defendants).
amount of discovery could change the legal reality [of] plaintiff’s claim.”\textsuperscript{119} Conversely, in the highest-odds cases, the proverbial “slam-dunks” (e.g., a clear breach of a contract to pay a certain sum), additional evidence will be of little use; the plaintiff often can win before trial, based on initial case filings.\textsuperscript{120}

Case merit (p) affects the optimal discovery amount because of the court’s opposing goals: (1) limit discovery cost (C_D); and (2) limit the error cost (C_E) of incorrect verdicts. More discovery raises discovery cost (C_D) while lowering error cost (C_E). The latter is a diminishing benefit; each additional piece of evidence likely is less useful (less helpful at preventing errors) than the prior one (e.g., deposition #1 is the key decisionmaker; #2, a key witness; #3, a peripheral witness).\textsuperscript{121} The court must choose the discovery amount (Q, quantity) that minimizes total cost, the sum of discovery and error costs (C=C_D+C_E). The court’s decision is illustrated by the following model, which economics-minded readers may recognize as based on the classic incomplete information Cournot duopoly model\textsuperscript{122}; however, the non-economics-minded reader “whose algebra is a bit rusty can skip . . . [these] algebraic expressions without any problem.”\textsuperscript{123}

\textsuperscript{119} See, e.g., Federico v. United States, 70 Fed. Cl. 378, 386 (2006) (dismissing pre-discovery for failure to state a claim because of “legal reality. . . that federal employees who serve by appointment may not bring contract claims”); Kloth v. Microsoft Corp., 444 F.3d 312, 324 (4th Cir. 2006) (same, on claim defendant “deprived consumers of competitive technology,” where claim of injury from deterred invention of new technology was “speculative and beyond the competence of a judicial proceeding,” so “discovery would not change or inform the nature of the alleged injuries”).\textsuperscript{120} See, e.g., New Rochelle Dodge, Inc. v. Bank of New York, 511 N.Y.S.2d 663, 665, 127 A.D.2d 638 (App. Div., 2d Dep’t, 1987) (granting plaintiff “summary judgment in lieu of complaint,” a procedure for plaintiffs to obtain pre-discovery judgment, where “defendant . . . acknowledged the debt” on retail installment contract); Chase Manhattan Bank, N.A. v. Markovitz, 392 N.Y.S.2d 435, 56 A.D.2d 763 (1st Dep’t 1977) (same, where defendant had agreed to pay legal expenses).\textsuperscript{121} See Richard A. Posner, An Economic Approach to the Law of Evidence, 51 STAN. L. REV. 1477, 1482 (1999) (“[As] more evidence is obtained, the effect of additional evidence . . . will tend to decrease, especially if the search[] begins . . . with the most probative evidence.”). A related reason that evidence offers diminishing marginal benefits is that “[i]f the searcher cannot determine in advance which evidence is . . . fruitful, his search procedure will resemble random sampling, and as the size of a sample grows, the value of additional sampling . . . [is] at a falling rate.” Id.\textsuperscript{122} In the Cournot model, two firms dominate a market, each facing an inverse demand curve and choosing production quantity by estimating the probability the other has high costs (low output) or low costs (high output). ROBERT GIBBONS, GAME THEORY FOR APPLIED ECONOMISTS 144 (1992). My model is analogous as to the probability a case has merit. With optimal discovery quantity rising and then falling as p increases from 0 to 1, I model p as a continuous, not discrete, variable (and thus use derivatives for optimization calculations). See Cooter & Ulen, Law & Economics 317 (4th ed. 2003) (“Economists often prefer to develop theory using continuous variables.”).\textsuperscript{123} BONE, ECONOMICS OF CIVIL PROCEDURE, supra note 90, at 89 (noting similarly).
Goal:  \( \min_{Q} C = C_D + C_E \)  

**Explanation:**

- Choose the amount of discovery (\( Q \)) that minimizes the sum of error costs and discovery costs.

Where:  \( C_D = aQ \)  

**Explanation:**

- Discovery cost (\( C_D \)) is proportional to the amount of discovery (\( Q \)), so it is modeled as \( Q \) times a constant (\( a \)) reflecting cost per additional unit of discovery.

& where:  \( C_E = \left( \frac{b}{Q} \right) (p-p_0^2) \)  

**Explanation:**

- When a case is clearly meritorious (\( p=1 \)) or meritless (\( p=0 \)), there is no error risk; the term \( p-p_0^2 \) models an error cost (\( C_E \)) that is zero when \( p \) is 0 or 1 and peaks when the case is as close a call as possible (\( p=0.5 \)).
- Error cost (\( C_E \)) decreases at a declining rate as the amount of discovery (\( Q \)) increases; this is modeled with the term \( \frac{b}{Q} \), where \( b \) is a constant reflecting the evidentiary benefit per additional discovery.

Substituting the above expressions for \( C_D \) and \( C_E \) into \( \min_{Q} C = C_D + C_E \) yields:

\[
\min_{Q} C = \min_{Q} aQ + \left( \frac{b}{Q} \right) (p-p_0^2)
\]

The optimization conditions: choose \( Q^* \), the optimal discovery amount, so as to minimize \( C \), i.e., any more discovery would raise discovery cost more than it reduces error cost, and any less would raise discovery cost less than it increases error cost:

\[
(1) \frac{\partial C}{\partial Q} = 0 \quad \text{and} \quad (2) \frac{\partial^2 C}{\partial Q^2} > 0
\]

Calculating \( Q \) for condition (1), i.e., the \( Q \) for which \( \frac{\partial C}{\partial Q} = 0 \):

\[
\frac{\partial C}{\partial Q} = a - \frac{b}{Q} (p-p_0^2) = 0
\]

\[
Q^* = \sqrt{\left( p - p_0^2 \right) \frac{b}{a}}
\]

**Explanation:** Optimal discovery (\( Q^* \)) depends on \( p \).

Calculating the term constituting condition (2):

\[
\frac{\partial^2 C}{\partial Q^2} = 2b(p-p_0^2)/Q^3 > 0
\]

**Explanation:** The condition is satisfied because all components of this term are positive; for all \( p \) between 0 and 1, \( (p-p_0^2) > 0 \), and \( b \) and \( Q \) are positive because error risk and discovery costs are positive.
In sum, the optimal discovery amount ($Q^*$) depends on $p$, the odds a claim will win. Optimal discovery is highest the closer the odds are to 50/50 (the closer $p$ is to 0.5), lowest when case merit is readily apparent (the closer $p$ is to 0 or 1).


For courts to make accurate decisions as to optimal discovery amount, they must consider $p$, the case merits – yet that may be the hardest task to undertake in a discovery dispute. This is a problem of decision-maker difficulty interpreting information signals (parties’ claims as to case merit), so it is useful to model the situation with game theory, the branch of economics that is “a very powerful tool for modeling information and studying its economic role.” In game theory terms, a decision (here, court discovery rulings) must base on some measure of merit (often in game theory the value of a good for sale, but here the merits of parties’ claims) that parties try to communicate. But during discovery, it is hard for courts to tell which cases truly have merit: all the evidence has not yet been gathered; and even if it had been, courts cannot review all of a case’s evidence (essentially holding a mini-trial) just to resolve a discovery dispute. As a result, in discovery, those of low merit often can falsely signal high merit.

With limited potential for effective communication of merit, the court’s decision must base on the average merit of all cases in the pool. Thus, a “pooling equilibrium” exists in which courts’ best available strategy is to rule the same regardless of case merit. Others have noted how parties discount each others’

---

124 To reiterate the preceding calculations: $Q^* = \sqrt{\frac{b}{a}(p - p^2)}$

125 See Eric Rasmusen, GAMES AND INFORMATION: An INTRODUCTION TO GAME THEORY § 11.1 at 320 (2006) (noting that “signaling costs must differ” between those of high and low worth “for signaling to be useful”).

126 See H. SCOTT BIERMAN & LUIS FERNANDEZ, GAME THEIRY WITH ECONOMIC APPLICATIONS § 18.2.6 at 337 (1993) (“a pooling equilibrium implies the informed player’s actions reveal nothing about what type of player he is”). The situation actually is likely a partially pooling equilibrium because some parties can signal merit effectively, such as with a powerful piece of evidence it unearthed early enough to submit to the court on a discovery motion. See Gibbons, supra note 122, at 213-18 (discussing partially pooling equilibria). The diagnosis of “pooling” remains, because in many cases, the evidence will be equivocal or disputed, and the court will have trouble sifting through both sides’ opposing arguments as to case merit, so many cases of varied merit levels will populate
bragging about the merits of their case or about a certain piece of evidence. This Article adds that not only parties, but also courts deciding discovery disputes, face the same dilemma of receiving useless signals of merit.

