TWOMBLY, PLEADING RULES, AND THE REGULATION OF COURT ACCESS

Boston University School of Law Working Paper No. 08-34  
(November 26, 2008)

Robert G. Bone

This paper can be downloaded without charge at:

**Twombly, Pleading Rules, and the Regulation of Court Access**

By Robert G. Bone†

[Forthcoming 94 Iowa L. Rev. ___ (2009)]

**ABSTRACT**

In *Bell Atlantic Corp. v. Twombly*, the Supreme Court reconsidered *Conley v. Gibson*’s very liberal notice pleading standard and held that the plaintiff must allege enough to support a plausibility of wrongdoing. This Article considers the *Twombly* decision within the broader framework of court access regulation and sketches a normative roadmap for designing optimal pleading and merits-based case-screening rules. The Article begins with an analysis of *Twombly* itself. It argues, contrary to much criticism of the decision, that the Court’s plausibility standard represents only a modest departure from traditional notice pleading and that its interpretation of Rule 8(a)(2) is consistent with the text and history of the Rule and in line with the pragmatic vision of the original Federal Rule drafters. The Article then addresses the broader normative issues involved in regulating court access through stricter pleading and other case-screening devices. It argues that a pleading requirement along the lines of *Twombly*’s thin plausibility standard might be justified by a process-based theory of fairness as reasoning, but that anything stronger must be evaluated on outcome-based grounds. Applying utilitarian and rights-based metrics of outcome quality, the Article then explores various methods of screening meritless suits. It highlights several issues that are often ignored or misunderstood, including the importance of carefully defining the undesirable lawsuits to be screened, correctly identifying the causes of the problem, and proceeding cautiously in the absence of empirical information by designing regulatory responses to fit the most probable causes. It argues that information asymmetry is likely to be a more important cause of meritless litigation than the commonly assumed cost asymmetry, and it outlines a hybrid approach to handle the information-asymmetry cases. The Article concludes by emphasizing the importance of using formal rulemaking or the legislative process to design case-screening rules and making those rules substance-specific rather than trans-substantive.

**TABLE OF CONTENTS**

I. **Introduction**

II. *Twombly* and Notice Pleading

† Robert Kent Professor in Civil Procedure, Boston University School of Law. I would like to thank workshop participants at the University of Texas School of Law and Boston University School of Law. Special thanks to Mitch Berman, Kris Collins, Ian Farrell, Keith Hylton, Alexi Lahav, Ken Simons, and Patrick Woolley, and to Anthony Dutra for his able research assistance.
A. The *Twombly* Decision
B. *Twombly*’s Impact

III. *Twombly* in Historical Context

A. Rule 8(a)(2), Code Pleading, and 1938 Advisory Committee Intent
B. The Vision of the Federal Rule Drafters

IV. Optimal Pleading and Case-Screening Rules: A Normative Analysis

A. *Twombly*’s Plausibility Standard from a Process-Based Perspective
   1. Fairness to Defendants
   2. Fairness to Plaintiffs

B. Stricter Pleading and Case Screening from an Outcome-Based Perspective
   1. Choosing an Outcome-Based Metric
      a. Utilitarian
      b. Rights-Based
   2. Defining the “Undesirable” Lawsuit
   3. Analyzing the Regulatory Options
      a. Targeting Causes
      b. Targeting Symptoms
      (1) Gatekeeping Rules
      (2) Incentive-Shaping Rules

C. Toward an Optimal Approach

V. Conclusion

I. Introduction

Pleading rules are once again a hot topic in civil procedure circles. At the end of its 2007 term, the Supreme Court decided three pleading cases,¹ and in June 2008, the

Court granted certiorari in another case that raises pleading issues.\textsuperscript{2} In what is surely the most controversial of these decisions, \textit{Bell Atlantic Corp. v. Twombly}, the Court held that to survive a motion to dismiss, antitrust class action plaintiffs had to state facts in their complaint sufficient to support a “plausible” inference of conspiracy.\textsuperscript{3} Many judges and academic commentators read the decision as overturning fifty years of generous notice pleading practice, and critics have attacked it as a sharp departure from the “liberal ethos” of the Federal Rules favoring decisions “on the merits, by jury trial, after full disclosure through discovery.”\textsuperscript{4}