Worse, judges might have an exaggerated (rather than accurate) perception of the extent to which, in discovery, they must assume all cases are low-merit ones not warranting costly discovery. If the pool contains more low- than high-merit cases, it is rational for judges to presume, early in a case, that the case likely has low merit and so does not deserve costly discovery. Presumably judges change such opinions as more case information emerges. But as anyone who has had an argument about politics knows, even relevant information may not convince people to change their initial estimates due to the confirmation bias—people’s tendency to be “not equally open to all information, but more open to that which comfortably confirms their views, more inclined to spin disconfirming evidence to fit those views.” Even if judges are less prone to confirmation bias, being experts at analyzing evidence and justifying conclusions with logic, there is little reason to think them immune from this well-documented quirk in human cognition, especially given recent experimental evidence that judges do use error-prone intuitive shortcuts to make decisions. Due to confirmation bias, judges’ early-stage inability to distinguish good from bad cases may persist even after they receive enough information about a case to start separating it from the pool.

Although ruling the same on all cases in the pool is the best available judicial strategy, it is merely the best among imperfect strategies; it is an imperfect

---

128 Bone, Economics of Civil Procedure, supra note 90, at 205 (noting that because “parties have incentives to misrepresent that they have favorable evidence when they do not . . . [and] verification is not always possible, . . . [recipients] discount the truth of the information disclosed”).

129 P.C. Wason, On the Failure to Eliminate Hypotheses in a Conceptual Task, 12 Q. J. Experimental Psychol. 129, 138-39 (1960) (finding that after people make an initial, premature guess as to a numerical pattern, they skew their interpretation of later data to preserve that guess).


[In our experimental research on judges[,] [w]e provide tests of judges’ general reasoning skills as well as their decision-making skills in legal contexts. Our results demonstrate that judges, like others, commonly make judgments intuitively, rather than reflectively, both generally and in legal contexts.]

Id. at 105.
strategy that yields sub-optimal results. In low-merit cases, ruling identically on all cases yields too much discovery, i.e., more discovery than justified by the need for more evidence to assess case merit. Ruling identically on all cases yields too little discovery in close-call cases, i.e., it disallows the extensive discovery that is justified when the case is a close call for the factfinder. In this scenario, the bad cases are treated too well and the good treated too badly, as in George Akerloff’s classic economic analysis of used car markets: due to “asymmetry in available information . . . good cars and bad cars must still sell at the same price since it is impossible for a buyer to tell the difference”; as a result, “bad cars drive out the good because they sell at the same price as good cars.”  

As with product decisions, in litigation the bad may come to drive out the good. If courts allow most cases similar discovery – the amount appropriate for an average case – bad (weak) cases will take up too much time and money, while good cases lose for inability to gather enough evidence, or may never be filed because they will be allowed only average-case discovery. This is to say that bad cases will drive out good cases; court dockets and parties’ litigation efforts may be filled with more bad cases, and fewer good cases, than is optimal (i.e., good cases will get too little attention, bad cases too much). Consequently, the discovery problem this Article diagnoses – courts’ inability to separate good and bad cases until after the discovery that accounts for so much litigation cost – may be a cause of the widely noted prevalence of frivolous litigation.

Could courts adjudicating discovery disputes undertake the necessary inquiries into the merits? One problem is that any merits analysis will be incomplete because it would lack at least some of the evidence, given that the analysis would be occurring during discovery and before the resolution of all discovery disputes. A second problem is that even if all the evidence is in, the information costs of undertaking a merits analysis to decide a discovery dispute is prohibitive; the court and the parties would have to spend a great deal of time, and the parties a great deal of money, holding a mini-trial presenting and arguing about all the evidence and inferences therefore. Thus, a merits analysis is necessary, but infeasible, for optimally accurate rulings on discovery disputes.

Under this analysis, the quest for better discovery limits has disappointed not due to bad decision-making or bad rule-making, as many argue. Typifying

---


133 See supra notes 81-88 and accompanying text (citing views on litigation and discovery excess).

134 See Henry S. Noyes, Good Cause is Bad Medicine for the New E-Discovery Rules, 21 HARV. J. L. & TECH. 49 (2007) (criticizing proportionality and e-discovery rules as too vague to rein in excess discovery that courts are too unwilling to limit); Redish, supra note 69, at 563-64 (2001)
arguments blaming rule-making, Thomas Rowe criticizes as too “vague” to “curb[] cost and excess” the 2000 narrowing of discovery from material relevant to the “subject matter” to material relevant to “claims and defenses”; one judge less charitably depicted “debating [that] difference . . . [as] the juridical equivalent to debating the number of angels that can dance on the head of a pin.”

Henry Noyes faults bad rule-making and bad judging:

[The] e-discovery amendments are the fourth recent attempt to contain discovery. The three prior . . . relied on increased judicial discretion, mistakenly assuming that judges would act to limit discovery. . . . [H]owever, courts have continued to rely on the default policy of “liberal discovery.” . . . [C]ontinued and expanded use of the good cause standard is problematic both for the new e-discovery rules and for the existing discovery rules.

Noyes concludes that “[t]he courts’ persistent reliance on the ‘liberal rules of discovery’ mantra will only be overcome with express instruction to limit discovery, which is absent from the e-discovery amendments.”

This Article, while agreeing that courts’ discovery decision-making is sub-optimal, and perhaps not disagreeing that different rules could help, disagrees as to whether better rules or better judicial decision-making truly could fix the problem. Even with the best of possible rules and judging, courts and parties would remain stuck in a pooling equilibrium; judges simply do not have the necessary information to make optimal decisions about exactly what discovery to allow. It is a fundamental information timing problem inherent in the discovery stage of litigation: optimal discovery depends on the merits, but the merits are knowable only after discovery. As in the folk song about the hole in the bucket fixable only with a machine requiring water poured from that bucket, the problem is a classic circularity; the problem prevents the solution.

(noting that “the rules’ drafters and revisers over the years . . . have failed to fashion a discovery process that satisfies most people,” and specifically criticizing discovery rules for lacking more cost-shifting or spoliation provisions); Thomas D. Rowe, Jr., A Square Peg in a Round Hole? The 2000 Limitation on the Scope of Federal Civil Discovery, 69 TENN. L. REV. 13 (2001) (criticizing federal rules’ discovery limits as vague and therefore unable to change judicial decision-making).

Rowe, supra note 134, at 14.


Noyes, supra note 134, at 51-52.

Id. at 52.

See supra note 30 and accompanying text.
IV.
SOLVING THE POOLING THAT PREVENTS BETTER DISCOVERY DECISIONS:
IN CLOSE CALLS ABOUT COSTLY DISCOVERY, PRESERVE THE EVIDENCE
BUT DELAY THE DISCOVERY UNTIL AFTER SUMMARY JUDGMENT

A. REPLACING THE POOLING EQUILIBRIUM WITH A SEPARATING
EQUILIBRIUM AS A CASE PROGRESS

In a pooling equilibrium, decisions are sub-optimal because it is hard to
distinguish between the meritorious and the unmeritorious, as discussed above.
More optimal decisions are possible in a separating equilibrium in which parties
are forced to “reveal their types . . . [to] previously uninforme d” decision-
makers.140 Courts could make more accurate discovery decisions if they could
better tell case merit, allowing more discovery in “close call” cases that, being
neither clear winners nor clear losers, warrant more extensive evidence-
gathering. But courts cannot separate the close calls from the broader case pool
unless parties can credibly “signal” merit by citing and asserting the evidence
supporting their positions.141 During discovery, parties have not yet gathered and
marshaled all their evidence, so low, mid-, and high-merit cases are hard to
distinguish; due to parties’ inability to signal merit level convincingly, courts are
stuck with a pooling, rather than a separating, equilibrium.142 The only way out
is for courts to conduct mini-trials in which parties argue case merits, detailing
and offering interpretations of the evidence,143 but the information costs (in time

equilibria actually may be sub-optimal if the signals have no intrinsic value except as a signal of
merit (e.g., obtaining a certain educational degree as a signal of work ethic or intellect). In such a
separating equilibrium, the cost of signal acquisition (e.g., time and tuition) could exceed the
improved ability to separate those of high and low merit. See Michael Spence, Job Market
Signaling, 87 Q. J. ECON. 355, 364-65 (1973). But this Article addresses forced disclosure of
evidence a party wishes to conceal, so the problem of wasteful acquisition of signals is inapposite.
141 Michal Barzuza, Lemon Signaling in Cross-Listing, at 15 (Oct. 2007) (available at
http://papers.ssm.com/sol3/papers.cfm?abstract_id=1022282) (discussing investor efforts to distinguish “Type
L” companies more susceptible to corruption and “Type H” ones less susceptible: “A separating
equilibrium will result if and only if managers of Type L companies choose not to mimic the
managers of Type H.”); Lucian A. Bebchuk, Asymmetric Information and the Choice of Corporate
be destabilized if “better” actors can make tangibly different offers); Robert H. Frank, PASSIONS
WITHIN REASON THE STRATEGIC ROLE OF THE EMOTIONS 96-113 (discussing how signals can
degenerate into cheap talk if listeners are uninformed and therefore unable to spot false signals).
142 See POSNER, ECONOMIC ANALYSIS OF LAW, supra note 90, at 726 (noting how pooling equilibria
occur when those with higher merit find it “difficult to separate themselves” from those with less).
143 See supra Part III(B)(3)(c).
and money) of that endeavor are prohibitive for resolving a discovery dispute.