This reaction is not surprising. In a society that relies heavily on the courts for rights vindication and social reform, the distribution of litigating power is vital to the

\textsuperscript{2} \textit{Iqbal v. Hasty}, 490 F.3d 143 (2d Cir. 2007), cert. granted \textit{Ashcroft v. Iqbal}, 2008 U.S. LEXIS 4906 (June 16, 2008).

\textsuperscript{3} \textit{Twombly}, 127 S.Ct. at 1964-65.

distribution of social and political power.\(^5\) As a result, rules that regulate court access often trigger intense controversy, especially when they make it more difficult for plaintiffs to sue.\(^6\) In civil procedure, for example, debates in the nineteenth century focused on the technical obstacles created by common law pleading, and in the early twentieth century on similar problems with code pleading.\(^7\) More recently, Rule 11 penalties and heightened pleading requirements have been major targets of concern.\(^8\)

This Article situates the Supreme Court’s *Twombly* decision in this broader framework. It views *Twombly* not so much as a pleading decision but rather as a court access decision, one that addresses a general problem of institutional design: how best to

---

\(^5\) In his sharp *Twombly* dissent, Justice Stevens hints at the impact he believes the decision will have on the distribution of power. *See Twombly*, 127 S.Ct. at 1989 (Stevens J., dissenting) (“The transparent policy concern that drives the decision is the interest in protecting antitrust defendants—who in this case are some of the wealthiest corporations in our economy—from the burdens of pretrial discovery.”).

\(^6\) *See, e.g.*, Phillips v. County of Allegheny, 515 F.3d 224, 230 (3d Cir. 2008) (“Few issues in civil procedure jurisprudence are more significant than pleading standards, which are the key that opens access to courts.”). Of course, concerns about court access extend well beyond the civil procedure examples mentioned in the text paragraph. In the 1970s, for example, a debate raged over the constitutionality of filing fees for certain types of cases. *See, e.g.*, LAURENCE H. TRIBE, AMERICAN CONSTITUTIONAL LAW 1008–10, 1111–12 (1978); Frank I. Michelman, *The Supreme Court and Litigation Access Fees: The Right to Protect One’s Rights—Part I*, 1973 DUKE L.J. 1153 (1973). Also, the Supreme Court’s standing cases have evoked heated controversy for decades, with the critics complaining that restrictive standing doctrines impede socially desirable forms of public law litigation. *See ERWIN CHEMERINSKY, CONSTITUTIONAL LAW: PRINCIPLES AND POLICIES 60–66 (3d ed. 2006).*


prevent undesirable lawsuits from entering the court system. From this broader perspective, screening more aggressively at the front door by demanding more from the complaint is just one approach, with its own costs and benefits, and should be evaluated relative to other alternatives. It makes no sense, for example, to strengthen pleading requirements if the same result can be achieved much better by bolstering Rule 11 sanctions, placing stricter limits on discovery in complex cases, shifting fees, or allowing truncated summary judgment determinations based on targeted and phased discovery.

The Supreme Court is in a poor position to make these choices in individual cases. Even conceding broad latitude for interpretation, the Justices cannot, for example, put teeth into Rule 11 at odds with its clear language or substantially alter the discovery rules. If these options are desirable, they must be implemented through the established process for making and amending the Federal Rules of Civil Procedure. Moreover, some alternatives, such as fee shifting, might require congressional action because of the Rules Enabling Act’s prohibition on Federal Rules that “abridge, enlarge, or modify any substantive right.”