A pooling equilibrium may become a separating equilibrium over time, as more information emerges that illustrates distinctions among the pool — a point noted by game theory analyses of information problems outside the litigation context, such as analyses of information about product quality and corporate corruption. Pretrial litigation is, at heart, a series of stages at which different information emerges. The paper pleadings stage, disclosing parties’ allegations, is followed by pre-discovery motions for judgment on the pleadings (most commonly motions to dismiss, either for failure to state a claim or for jurisdictional failings) that disclose some of the parties’ legal arguments and weed out the cases whose (lack of) merit is clearest, which is followed by fact disclosures in discovery, which is followed by summary judgment motions that further weed out weak claims, and then the trial that resolves remaining claims.

In sum, as a case progresses through the pretrial stages, it gets easier to distinguish it from the pool. This is why, even though most cases settle, some do not settle until some motion litigation or discovery; the outcomes of certain pretrial skirmishes, or disclosures in early-stage discovery (e.g., the initial, key depositions), may allow parties to signal merit more meaningfully than they could earlier.

In this sense, moving from one litigation stage to the next –

---

144 The classic article is George A. Akerlof’s The Market for “Lemons”: Quality Uncertainty and the Market Mechanism, which notes that in used car markets, “bad cars drive out the good because they sell at the same price as good cars,” but over time better information emerges:

After owning a specific car . . . , the car owner can form a good idea of the quality . . . ;

i.e., the owner assigns a new probability . . . that his car is a lemon. This estimate is more accurate than the original estimate. . . . But good cars and bad cars must still sell at the same price since it is impossible for a buyer to tell the difference.


145 Barzuza, supra note 141, at 8-9 (available at http://papers.ssm.com/sol3/papers.cfm?abstract_id=102282) (discussing how a pooling equilibrium might become a separating equilibrium if law forces a decision on parties (i.e., whether to list stock on an exchange imposing intrusive regulation) that high- and low-value companies decide differently, thereby credibly signaling their value).

146 Or plaintiffs’ motions for judgment on the pleadings, which are rarer because “federal courts have followed a fairly restrictive standard in ruling on motions for judgment on the pleadings.” Charles Alan Wright et al., 5 Fed. Prac. & Proc. Civ. 3d § 1368 (2007) (collecting cases); see also id. at § 1367 (“[J]udgment on the pleadings only has utility when all material allegations of fact are admitted or not controverted in the pleadings and only questions of law remain . . . .”).

147 See Swierkiewicz, supra note 34, passim (discussing this sequence of pretrial stages).

148 See Bone, Economics of Civil Procedure, supra note 90, at 90; Scott A. Moss, Illuminating
pleadings, dismissal motions, discovery, etc. – is the bearing of the information costs necessary to separate by merit an initially hard-to-distinguish pool of cases.

B. SUMMARY JUDGMENT AS THE KEY “SEPARATING” PROCESS THAT ALLOWS COURTS TO DISTINGUISH CASES BY LEVEL OF MERIT (P)

Summary judgment is the critical stage for redressing the problem of cases existing in a pooling equilibrium. Typically coming at the end of discovery, summary judgment is the next point, after most discovery disputes, when the court can more meaningfully distinguish among cases. It is exactly the sort of “mini-trial” – reviewing all the evidence to assess case merit – necessary to decide discovery disputes accurately. In deciding summary judgment, courts allow to proceed to trial only claims a reasonable jury could decide either way, thereby weeding out both claims with the lowest probability of merit (i.e., by granting defendants summary judgment) and claims with the highest probability (i.e., by granting plaintiffs summary judgment). After summary judgment, the only claims left are the “close calls” in which additional evidence is most useful; summary judgment separates those close-call cases out of the pool.

Courts are stuck with a low-information pooling equilibrium until summary judgment, as illustrated by the following model of how the estimated probability that a lawsuit is meritorious varies as litigation progresses, based on how many cases get weeded out of the pool at each litigation stage and why (i.e., whether the reason cases are weeded out is that they lack merit). The following are the variables that influence estimates of the probability that a lawsuit is meritorious:

let: $d_1 =$ fraction of cases dismissed before discovery (on motions to dismiss)

$d_2 =$ fraction dismissed after discovery (on summary judgment motions)

$s_1 =$ fraction settling before discovery disputes arise

thus: $d_1 + s_1 =$ fraction not reaching end of discovery or summary judgment motion

$1 - d_1 - s_1 =$ fraction reaching end of discovery (“Stage II” below)

$d_1 + s_1 + d_2 =$ fraction not surviving past summary judgment motions

$1 - d_1 - d_2 - s_1 =$ fraction surviving summary judgment and thus going to trial or settling just before trial (“Stage III” below)

According to data and theory, there is a roughly 50/50 chance that a plaintiff
will prevail in a case going to trial – i.e., cases that survive dispositive motions (dismissal and summary judgment) and that the parties do not settle. But at no earlier stage does the court have much meaningful sense of the merits. Following is a discussion of what information the court has, or can infer, about case merit at three key stages of the path to trial: first, at the start of the case, before discovery, motions, or during-litigation settlement efforts (“Stage I”); next, after dismissal motions and early-litigation settlements, including settlements during discovery (“Stage II”); and next, after summary judgment motions (“Stage III”).

**Stage I – Start of case, before discovery, motions, or during-litigation settlements:** At this early stage, all the court knows is that the parties’ pleadings allege exactly opposite facts, and that there are various possible case outcomes: a pretrial finding that the case has no merit, either on a motion to dismiss or on a summary judgment motion; or that the case is the sort of close call (presumably with 50/50 odds for the plaintiff) that will survive dismissal motions; or that the parties will settle the case. This intuitive sense of the range of possibilities is easily formalized: the probability that a case is meritorious at Stage I ($P_1$)

---

150 George L. Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 J. LEGAL STUD. 1, 18-19 (1984) (noting that as litigation features fewer trials, “the proportion of plaintiff victories will approach 50 percent” under certain assumptions, such as that “plaintiff and defendant possess information that is on average of equal precision, and if the application of legal standards is, on the whole, coherent and predictable, . . . [and] to the extent [there is a] cost advantage of settlement over litigation”). Priest and Klein collect “substantial evidence” for their “selection hypothesis” that cases selected for trial will tend to be close calls. Id. at 55 (recounting evidence, id. at 31-53).

Others dispute the data and theory underlying the 50 percent hypothesis. See, e.g., Steven Shavell, *Any Frequency of Plaintiff Victory at Trial is Possible*, 25 J. LEGAL STUD. 493, 494 (1996) (“Data on the frequency of plaintiff victory does not clearly support the 50 percent tendency.”). Shavell notes that the 50/50 hypothesis may fail if there are certain information problems or if the majority of lawsuits are meritorious. These conditions seem likely, if at all, only in certain case types, such as cases that are especially uncertain, and thus hard to settle (or hard for courts to dismiss when unmeritorious), because they arise under a new statute. See, e.g., Ruth Colker, *The Americans With Disabilities Act: A Windfall For Defendants*, 34 HARV. C.R.-C.L. L. REV. 99, 100 (1999) (reporting, several years after Americans with Disabilities Act, that “defendants prevailed” in more than ninety-three percent of reported ADA employment discrimination cases decided on the merits at the trial court level. Of those cases that are appealed, defendants prevail in eighty-four percent of reported cases. These results are worse than results found in comparable areas.”).

151 Courts might also rule based on tentative hunches, of questionable accuracy in any individual case, about case merit based on proxies such as case type. See, e.g., In re First Constitution S’holder Litig., 145 F.R.D. 291, 293 (D. Conn. 1991) (“Securities fraud actions are recognized as being particularly vulnerable to strike suits . . . . [T]his action belongs to a class that is subject to strike suits.”). Most courts, however, deny considering case merit in discovery decisions. See supra note 116 and accompanying text. Courts also might see in the pleadings an apparent flaw that justifies limiting discovery until a dismissal motion. See, e.g., Flaim v. Med. Coll. of Ohio, 418 F.3d 629 (6th Cir. 2005) (upholding limited discovery pending dismissal motion, given that complaint seemed highly questionable and qualified immunity defense seemed promising).
depends on the fraction of cases weeded out on dismissal and summary judgment ($d_1$ and $d_2$, respectively) and weeded out via settlement ($s_1$), as well as the likelihood that a settled case was meritorious ($P_{s1}$):

$$
P_I = \sum \text{the fraction of cases with each possible outcome multiplied by the probability that a case with that outcome is meritorious}
$$

$$
P_I = (0)(d_1 + d_2) + P_{s1}(s_1) + (.5)(1 - d_1 - d_2 - s_1)
$$

$$
= P_{s1}(s_1) + (.5)(1 - d_1 - d_2 - s_1)
$$

$$
= P_{s1}(s_1) + .5 - .5(d_1 + d_2) - .5s_1
$$

The above shows that the probability a case is meritorious at the start of litigation ($P_I$) is less than 0.5, except under two unlikely assumptions: (1) there would have to be few enough cases dismissed on motions (i.e., low $d_1+d_2$) that removing good cases from the pool by settlement dominates the opposite effect of removing weak cases by dismissal, contrary to estimates that over one-third of federal cases are dismissed on motions;152 and (2) settled cases would have to be on average highly meritorious (i.e., high $P_{s1}$), contrary to recent (limited) data indicating that many confidential settlements are for modest sums.153

Thus, judges’ likely intuition is that initially, the probability that a case is meritorious is low (i.e., $P_I < 0.5$), but that assessed probability increases during pretrial processes, eventually reaching 0.5 for cases surviving summary judgment (Stage III). The question is whether the court’s estimate of case merit rises primarily from Stage I to Stage II (i.e., as a result of dismissal motions and early settlements), or primarily from Stage II to Stage III (i.e., as a result of summary judgment motions); as discussed below, it is primarily the latter.

**Stage II – After dismissal motions and early settlements (including settlements during discovery):** At this stage, when most discovery disputes occur, the court has two pieces of information it lacked at the start of litigation: (1) the case survived pre-discovery dismissal motions; and (2) it did not settle early. But early dismissals and settlements, taken together, clarify little about case merit. Dismissal motions weed out only the most clearly meritless lawsuits, given that Supreme Court has cautioned against granting them too readily.154

---


153 See Kotkin, *infra* note 155.

154 In *Swierkiewicz*, supra note 34, the Court held that motions to dismiss for failure to state a claim rarely should be granted in discrimination suits, one of the most common lawsuit types (see *infra* note 193). The Court may have shown more willingness to allow such dismissals in *Bell Atlantic v.*
Settlements typically are confidential, preventing the court from knowing the terms of settlement or looking any further into the merits, so the court knows nothing meaningful about the merits of settled cases.

The probability a case is meritorious at Stage II ($P_{II}$) can be estimated by noting that the probability of merit of a Stage I (just-filed) case is the weighted average of the following possibilities: (1) that a case survives to Stage II (the fraction $1-d_1-s_1$ of all cases, with $P_{II}$ probability of merit); (2) that a case loses on a dismissal motion (fraction $d_1$, which by definition have zero probability of merit); and (3) that a case settles early (fraction $s_1$, with $P_{s1}$ probability of merit):

\[
P_I = (P_{II})(1 - d_1 - s_1) + (0)(d_1) + (P_{s1})(s_1)
\]

\[
(P_{II})(1 - d_1 - s_1) = P_I - (P_{s1})(s_1)
\]

\[
P_{II} = ( P_I - P_{s1}s_1 ) / (1 - d_1 - s_1)
\]

This equation shows that by Stage II, we know somewhat, but not much, more about the probability that a case has merit. Knowing how many cases lose on motions to dismiss helps: the probability that a Stage II case has merit ($P_{II}$) is higher when the more cases lose on dismissal motions (i.e., high $d_1$), because weeding out unmeritorious cases leaves the remaining pool more meritorious. While we know the effect of more dismissals ($d_1$), we do not know the effect of higher settlement rates ($s_1$) because we do not know the merits of settled cases:

- If early-settling cases are mostly unmeritorious (e.g., if defendants mostly pay small nuisance-value settlements of a few thousand dollars in weak cases), then early settlement weeds out weak cases, leaving the remaining pool (Stage II cases) of higher merit; i.e., $P_{II} > P_I$.

Twombly, which dismissed an antitrust complaint that insufficienly alleged conspiracy. 127 S.Ct. 1955 (2007). Yet Twombly denied abrogating Swierkiewicz and may be more of a heightened antitrust pleading standard than a major change to general standards for dismissal motions.

155 See Moss, Illuminating Secrecy, supra note 148, at 867-69 (2007) (noting prevalence of confidentiality clauses in settlements); Bone, Economics of Civil Procedure, supra note 90, at 19 (“Empirical research in this area is extremely difficult to conduct because most lawsuits settle and settlements mask evidence of frivolousness.”). The one known study of confidential settlements found that in one federal district, the median confidential settlement size was $30,000 in employment discrimination, and $181,500 in personal injury, cases. Minna J. Kotkin, Outing Outcomes: An Empirical Study Of Confidential Employment Discrimination Settlements, 64 WASH. & LEE L. REV. 111 (2007). But most such settlements were late in litigation, after discovery or summary judgment, so the study sheds only a little light on the merits of cases that settle early.

156 See, e.g., Fletcher v. City of Fort Wayne, 162 F.3d 975, 976 (7th Cir. 1998) (finding plaintiffs not “prevailing parties” due to size of $2,500-$5,000 settlements; settlement “for less than the costs of defense is a good working definition of a nuisance-value settlement, unless . . . the stakes of the case are themselves small.”); see generally Moss, Illuminating Secrecy, supra note 148, at 899-900.
• If early-settling cases are *mostly meritorious* (e.g., if defendants pay mostly to avoid liability and incriminating disclosures), then settlement decreases average Stage II merit; if $P_{s1}$ is high, then as $s_1$ rises, $P_H$ falls.
With dismissals weeding out the unmeritorious while settlements weed out the meritorious, we cannot say if $P_I$ or $P_H$ is higher.\(^{157}\)

In short, between Stage I and Stage II, dismissal motions weed out weak cases while settlements weed out cases of unclear merit. It seems likely that the dismissal of weak cases dominates the theoretically possible effect of settling strong cases,\(^ {158}\) which would mean that Stage II cases have higher average merit (i.e., $P_H > P_I$). But so little information about settlements,\(^ {159}\) we cannot make any truly confident statements. Accordingly, courts face much the same dearth of information about Stage II cases that they face in Stage I.

**Stage III – After summary judgment motions:** At this stage, as discussed above, theory and data suggest that the remaining cases have a roughly 50/50 probability of merit; i.e., $P_{III} = 0.5$.\(^ {160}\) With case merit largely unknowable at Stages I and II, Stage III is the first point in time at which courts meaningfully

\[ \frac{\partial P_H}{\partial s_1} < 0? \]
\[ \frac{\partial P_H}{\partial s_1} = \frac{\partial}{\partial s_1} \left[ \frac{P_1 - P_{s1}}{1 - d_1 - s_1} \right] \]
\[ = \frac{[1 - d_1 - s_1](P_{s1}) - (P_1 - (P_{s1})(s_1))(-1)] / (1 - d_1 - s_1)^2}{1 - d_1 - s_1} \]
\[ = \left[ -P_{s1} + P_{s1}d_1 + P_{s1}s_1 + P_1 - P_{s1}s_1 \right] / (1 - d_1 - s_1)^2 \]
\[ = \left[ P_{s1}(d_1 - 1) + P_1 \right] / (1 - d_1 - s_1)^2 \]

\[ \frac{\partial P_H}{\partial s_1} > 0? \]
\[ \text{if } P_{s1}(d_1 - 1) + P_1 > 0 \]
\[ P_{s1}(d_1 - 1) > -P_1 \]
\[ P_{s1} > P_1(1-d_1) \]

\[ \frac{\partial P_H}{\partial s_1} > 0 \text{ only if } P_{s1} \text{ is higher than } P_1 \text{ discounted by the fraction of cases not dismissed.} \]

\(^{157}\) With settled case merit unknown, varied settlement frequency has indeterminate effects; we cannot tell whether increasing settlements leaves the remaining case pool higher- or lower-merit:

\[^{158}\text{See supra notes 153-152 and accompanying text (discussing evidence of high rates of dismissals of weak cases and of limited value of settled cases).}\]

\[^{159}\text{See supra note 155.}\]

\[^{160}\text{See supra note 150 (discussing Priest/Klein hypothesis and its critics).}\]
can assess case merit, and therefore the first point when cases exist largely in a separating rather than a pooling equilibrium. It is the stage when courts finally can know enough about case merit to decide discovery disputes accurately.

Yet delaying discovery decisions until summary judgment seems to conjure up the hole-in-the-bucket problem again: summary judgment should base on all the evidence, so how can evidence-gathering decisions wait until summary judgment? As discussed below, there is room for a narrow but important practice of making some discovery decisions post-summary judgment.

C. THE PRESCRIPTION: IN CLOSE CALLS, PRESERVING THE EVIDENCE BUT DELAYING THE DISCOVERY UNTIL AFTER SUMMARY JUDGMENT

Because summary judgment motions ideally are evaluations of all the evidence, they typically come after all the evidence is gathered in discovery. Summary judgment before discovery closes is, and should be, exceptional because “discovery should precede consideration of dispositive motions when the facts sought to be discovered are relevant to consideration of the particular motion.” For this reason, the summary judgment rule provides that if additional discovery is reasonably available, courts should not grant summary judgment without that discovery, but instead should “deny the motion [or] order a continuance to enable . . . other discovery to be undertaken.” Courts granting summary judgment before completion of discovery risk reversal, as in Gary Plastic Packaging Corp. v. Merrill Lynch, Pierce, Fenner & Smith, which

161 See Bone & Evans, supra note 29, at 1284 (“[T]he procedural system seems to favor postponing a serious evidentiary review until after substantial discovery has been completed. Summary judgment, for example, usually takes place only after the parties have had ample opportunity to uncover information and evidence . . . .”); Natural Res. Def. Council v. Curtis, 189 F.R.D. 4, 8 (D.D.C. 1999) ([L]eaving the question of the suffic iency of plaintiffs' case as a matter of law to a point after discovery closes is the way in which the federal courts handle such matters. Therefore, plaintiffs are correct . . . that they are not required to establish a legally sufficient case . . . of the applicability of FACA . . . as a condition of securing discovery and that resolution of the legal issues concerning that applicability is premature until discovery ends.”) (emphases added).