---

9 The Rules Enabling Act, 28 U.S.C. §§ 2072–2073, vests the authority to make procedural rules for the federal courts in the United States Supreme Court and creates a multi-stage rulemaking process with opportunities for public input. A proposed amendment to an existing Federal Rule of Civil Procedure or a new Rule is first considered by the Advisory Committee on Civil Rules. If the proposal meets with the Advisory Committee’s approval, it is forwarded to the Standing Committee and then to the Judicial Conference, and eventually on to the United States Supreme Court. If the Supreme Court approves the proposal, it is forwarded to Congress, which has the opportunity to veto it. If Congress does not exercise its veto within the prescribed period, the proposal goes into effect. See generally Robert G. Bone, The Process of Making Process: Court Rulemaking, Democratic Legitimacy, and Procedural Efficacy, 87 GEO. L. J. 887, 892 (1999).

10 28 U.S.C. § 2072(b). All types of fee shifting involve one party paying all or a portion of the opposing party’s attorney fees. The obligation to pay fees can be triggered by different conditions, such as a trial loss, failure to improve on a settlement offer at trial, or filing a meritless suit.
Given these limitations, the Court would be well advised to exercise restraint in adapting existing Federal Rules to screening cases more aggressively. And that is just what it has done. In fact, as I argue below, the Supreme Court’s decision in *Twombly* does not alter pleading rules in as drastic a way as many of its critics, and even some of its few defenders, suppose.

This is a critical time to examine these issues. *Twombly* has already had a major impact. The case was cited a startling 4000 times during its first nine months,\(^{11}\) and the Supreme Court’s recent grant of certiorari in *Ashcroft v. Iqbal* sets the stage for a possible clarification of the plausibility standard.\(^{12}\) The Standing Committee on Civil Rules discussed *Twombly*’s impact at its January 2008 meeting,\(^ {13}\) and scholarly commentary promises to continue for some time to come.\(^ {14}\)

The remainder of this Article is divided into four parts. Part II examines the *Twombly* decision with care. It argues that, contrary to much criticism of the case, the Court’s plausibility standard marks only a modest departure from notice pleading. Properly interpreted, it requires no more than that the allegations describe a state of affairs that differs significantly from a baseline of normality and supports a probability of

---


\(^ {12}\) *Iqbal v. Hasty*, 490 F.3d 143 (2d Cir. 2007), *cert. granted* *Ashcroft v. Iqbal*, 2008 U.S. LEXIS 4906 (June 16, 2008). The certified questions also raise substantive issues relating to the scope of a *Bivens* action against federal officials, so it is not certain that the Court will address the pleading issue.


\(^ {14}\) The AALS Civil Procedure Section chose a topic for its January 2009 annual section meeting, “The Changing Shape of Federal Pretrial Practice,” that is inspired in part by the *Twombly* decision. *See Section on Civil Procedure: Call for Papers*, July 10, 2008 (copy on file with author). At least two of the three papers selected for the session deal with pleading and *Twombly*. 
wrongdoing greater than the background probability for situations of the same general type.

Part III places Twombly in historical perspective. It describes the genesis of Rule 8(a), the federal pleading rule, and argues that Twombly’s pleading standard neither signals a return to code fact pleading nor represents a sharp break from the vision of the 1938 Federal Rule drafters, as some critics have argued.\(^\text{15}\) The drafters were pragmatists, who assumed that procedural rules would be “continually changed and improved” as litigation conditions changed.\(^\text{16}\) Twombly reflects a similar pragmatism.

Part IV addresses the normative issues involved in designing optimal pleading and case-screening rules. In what is perhaps the most important part of the Twombly opinion, the Court concedes, in the clearest possible way, that trial judges are likely to have difficulty managing complex litigation on a case-specific basis given informational and strategic constraints. Part IV argues that the Twombly Court is correct to question the efficacy of case-specific discretion and that this calls for consideration of more rule-based regulatory approaches, such as strict pleading. Moreover, the shift from discretion to rule forces explicit attention to the normative issues involved in constructing an optimal system.