162 Coastal States Gas Corp. v. Dep't of Energy, 84 F.R.D. 278, 282 (D. Del. 1979) (citing Canavan v. Beneficial Fin. Corp., 553 F.2d 860, 865 (3d Cir. 1977)); see also Concord Labs. v. Concord Med. Ctr., 552 F. Supp. 549 (N.D. Ill. 1982) (holding that if case entails “knowledge and intent” issues, “material evidence is almost entirely in the hands of the defendants, and where plaintiff can establish a fair likelihood that it can obtain material evidence through discovery, we think it unfair to grant defendants summary judgment until plaintiff has had a full opportunity”); United States v. Price, 577 F. Supp. 1103, 1115-16 (D.N.J. 1983) (“where a plaintiff must obtain a good deal of information from the opposing party, judgment should be withheld until the discovery process has been completed”) (citing Nat'l Life Ins. v. Solomon, 529 F.2d 59, 61 (2d Cir.1975)).


explained why complete discovery should precede a grant of summary judgment:

While summary judgment is a valuable procedural device . . . , it is also a drastic remedy that cuts off the right to have one's day in court. The harshness of the remedy is exacerbated when the trial court refuses to allow plaintiff to conduct discovery. Discovery serves important purposes, such as . . . fully disclosing the nature and scope of the controversy, . . . framing the issues involved, and enabling parties to obtain the factual information needed to prepare for trial. . . . Summary judgment should be sparingly granted . . . when discovery is incomplete and . . . defendants have exclusive possession of the material facts.\textsuperscript{165}

Consequently, there is good reason that “[m]ost courts are reluctant to grant summary judgment prior to the termination of discovery.”\textsuperscript{166}

But putting summary judgment after all discovery is just a common-sense convention, not a rule. “[T]here is no requirement in Rule 56 . . . that summary judgment not be entered until discovery is complete.”\textsuperscript{167} In appropriate cases, courts entertain limited-scope summary judgment motions after only partial discovery; examples include motions for summary judgment limited to threshold questions such as a governmental defendant’s claim of immunity from suit\textsuperscript{168} or a libel defendant’s assertion that only limited evidence is necessary to undercut plaintiff’s required showing of the falsity of the allegedly libelous statement.\textsuperscript{169}

\textsuperscript{165} Gary Plastic Packaging Corp., 756 F.2d at 236. \textit{Accord} Conn. Bank of Commerce v. Republic of Congo, 309 F.3d 240, 264 (5th Cir. 2002) (“[the] court should have allowed full discovery [to] . . . allow[] the Bank a fair opportunity to present all available material evidence pertinent to its opposition to . . . summary judgment”); Weiss v. Reebok Int’l, 91 Fed. Appx. 683, 690, 2004 WL 434158, at *4 (Fed. Cir. 2004) (“Weiss has not had ample opportunity for discovery[, which] . . . was stayed pending resolution of Reebok’s summary judgment motion that was narrowly focused on the structural aspects [of the disputed shoes] . . . . Weiss should be granted the time that all litigants receive to gather . . . evidence that the accused shoes can perform the claimed functions”).


\textsuperscript{167} Pub. Serv. Co. of Colo. v. Cont’l Cas. Co., 26 F.3d 1508, 1518 (10th Cir. 1994) (emphasis added); see also Paul Kadair, Inc. v. Sony Corp., 694 F.2d 1017, 1029-30 (5th Cir. 1983) (“[A] plaintiff’s entitlement to discovery prior to a ruling on a motion for summary judgment is not unlimited, and may be cut off when the record shows that the requested discovery is not likely to produce the facts needed by plaintiff to withstand . . . summary judgment.”).

\textsuperscript{168} See, e.g., Moore v. Busby, 92 Fed. Appx. 699, 2004 WL 389461 (10th Cir. 2004) (holding that district court was permitted to stay discovery pending disposition of the summary judgment motion by the defendant, a judge, on the threshold question of his immunity from suit as a judge).

\textsuperscript{169} See, e.g., Living Will Ctr. v. NBC Subsidiary (KCNC-TV), 857 P.2d 514 (Colo. Ct. App. 1993) (holding that “[l]imited discovery on the issue of falsity is therefore appropriate” before summary judgment motion: “[D]iscovery pertaining to defendants’ state of mind, . . . is not pertinent to the
In certain cases, some burdensome discovery could be allowed only after summary judgment. The premise of this suggestion is not that more discovery should be delayed. Rather, it is that in a meritorious case, certain burdensome discovery is regularly denied under current practice—and must be denied because courts cannot tell the case is meritorious (and therefore is deserving of more discovery than usual) during the pooling equilibrium that exists before summary judgment. In such a case, the summary judgment denial is a determination that the case is one in which more discovery is warranted than in the broader pool of all cases in discovery; it is a determination that relatively more discovery is warranted than the court could have assumed during discovery.

Notably, courts’ existing broad case management powers over discovery and summary judgment make a new rule technically unnecessary. There already is “a great deal of discretionary power in the trial court” as to discovery,\(^\text{170}\) including as to “the control and scheduling of discovery”\(^\text{171}\) and “the appropriateness and timing of summary adjudication under Rule 56.”\(^\text{172}\)

In short, if a court denies costly discovery when a case is hard to distinguish from the “pool” of all cases in discovery (Stage II), it should reconsider that denial of discovery if the case survives summary judgment (\(i.e.,\) reaches Stage III). Surviving summary judgment separates a case from a broader pool (all cases in discovery) into a narrower one (cases reaching trial). More specifically, a summary judgment denial means a reasonable jury could decide either way, \(i.e.,\) \(p\) is roughly 0.5—higher than in most cases, which means that more evidence is more valuable than in most cases (\(i.e.,\) \(Ap\) of additional evidence is high). The key problem courts face in deciding discovery disputes is that they would need “mini-trials” to assess case merit sufficiently; summary judgment is the existing point in litigation when the court already undertakes that effort. In deciding summary judgment, the court is bearing the information costs necessary to switch from a pooling equilibrium (where \(p\) and \(Ap\) are hard to discern) to a separating equilibrium (where it is clearer which are high-\(p\), high-\(Ap\) cases). Based on this analysis, courts’ effort to assess case merit on summary judgment can serve double-duty, helping courts decide discovery disputes that, earlier in litigation, they had trouble deciding because case merit and evidentiary value was unclear.

---


\(^{171}\) Fed. R. Civ. P. 16(c)(6).

\(^{172}\) Fed. R. Civ. P. 16(c).
D. THE DEVIL IS IN THE DETAILS: MAKING WORKABLE THE PROPOSAL FOR POST-SUMMARY JUDGMENT REVISITING OF DISCOVERY THE COURT HAD DENIED EARLIER

In proposing a different way for courts to handle nuts-and-bolts practical matters like discovery disputes, the details matter. This subpart discusses five legitimate concerns about this Article’s proposal and responds to those concerns by fleshing out how this proposal best could be implemented.

CONCERN #1: SUMMARY JUDGMENT MAY NOT CLARIFY CASE MERIT

A denial of summary judgment does not always mean a case is a close call; weak cases can survive summary judgment when they are fact-intensive or depend on debatable inferences from the facts.

This Article’s premise is that cases surviving summary judgment have higher – roughly 50/50 – odds of success than the broader pool of all filed cases. Some summary judgment denials, though, do not indicate such high odds of success. Decisions denying summary judgment sometimes actually say that the case “barely” survives summary judgment,173 or that it does so despite “weak” evidence.174 Further, summary judgment “is not commonly interposed, and even less frequently granted,” in certain areas of law. For example, in negligence lawsuits, “the judge and jury each have a specialized function”;175 such cases often turn on pure factual disputes (e.g., drivers disputing who entered an intersection first) or “reasonableness” and “due care” inquiries fuzzy enough that even in cases that seem weak, it is hard for the court to say no reasonable jury could find for the plaintiff. The same may hold for other areas of law featuring similar “reasonableness” tests, such as “unreasonable” use of force by police.176

But the ability of a party to avoid summary judgment by citing factual disputes is less than it once was. Since the 1980s, the Supreme Court has “signal[ed] to the lower courts that summary judgment can be relied upon more

173 Sylvester v. SOS Children’s Villages, 453 F.3d 900, 904 (7th Cir. 2006) (“There is no rich mosaic of circumstantial evidence of retaliation in this case, but there is enough (though maybe barely enough) to preclude summary judgment.”); Smith v. Mattox, 127 F.3d 1416, 1419 (11th Cir. 1997) (holding the evidence of excessive use of force “barely” enough because the “hazy border between permissible and forbidden force is marked by a multifactored, case-by-case balancing test,” precluding ruling on the level of force “within the confines of summary judgment review”).

174 MetroNet Servs. Corp. v. U.S. West Commc’n, 329 F.3d 986, 1008 (9th Cir. 2003) (“Although the evidence of the financial harm to MetroNet is weak, it is sufficient to withstand summary judgment.”); Colburn v. Trustees of Indiana Univ., 739 F. Supp. 1268, 1293 (S.D. Ind. 1990) (“Plaintiffs’ evidence . . . is weak, but . . . just enough to get them past summary judgment.”)