In analyzing these normative issues, Part IV begins, in Section A, by considering whether a thin version of Twombly’s plausibility standard is required by a process-based principle of fairness as reason-giving. Section B then considers pleading rules that require more factual detail than thin plausibility, including the heightened level of

---

15 See, e.g., Spencer, supra note 4, at 431–33, 445–46.
specificity that critics of *Twombly* fear lower courts might interpret plausibility to require. These stricter rules must be justified, if at all, on outcome-based grounds, as tools to screen undesirable lawsuits.

Section B also highlights four central points, often overlooked, that are crucial to an outcome-based analysis. First, any such analysis must choose between utilitarian and rights-based metrics for evaluating outcome quality. These two metrics weigh costs and benefits differently and have different implications for regulatory choices. Second, it is also important to define the case-screening goal clearly. In particular, there is a crucial difference between screening meritless suits and screening weak suits that nonetheless have merit. Third, it is important to identify the cause of undesirable filings as precisely as possible. The *Twombly* Court and many commentators assume cost asymmetry is the principal cause, but informational asymmetry is likely to have a much stronger effect, and informational asymmetry requires a different response. Finally, it is important to consider all regulatory options, not just strict pleading; compare the costs and benefits of each, and choose the combination that best addresses the problem given its most likely cause.

Part V concludes. It draws the different strands together and highlights two important general lessons. First, adjustments to pleading burdens require a global and systemic analysis that is possible only through the formal committee-based rulemaking process or through Congress and not by common law adjudication or creative interpretation of the Federal Rules. Second, any approach to regulation of court access should be substance-specific in order to take account of the different cost-benefit tradeoffs for different types of cases.
II. Twombly and Notice Pleading

A. The Twombly Decision

Bell Atlantic Corp. v. Twombly was a consumer antitrust class action brought against local telephone and telecommunications carriers, alleging a conspiracy in violation of Section 1 of the Sherman Act. To understand the allegations, one must know something about the history of the telephone industry. In 1984, the American Telephone & Telegraph Company (“AT&T”) was ordered to divest its local telephone business. The local business was taken over by regional monopolies, known as incumbent local exchange carriers (“ILECs”). In 1996, Congress forced the ILECs to share their networks with competing companies seeking entry into the local markets. These competing companies are known as competitive local exchange carriers (“CLECs”).

The antitrust claim in Twombly was brought on behalf of a class consisting of “at least 90 percent of all subscribers to local telephone or high-speed Internet services in the continental United States” against “America’s largest telecommunications firms” operating as ILECs. The plaintiffs sought treble damages as well as declaratory and injunctive relief. The complaint alleged that the ILEC defendants restrained trade by adopting common measures to “inhibit the growth of upstart CLECs” and by agreeing not

---

18 This history is taken from the summary in the Twombly opinion. Id. at 1961–62.
19 Id. at 1961
20 Id.
21 Id.
22 Twombly, 127 S.Ct. at 1967.
23 Id. at 1962.
to compete in one another’s markets.\textsuperscript{24} As for the critical element of conspiracy, paragraph 51 of the complaint stated that the defendants “entered into a contract, combination or conspiracy to prevent competitive entry . . . and have agreed not to compete with one another and otherwise allocated customers and markets to one another.”\textsuperscript{25}

The defendant ILECs filed a motion to dismiss for failure to state a claim, arguing that the complaint did not allege facts sufficient to infer the express or tacit agreement required for a violation.\textsuperscript{26} The district judge granted the motion and dismissed the complaint.\textsuperscript{27} The Second Circuit Court of Appeals reversed, holding that the complaint need only allege facts sufficient to put the defendant on notice of the general nature of the dispute.\textsuperscript{28} The Supreme Court granted certiorari and reversed the Court of Appeals over the vigorous dissent of Justices Stevens and Ginsburg.\textsuperscript{29} The seven-Justice majority held that the complaint failed to allege enough to support a “plausible” inference of agreement.\textsuperscript{30}

\textit{Twombly} triggered a sharp response from the academic community almost immediately, most of it criticizing the Court for tightening up on pleading requirements.\textsuperscript{31} The main reason for the furor has to do with the Court’s discussion of the

\textsuperscript{24} \textit{Id.}
\textsuperscript{25} \textit{Id.} at 1962–63 (quoting from the complaint).
\textsuperscript{26} \textit{See id.} at 1964 (noting that Section 1 requires more than parallel conduct).
\textsuperscript{28} \textit{Twombly} v. Bell Atlantic Corp., 425 F.3d 99 (2d Cir. 2005).
\textsuperscript{29} \textit{Twombly}, 127 S.Ct. 1955.
\textsuperscript{30} \textit{Id.} at 1970–74.
\textsuperscript{31} \textit{See} Hoffman, \textit{supra} note 4, at 7 (“The majority view among academics has been that robust efforts to regulate at the pleading stage are wrongheaded, inconsistent with the traditional pleading standard the Court has followed since \textit{Conley}.”).
old civil procedure chestnut, *Conley v. Gibson.*\(^{32}\) *Conley,* decided in 1957, is the hallmark case that single-handedly established Rule 8(a)(2) as a notice pleading rule.\(^{33}\)

Notice pleading requires only a general description of the events sufficient to give fair notice to the defendant (and the court) of what the dispute generally is about so the defendant can file an Answer or otherwise respond.\(^{34}\) It assigns the fact-gathering process to the discovery rather than the pleading stage, and most important, it rejects case screening as a pleading function.

Notice pleading proponents focus their concern on *Twombly*’s treatment of language in *Conley* quoted for fifty years as the key statement of the notice pleading standard: “a complaint should not be dismissed for failure to state a claim unless it appears beyond doubt that the plaintiff can prove no set of facts in support of his claim which would entitle him to relief.”\(^{35}\) The *Twombly* majority rejected a literal interpretation of this passage\(^{36}\) and held that a complaint’s allegations must support a “plausible” not merely a “possible” inference, one that rises above a “speculative level.”\(^{37}\) It defended the plausibility standard not only as good policy, but also as a proper interpretation of Rule 8(a)(2).\(^{38}\)


\(^{33}\) Id. at 47.

\(^{34}\) Id.

\(^{35}\) Id. at 45–46.

\(^{36}\) *Twombly,* 127 S.Ct. at 1968 (“This ‘no set of facts’ language can be read in isolation as saying that any statement revealing the theory of the claim will suffice unless its factual impossibility may be shown from the face of the pleadings.”).

\(^{37}\) Id. at 1964–65.

\(^{38}\) Specifically, the Court reasoned that *Conley*’s broad formulation was inconsistent with Rule 8(a)(2)’s requirement that the plaintiff provide not only a “short and plain statement of the claim” but also a statement “showing that the pleader is entitled to relief.” Id. at 1964–65 & n.3.
The Court’s criticism of Conley has caused a great deal of confusion. Lower court judges and commentators at first questioned whether Twombly’s holding might be confined to antitrust cases like Twombly itself; but the lower courts appear to have answered this question in favor of a more general application. A much more difficult issue involves determining exactly how the plausibility standard changes previous 8(a)(2) pleading law. The term “plausible” obviously refers to the strength of the inference from allegation to necessary factual conclusion. It is useful to think about an inference as a conditional probability: in this case, the probability that the defendants entered into an agreement conditional on the allegations in the complaint being true. “Plausible” corresponds to a probability greater than “possible.” Exactly how much greater is uncertain.


40 See, e.g., Phillips v. County of Allegheny, 515 F.3d 224, 234 (3d Cir. 2008) (“[W]e decline at this point to read Twombly so narrowly as to limit its holding on plausibility to the antitrust context.”); Davis v. Coca-Cola Bottling Co. Consol., 516 F.3d 955, 974 n.43 (11th Cir. 2008) (“We understand Twombly as a further articulation of the standard by which to evaluate the sufficiency of all claims brought pursuant to Rule 8(a).”); Perez-Acevedo v. Rivero Cubano, 520 F.3d 26, 29 (1st Cir. 2008) (section 1983 claim must “raise right to relief above a speculative level”); Williams v. Gerber Products Co., 523 F.3d 934 (9th Cir. 2008) (applying plausibility standard to tort and breach of warranty claims); Sonnier v. State Farm Mutual Automobile Insurance Co., 509 F.3d 673 (5th Cir. 2007) (extending plausibility standard to civil contract disputes); see also Kendall W. Hannon, Note, Much Ado About Twombly? A Study of the Impact of Bell Atlantic v. Twombly on 12(b)(6) Motions, 83 NOTRE DAME L. REV. 101 (forthcoming 2008) (finding that reported cases apply Twombly to a wide range of different types of claims).