176 See, e.g., Smith v. Mattox, supra note 173.
so than in the past to weed out frivolous lawsuits and avoid wasteful trials, and the lower courts have responded accordingly. As one much-cited case noted, courts “cannot resolve factual disputes that could go to a jury at trial, but weak factual claims can be weeded out through summary judgment motions,” because the mere “existence of a triable [fact] issue” is insufficient to avoid summary judgment; “the triable issue must be evaluated in its factual context, which suggests that the test for summary judgment is whether sufficient evidence exists in the pre-trial record.”

Similarly, “the fact that a summary judgment is difficult to obtain in actions in which the parties’ states of mind are relevant does not mean that it will never be granted. . . . [S]ummary judgment has been granted to defendants in suits involving fraud, conspiracy, and other claims turning on state of mind when plaintiffs’ allegations were not sufficiently supported.” Thus, courts do meaningfully assess case merit on summary judgment even on claims that are quite fact-specific or that turn on state of mind.

Still, with some weak cases surviving summary judgment, this Article’s proposal will not be useful in every case surviving summary judgment. It is unsurprising that this Article’s proposal is imperfect, because its premise is that there is no perfect fix. Judges make sub-optimal discovery decisions not because they are bad at their jobs or because the rules are badly written, but because of the nature of the information-timing problem: courts lack sufficient information on case merit and evidentiary value to make optimal discovery decisions.

When summary judgment denials do not indicate case merit, judges should not, and will not, view that denial as sufficiently informative to affect their prior discovery rulings. Such uninformative summary judgment denials mean that the pooling equilibrium, in which judges have too little information to make optimal discovery decisions, will persist until trial, because summary judgment does not move the case from a pooling equilibrium into a separating equilibrium consisting mainly of higher-than-average merit “close call” cases warranting more discovery. But the judge will know this; after all, the judge, having sifted

178 Collins v. Associated Pathologists, Ltd., 844 F.2d 473, 476 (7th Cir. 1988); see also Thompson Everett, Inc. v. Nat’l Cable Advertising, 57 F.3d 1317, 1322 (4th Cir. 1995) (deeming “mere existence of some disputed facts” insufficient, because “the quality and quantity of the evidence offered to create a question of fact must be adequate to support a jury verdict. Thus, if the evidence is ‘merely colorable’ or ‘not significantly probative,’ it may not be adequate to oppose entry of summary judgment.”) (quoting Anderson v. Liberty Lobby, Inc., 477 U.S. 242, 249-50 (1986)).
179 CHARLES ALAN WRIGHT ET AL., 10A FED. PRAC. & PROC. CIV.3d § 2730 (2007) (collecting cases); see Betkerur v. Aultman Hosp. Ass’n, 78 F.3d 1079, 1087 (6th Cir. 1996) (“Cases involving state of mind issues are not necessarily inappropriate for summary judgment.”) (citation omitted).
through each party’s evidence and arguments to assess how a reasonable jury could rule, is well-positioned to know whether his or her summary judgment denial was or was not based on the merits of the case.

Consequently, the fact that some summary judgment denials do not indicate case merit means this proposal will not be useful in all cases. Importantly, though, it does not create a risk of bad post-summary judgment discovery grants, because judges will know when their summary judgment denials indicate enough about case merit to warrant reconsideration of their denials of discovery.

**Concern #2: Courts Should Use Alternatives Such as Sampling and Cost-Shifting**

Where courts hesitate to allow potentially relevant but costly discovery, they need not postpone it until summary judgment, because they have two alternatives more in conformity with existing discovery practice: ordering cost-shifting that allows the discovery only if the requesting party is willing to pay some or all of the cost; or ordering a partial sampling of high-volume discovery.

Sampling and cost-shifting are useful discovery tools but, as discussed below, they are not a panacea and do not eliminate the information-timing problem that makes post-summary judgment discovery potentially useful.

One tool courts currently use is sampling, offering a fraction of the data first, then the rest if the sample is promising.\(^{180}\)

A phased approach will allow the Court to engage in a more meaningful benefit-burden analysis before determining whether to require cost-shifting . . . . After Defendant restores a portion of the back-up tapes . . . . Plaintiff will then have the opportunity . . . . to determine if it contains relevant evidence and if additional restoration of back-up tapes is warranted. . . . [R]estoration of one-fourth . . . should be adequate to determine whether the tapes are likely to possess relevant evidence.\(^{181}\)

Yet sampling is useful only under certain conditions: (1) if a sample is much cheaper than all the evidence (which is not true if, e.g., the main cost is finding a way to read old data); and (2) if the key evidence is likely present in a limited sample (which is not true if, e.g., plaintiff seeks one particular email, because its absence from a sample will not prove it does not exist). Sampling is only a

---

\(^{180}\) For cases requiring sampling of high-volume deleted data, see McPeek v. Ashcroft, 202 F.R.D. 31 (D.D.C. 2001); Zubulake III, 216 F.R.D. at 281-82; AAB Joint Venture v. United States, 75 Fed. Cl. 432, 443-44 (Fed. Cl. 2007).

\(^{181}\) AAB Joint Venture, 75 Fed. Cl. at 443-44 (citing McPeek, 202 F.R.D. at 34-35) (citations omitted).
limited fix because of the above conditions and because it does not redress the key problem of courts’ difficulty deciding whether to allow costly discovery before seeing much of the evidence.

Cost-shifting, whether under current rules\(^{182}\) or to a greater degree,\(^{183}\) gives courts a wider range of options for costly discovery than “yes, you can obtain it” and “no, you cannot” – but it is a limited fix that does not redress the information timing problem. To begin with, requiring requesting parties to pay for responding parties’ production costs jeopardizes non-wealthy plaintiffs’ ability to serve the important social function of suing to unearth and redress important violations of law.\(^{184}\) More fundamentally, allowing cost-shifting in limited circumstances does not eliminate courts’ need to make difficult-to-impossible decisions about the “value” of requested discovery.

That is, even 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.

In sum, both sampling and cost-shifting have their place as important tools that, in some cases, can help courts expand their options for discovery rulings. But neither of these tools fully redresses the information-timing problem this Article diagnoses, so neither eliminates the utility of this Article’s proposal.

**CONCERN #3: JUDGES MIGHT EXCESSIVELY DENY DISCOVERY AND GRANT SUMMARY JUDGMENT**

Judges might respond to this proposal by denying more discovery (as a way out of difficult proportionality decisions) and by granting summary judgment more often (both to avoid cumbersome post-summary judgment discovery and because plaintiffs will be less able to obtain evidence they need to oppose summary judgment).

This concern is real but, for three reasons, should not be overstated. First, a court using this Article’s proposal to deny too much discovery risks reversal on

\(^{182}\) See supra notes 62-65 (as to cost-shifting rules), 196-208 (as to cost-shifting case law).

\(^{183}\) See supra note 69.

\(^{184}\) See Hay, supra note 87 (discussing how discovery helps plaintiffs prove and redress illegality).
appeal. While courts cannot always allow all discovery parties want before summary judgment, appellate courts do enforce the rule of as full discovery as possible before summary judgment, reversing courts that grant summary judgment after unduly denying discovery. Courts are aware of this presumption that pre-summary judgment discovery should be as full as possible; this awareness would not disappear if courts adopt this Article’s proposal.

Second, this Article is not suggesting postponing most e-discovery. The media focus on the costliest cases, but much e-discovery is modest and should remain in standard (pre-summary judgment) discovery. A simple, non-technical search can respond to a request for all e-mails with certain text; some “deleted” emails remain easily accessible on company servers; and some backup file restoration is affordable. Even costly discovery (e.g., many deleted files, or “metadata” in fraud cases addressing when a document was created or altered) need not always be delayed, because courts can initially allow a partial sample.

Third, courts already deny much costly discovery, so this Article would give parties a better chance at such discovery – just later, after summary judgment. In addition to the cases denying relevant discovery due to cost, e-discovery decisions in employment cases (which are 12-14% of federal cases) show that currently, the best-case scenario for a plaintiff in even a high-value case may be a court order allowing costly e-discovery only with cost-shifting, i.e., only if the

---

185 See supra notes 163-166.
186 Janet Novack, “Control/alt/discover,” FORBES, pg. 60 (Jan. 13, 1997) (telling how one e-discovery consultant charged over $1 million for a court-ordered search of 50,000 tapes, which the article cast as “blackmail” (by plaintiff, not the consultant), a “form of legal torture,” and a violation of some “gentleman’s agreement” not to go after each other’s electronic data” among lawyers generally).
187 See, e.g., Zubulake I, 217 F.R.D at 315 (“[S]imply . . . create a plain language search . . . [for] ‘header’ information, such as the date or the name of the sender, . . . [or] the text of the e-mail . . . . UBS personnel could easily run a search for e-mails containing the words ‘Laura’ or ‘Zubulake.’”).
188 For example, this author has no technical skills but once recovered many “deleted” emails that remained accessible from university servers in an email account sub-folder.
189 See, e.g., Semroth v. City of Wichita, 239 F.R.D. 630, 638 (D. Kan. 2006) (rejecting argument that e-mails on backup tapes were “not readily accessible” where estimated cost was $3,374.95).
190 Panel Discussion, supra note 18, at 22 (2007) (comments by Hon. James C. Francis IV) (noting that “metadata” includes “changes to the document over time [and] who the author of the document is,” as well as when the computer was used on the document, which may help assess “the authenticity of documents” and a party’s “intent . . . in drafting” them).
191 See supra notes 180-181 and accompanying text.
192 See supra notes 75-76, 79 (collecting cases denying seemingly relevant discovery due to cost).
193 See Ann Hodges, Mediation and the Transformation of American Labor Unions, 69 Mo. L. Rev. 365, 369 (2004) (also noting that such cases “also substantially increased in many state courts”).
plaintiff pays an often prohibitively high share of defendants’ production costs.