41 See, e.g., Robbins v. Oklahoma, 519 F.3d 1242, 1247 (10th Cir. 2008) (calling the new formulation “less than pellucid”); Phillips v. County of Allegheny, 515 F.3d 224, 230 (3d Cir. 2008) (describing Twombly’s impact as “confusing”); Iqbal v. Hasty, 490 F.3d 143, 155, 157 (2d Cir. 2007) (noting the Supreme Court’s “conflicting signals” and the “considerable uncertainty” created by Twombly).

42 The Court does make clear that it is not imposing a “probability” standard, perhaps suggesting that plausible lies somewhere between possible and probable. See Bell Atlantic Corp. v. Twombly, 127 S.Ct. 1955, 1965 (2007). It is also clear that the Court does not mean to impose a standard as strict as the “strong inference” that the Private Securities Litigation Reform Act requires for scienter allegations in a securities fraud claim. 15 U.S.C. § 74u-4(b)(2). See Tellabs, Inc. v. Makor Issues & Rights, Ltd., 127
B. **Twombly’s Impact**

In evaluating *Twombly*’s impact, one must first recognize that there are two distinct issues any system of pleading has to resolve: first, it has to identify which elements of a dispute the plaintiff must allege, and second, it has to determine how much the plaintiff must allege about each such element. Even in a very liberal notice pleading system, the complaint must do more than tell an interesting story. While it need not state each element explicitly, it must offer some reason to believe that the story it tells is linked to the elements of a legal claim. What this means is that a judge evaluating whether a given complaint meets the relevant pleading requirement must have a standard that specifies the elements that should appear in a complaint and a criterion for evaluating whether the complaint adequately reveals those elements. *Twombly* deals with the second part—how strict the criterion of adequacy should be.

What *Twombly* actually says about this issue is difficult to determine because the Court appears to send conflicting signals on the subject. On the one hand, it condemns *Conley*’s “no set of facts” language. On the other, it emphasizes that its plausibility

---

43 See *Conley v. Gibson*, 355 U.S. 41, 47 (1957) (holding that Rule 8(a) requires that the complaint give the defendant “fair notice of what the plaintiff’s claim is and the grounds upon which it rests”); 5A WRIGHT & MILLER, supra note 8, § 1215 at 194, § 1216 at 236. Even Justice Stevens in his sharp *Twombly* dissent recognized that “had the amended complaint in this case alleged only parallel conduct, it would not have made the required ‘showing’ . . . Similarly, had the pleadings contained only an allegation of agreement, without specifying the nature or object of that agreement, they would have been susceptible to the charge that they did not provide sufficient notice . . . .” *Twombly*, 127 S.Ct. at 1979 n.6.

44 See *Iqbal v. Hasty*, 490 F.3d 143, 157 (2d Cir. 2007) (noting the “conflicting signals”).

45 *Conley*, 355 U.S. at 45–46.
standard is not a form of "heightened pleading." Moreover, it endorses the skeletal complaint in Form 9 (now Form 11) appended to the Federal Rules of Civil Procedure as an example of acceptable pleading, and cites with approval several previous decisions that seem to apply extremely liberal pleading standards.

To confuse matters even further, just three weeks after the Court decided *Twombly*, it upheld the sufficiency of a complaint in *Erickson v. Pardus* without even mentioning the plausibility standard. The Court went out of its way to chastise the lower courts for a "stark" departure from what it later in the opinion called "the liberal pleading standards" of Rule 8(a)(2). Moreover, the Court characterized the 8(a)(2) standard in language reminiscent of traditional notice pleading.