Consider *Zubulake*. Title VII (employment discrimination) monetary relief is only lost pay plus capped emotional distress and punitive damages, but *Zubulake*’s damages “undoubtedly” were “higher than that of the vast majority of Title VII” cases: with millions in lost pay, plaintiff claimed damages of $15 million, and defendant counter-estimated $1 million. Though the $165,000 e-discovery cost was “surely not ‘significantly disproportionate’” to case value and “weighed against cost-shifting,” the court still shifted 25% of that cost to plaintiff, even though the evidence was relevant and plaintiff made a “limited and targeted request”: emails about her from five individuals. Defendant initially produced a small sample, five of 94 backup tapes, and “a review of these e-mails reveal[ed] that they are relevant” to *Zubulake*’s claim of termination not for performance but due to discrimination by her supervisor Chapin and others:

> [T]hey tell a compelling story of the dysfunctional atmosphere . . . [and] Chapin’s behavior . . . . [T]he e-mails contradict testimony given by UBS employees . . . . An e-mail from Chapin . . . acknowledg[ed] that *Zubulake*’s “ability to do a good job . . . is clear,” and that she is “quite capable.” . . . [E]-mail contains the precise words used by the author[,] . . . a particularly powerful form of proof at trial . . . as an admission.

The “marginal utility” of the evidence “may be quite high,” but just “potentially,” because the sample lacked “direct evidence of discrimination” – an oddly high threshold, given a recent Supreme Court holding that “direct evidence” is not necessary to prove discrimination. Faulting plaintiff for inability to prove with certainty that it would find a smoking gun in as-yet-unseen discovery, the court held that despite the “powerful” admissions in emails that tell a “compelling story,” marginal utility analysis weighed only “slightly against cost-shifting.”

In sum, *Zubulake* imposed cost-shifting despite finding that of the seven rank-ordered cost/benefit factors, the first four tipped against cost-shifting, the

---

194 42 U.S.C. § 1981a(b)(3) (capping emotional distress and punitive damages at $50,000-$300,000).
195 *Zubulake III*, 216 F.R.D. at 281, 288 (recounting that *Zubulake* lost a $650,000 annual salary as an equities trader at a major New York securities firm).
196 *Id*. at 288, 291.
197 *Id*. at 281-82, 285.
199 *Id*. at 287.
201 *Zubulake III*, 216 F.R.D. at 287.
next two were neutral, and only the seventh supported cost-shifting. The odds are low of any discrimination plaintiff obtaining costly e-discovery without cost-shifting, if Zubulake could not do so despite proven relevance from a partial sample, high case stakes, a carefully tailored discovery request, and all the weightiest factors in the seven-factor test militating against cost-shifting.

Even high-dollar, high-import class actions do not always obtain costly e-discovery. In *Wiginton v. CB Richard Ellis*, 202 a nationwide class claimed sex harassment of over 1,000 employees, 203 and the court followed *Zubulake* but ordered plaintiffs to pay even more: 75% of the $249,000 cost to search backup tapes for high-relevance evidence “relating to CBRE’s workplace environment,” such as “pornograph[y] . . . distributed electronically (i.e., via e-mail) and displayed on [office] computers.” 204 A “test search” (sampling) “result[ed] in relevant documents that had not been produced” earlier – 1.64% to 6.5% of sampled emails were relevant – so the evidence clearly was “only available through restoring and searching the backup tapes.” 205 Yet the court viewed those statistics negatively: “marginal utility” was low because the sample “revealed a significant number of unresponsive documents.” 206 To say that finding hundreds of emails required searching thousands, however, is a criticism not of utility but of cost. Further, *Wiginton* deemed the case stakes of a class action under a major remedial federal statute (Title VII) insufficient to justify the discovery:

Plaintiffs claim that should a class be certified, their class recovery could extend into the tens of millions of dollars. While the Court cannot completely accept Plaintiffs’ speculative estimate . . . , neither can it accept that their claims are worthless . . . . Nevertheless, several hundred thousand dollars for one limited part of discovery is a substantial amount . . . . Therefore, this factor weighs in favor of cost-shifting. 207

Neither did claims of mass harassment help prove “[t]he importance of the issues at stake”; quoting *Zubulake*, *Wiginton* held that “this factor ‘will only rarely come into play . . . . [D]iscrimination in the workplace . . . is hardly unique.’” 208

“Publication bias” – the fact that published decisions are just the tip of the

---

204 *Wiginton*, 229 F.R.D. at 570.
205 Id. at 571, 574.
206 Id. at 575 (emphasis added).
207 Id.
208 Id. at 576 (quoting *Zubulake II*, 216 F.R.D. at 289).
iceberg, and unpublished or unwritten orders may be very different\(^\text{209}\) – is especially salient for discovery rulings, which often deny discovery in unwritten, unappealable\(^\text{210}\) oral orders at court conferences.\(^\text{211}\) In one unrecorded court conference in a typical Title VII case,\(^\text{212}\) the judge said, “that is insane, insane!” when the plaintiff’s attorney\(^\text{213}\) stated plans to depose ten employees (the number the rules deem permissible without court permission\(^\text{214}\)); the judge also said Title VII plaintiffs “never” get to see the personnel files of “comparators” (those who got the disputed job), which commonly are part of discovery in Title VII cases\(^\text{215}\).

With courts often grudging about even run-of-the-mill discovery, such as ten depositions and relevant personnel files, and with courts refusing to allow much e-discovery without cost-shifting in even high-value, high-import cases like Zubulake and Wiginton, there is little current hope for plaintiffs in most cases to obtain costly e-discovery – unless, under this Article’s proposal, the case proves its merit to the judge by surviving summary judgment. The effect of this Article’s proposal on most cases would be to allow plaintiffs a second chance, post-summary judgment, to seek discovery they rarely get now.

Nevertheless, as stated above, a valid concern does remain that courts might misuse this proposal to deny more discovery. To address that concern, a new rule, though not required, would be advisable. A new rule could make the intent of this proposal as clear and enforceable as possible to district courts making discovery decisions and appellate courts reviewing grants of summary judgment. Alternatively, a new Advisory Committee note to the rules on case management, discovery, or summary judgment\(^\text{216}\) could clarify similarly.\(^\text{217}\) Whatever the


\(^{210}\) See supra note 50.

\(^{211}\) See Fed. R. Civ. P. 16 (providing for court conferences on discovery, trial scheduling, etc.).


\(^{213}\) The author of this Article was that unfortunate plaintiff’s lawyer.


\(^{216}\) Fed. R. Civ. P. 16, 26 and 56, respectively.

\(^{217}\) Although the Judicial Conference has a policy of not issuing new Advisory Committee Notes without a new rule, that policy is not mandated by any law or rule of civil procedure.
the message is that post-summary judgment discovery should be a vehicle not for restricting discovery, but for allowing more discovery – when a case proves worthy of more evidence-gathering by surviving summary judgment.

**CONCERN #4: JUDICIAL RELUCTANCE TO DELAY TRIAL TO REARGUE A DISCOVERY DISPUTE**

Judges might be reluctant to allow redundant rearguments of already-decided discovery disputes after summary judgment; such re-litigation of discovery disputes might undesirably delay trial.

Obviating this concern is the fact that the procedure for post-summary judgment reconsideration of discovery can be simpler than a full motion for reconsideration. One method is an expedited, streamlined motion: the court could entertain a short reconsideration motion on any discovery previously denied within one week of denying summary judgment. An even more streamlined method would be to allow parties to add to their summary judgment briefings a short discussion of possible post-summary judgment discovery:

1. the party opposing summary judgment (typically plaintiff) could submit, with its summary judgment opposition filing, a short (e.g., three-page) supplement to its summary judgment brief, stating what more discovery it wants if the case survives summary judgment;

2. the party moving for summary judgment (typically defendant), in its reply papers, could submit a similarly concise supplement arguing against that additional discovery; and

3. the court, if it denies summary judgment, could include in its decision an order stating what, if any, additional discovery is being granted and by when (e.g., within X weeks of the summary judgment ruling) that discovery must occur.

More broadly, courts already are creative in managing discovery, such as in sharing work between trial and magistrate judges, in interspersing limited-scope dispositive motions with partial discovery, and in allowing class actions

[^218]: The following is a possible bare-bones phrasing: “Where the court denies discovery it might have allowed were it clearer that additional evidence would prove helpful to the factfinder at trial, the court may reconsider that discovery denial if the case survives summary judgment.”

[^219]: See Richard A. Posner, *Coping with the Caseload: A Comment on Magistrates and Masters*, 137 U. PA. L. REV. 2215, 2216 (1989) (discussing Magistrate Judges’ discovery expertise: “Abuse . . . is more likely to occur in a case supervised by a district judge, whose primary responsibilities lie in trying cases and managing . . . docket[s], than in a case supervised by a magistrate, whose most challenging and responsible task is, precisely, to manage discovery in big civil cases.”).

[^220]: See supra note 151 (discussing decisions limiting discovery based on anticipated merits).
partial discovery limited to discerning the presence or absence of a true “class” before adjudicating the question of whether a class should be certified.\textsuperscript{221} There is no reason to think courts could not use the above two procedures, or quite likely better ones, to minimize any possible disruption or redundancy that might result from reexamining a discovery dispute after summary judgment.

**CONCERN #5: LOSS OF EVIDENCE WHILE DISCOVERY IS DELAYED**

Evidence might be lost or destroyed between a discovery dispute and a summary judgment denial: summary judgment might not occur until weeks or months after a discovery dispute; it can take months just to brief and argue summary judgment; and it can take months or over a year for courts to decide summary judgment motions.

In an order denying burdensome discovery, the court should issue a preservation order stating that the evidence requested should be preserved until the court decides any summary judgment motions. The extent of parties’ duties to preserve evidence is a key e-discovery battleground\textsuperscript{222} but not a new issue; preservation has been a high-stakes bone of contention among parties for decades in disputes about destroying evidence ranging from body parts\textsuperscript{223} to records of

\textsuperscript{221} See Charles Alan Wright et al., 7B Fed. Prac. & Proc. Civ. 3d § 1796.1 (2007) ("[I]nitially, . . . discovery should be limited to what is necessary for determining whether a proper class action exists."); see, e.g., Parker v. Time Warner Ent. Co., 331 F.3d 13, 21 (2 d Cir. 2003) ("[I]t is likely that at least minimal class discovery must be conducted in order to provide the court with the factual information necessary to decide whether or not to certify a Rule 23(b)(2) class.").

\textsuperscript{222} See, e.g., Kronisch v. United States, 150 F.3d 112, 126 (2d Cir. 1998) (requiring preservation of data upon notice of relevance to ongoing or impending litigation). Zubulake imposes a broad but not unlimited duty "to suspend [a] routine document retention/destruction policy" to preserve data:

Must a corporation, upon recognizing the threat of litigation, preserve . . . every e-mail or electronic document, and every backup tape? . . . Such a rule would cripple large corporations . . . that are almost always involved in litigation.

Once a party reasonably anticipates litigation, it must suspend its routine document retention/destruction policy . . . [with] a “litigation hold” . . . As a general rule, that litigation hold does not apply to inaccessible backup tapes . . . maintained solely for the purpose of disaster recovery[, which may continue to be recycled . . . [per] company[] policy. On the other hand, if backup tapes are accessible (i.e., actively used for information retrieval), then such tapes would likely be subject to the litigation hold.

However, . . . [i]f a company can identify where particular employee documents are stored on backup tapes, then the tapes storing the documents of “key players” to the existing or threatened litigation should be preserved if the information contained on those tapes is not otherwise available. This exception applies to all backup tapes.


\textsuperscript{223} See, e.g., Welsh v. United States, 844 F.2d 1239, 1244 (6th Cir. 1988) ("[Defendant’s] act of discarding the skull flap was, if not intentional, at least seriously negligent.").
Cold War-era CIA programs.\textsuperscript{224} Sometimes, a court actually subjects a party who destroyed necessary evidence to an “adverse inference” that the evidence would have been favorable to the other side.\textsuperscript{225}

Notably, any preservation controversies arising out of this Article’s proposal would be more limited than the usual preservation dispute. Currently, such disputes occur early in litigation, when a party demands preservation of all data on every computer system or data device,\textsuperscript{226} because it does not yet know what data or devices will prove relevant, and it does not want to lose data day by day while, over the first few weeks and months of litigation, it figures out exactly which data or devices actually are most relevant. Courts hesitate to make pre-discovery preservation orders unlimited in scope but do issue quite broad orders because of the uncertainty about what eventually will be discoverable.\textsuperscript{227}

The sort of preservation order most likely under this Article’s proposal, however, would be narrow, extending not to all data and all devices, only to specific devices with the specific data on which discovery was denied. In contrast, early in litigation a plaintiff might seek a preservation order against deleting any files or emails, disposing of any computers or personal data devices, or disposing of any paper files with any personnel matters. The preservation this

\textsuperscript{224} See, e.g., Kronisch v. United States, 150 F.3d 112, 116-17 (2d Cir. 1998) (allowing adverse inference against government defendant where “records were destroyed,” by CIA personnel, as to CIA program of “surreptitious administration of LSD to unwitting nonvolunteer subjects”).

\textsuperscript{225} See Welsh, 844 F.2d at 1244 (upholding “adverse inference” as to liability, because while “[defendant’s] negligent destruction of the skull flap does not lead to a conclusion that the medical care of Mr. Welsh was negligent[,] . . . [t]he destruction did, however, prejudice his legal rights”).

\textsuperscript{226} See, e.g., AAB Joint Venture v. United States, 75 Fed.Cl. 432, 443 (Fed. Cl. 2007) (holding that defendant had “a duty to preserve e-mails from July 2002 to the present, and that Defendant's decision to transfer the e-mails to back-up tapes does not exempt Defendant from its responsibility to produce relevant e-mails”). As one federal judge explained:

\textquote{[T]he costs of preservation can be exorbitant, not just . . . not recycling back-up tapes, but . . . implementing a litigation hold, just contacting everybody, finding out where the information resides . . . . [T]here is inevitably uncertainty about the scope . . . . Are you going to have to preserve back-up data? How far back are you going to have to preserve it? What are your employees going to be able to do in terms of deleting their e-mails?}

Panel Discussion, supra note 18, at 17 (comments by Hon. James C. Francis IV).

\textsuperscript{227} One judge explained the need for broad preservation orders as follows:

\textquote{In the paper realm, I can pretty well say, ‘And thou shalt not destroy any documents of this type’ . . . In the electronic arena, I am probably going to have to know which servers the data is likely to reside on, and perhaps who the individuals are whose e-mails have to be preserved. . . . [Y]ou may well have to preserve inaccessible data even though you will make an argument later on that you do not have to produce it.}

Panel Discussion, supra note 18, at 19 (comments by Hon. James C. Francis IV).
proposal would require would be more limited, covering only the particular evidence denied in discovery (e.g., preserving only emails) and only a limited duration – from the discovery dispute to the time summary judgment is decided (when the court would either allow the discovery or end the preservation order).

Further limiting the burden of the necessary preservation is that this Article’s proposal is not asserting that all disputed evidence must be preserved from the time of a discovery dispute to the time summary judgment is decided; preservation is necessary only if the court both (1) sees a real risk of evidence destruction and (2) sees the particular discovery as the sort of close call that it should deny to the requesting party in most cases but perhaps should grant if the case proves its worth by surviving summary judgment.

V. CONCLUSION

The prevalence of cost-benefit analyses of discovery presents a mismatch of problem and solution. The main problem with discovery decision-making is not that judges lack the skill to decide what discovery is insufficiently beneficial, or is beneficial but too costly. The problem is that courts face discovery decisions before they have cost-effective access to information needed to make those decisions – before cases separate from a pooling equilibrium in which there are many low-merit cases and in which individual case merit is hard to discern.

A problem of timing requires a solution of timing, such as this Article’s proposal for when a court finds a discovery dispute a close call, typically because the requested evidence is likely to be useful but too costly to be warranted in the mass of low-merit cases. In such situations, a court could postpone the discovery until after summary judgment, the point at which a case likely has 50/50 odds, justifying more costly evidence-gathering. The proposal aims not to deny or postpone more discovery, but rather to allow, after summary judgment, the sort of helpful but costly discovery that courts currently deny on a regular basis.

This proposal is not a perfect solution, but that is the point: there is no perfect solution to the information timing problem: that optimal discovery depends on case merit, which cannot be assessed until after discovery. The way to shift from a pooling to a separating equilibrium is for discovery rulings to consider as much evidence as possible – a prohibitively costly endeavor during discovery, but exactly what courts do on summary judgment. Postponing certain costly discovery until summary judgment is an imperfect solution but better than the prevailing impractical alternative: insisting that rules and judges make accurate discovery cost-benefit decisions without the information necessary to do so.
More broadly, economics is getting better at recognizing one of the defining aspects of litigation: how information emerges over time. For economics to provide accurate diagnoses and useful proposals, it must do more than just prescribe cost-benefit analysis; it must consider critical information-timing matters, whether by modeling cases with options theory, by noting differences in information disclosure by litigation stage, or – as this Article attempts with discovery – by modeling litigation disputes based on how information costs and merit signals change as litigation progresses.


229 See, e.g., Moss, Illuminating Secrecy, supra note 148 (analyzing settlement confidentiality based on information distinctions between settlements reached before and after litigation commences).