Despite these seemingly contradictory signals, evaluating *Twombly*’s impact on notice pleading is not as difficult as some critics believe. The Court’s signals appear conflicting only if one assumes that *Twombly* substantially tightens pleading requirements. But this assumption is incorrect. To be sure, *Twombly* departs from notice pleading on a policy level. It elevates case screening to an important pleading function in

---

46 *Twombly*, 127 S.Ct. at 1973 n.14 (noting that the creation of heightened pleading standards is for the formal rulemaking process or for Congress).

47 After the recent style amendments, Form 9—a model complaint for automobile negligence—appears as Form 11 and the specific date and location in the original have been replaced with placeholders. See FED. R. CIV. P. Form 11.

48 *Erickson* v. *Pardus*, 127 S.Ct. 2197 (2007) (per curiam). Erickson, a prisoner acting *pro se*, alleged that the prison administrators’ decision to withhold his Hepatitis C treatment violated the Eighth Amendment’s prohibition on cruel and unusual punishment. *Id.* at 2198-99. The lower courts held that Erickson had not adequately pleaded substantial harm from the withheld treatment distinct from whatever harm he might otherwise have suffered as a result of his existing Hepatitis C condition. *Id.* at 2199. The Supreme Court reversed. *Id.* at 2200.

49 *Id.* at 2198, 2200.

50 *Id.* at 2200 (emphasizing that “specific facts are not necessary”).
sharp contrast to notice pleading’s exclusive focus on giving notice. This is a significant change, but equally significant is the modest nature of the Court’s move in this direction. The Court’s approval of liberal pleading does not contradict its holding; it qualifies and explains that holding by signaling that “plausibility” should not be interpreted as a demanding standard.

Indeed, *Twombly* itself was a particularly suitable candidate for a pleading stage dismissal because of two factors. First, the suit has many of the features usually associated with a significant risk of nonmeritorious filings and huge potential discovery costs. Second, the allegations of the complaint do not provide a strong enough signal that the case is in fact meritorious.

As to the first factor, *Twombly* involves an antitrust conspiracy claim with the possibility of treble damage recovery. It is brought as a nationwide class action with

---

51 *Id.* at 1966–67 (arguing that the plausibility requirement is needed to avoid high discovery costs and to weed out meritless suits before the threat of these high costs pressures defendants to settle). It is easy to miss this important distinction if one focuses only on doctrine and the linguistic shift to “plausibility.” Indeed, some courts and commentators have tried to subsume the “plausibility” standard under notice pleading by arguing that the Court merely specified what was required to give the defendant fair notice. *See, e.g.,* Phillips v. County of Allegheny, 515 F.3d 224, 233–34 (3d Cir. 2008) (linking together “plausibility,” a Rule 8 “showing,” and notice); Airborne Beepers Video, Inc. v. AT&T Mobility LLC, 499 F.3d 663, 668 (7th Cir. 2007) (“At some point the factual detail in a complaint may be so sketchy that the complaint does not provide the type of notice of the claim to which the defendant is entitled under Rule 8”); Smith, *supra* note 4, at 24 (arguing that *Twombly* simply sets out the requirements for adequate notice: “It is only where the complaint states a logically coherent theory of liability that a defendant is truly ‘on notice’ of the claim against it.”). This effort might reflect an implicit realization that pleading standards are about fairness to defendants, not just system efficacy, fairness to plaintiffs, or efficiency. I discuss this point *infra* notes 135-161 and accompanying text. Whether or not *Twombly*’s plausibility standard can be justified in fairness terms, however, the Court’s opinion clearly mentions case screening as an appropriate pleading function. *See generally* Marcus, *supra* note 4, at 452 (observing that the traditional notice-giving function requires very little by way of allegations when the defendant has access to broad discovery).

It targets events that occurred over a seven-year period and are likely to generate massive amounts of discoverable information, including e-mail and electronic records. And it names as defendants some of the largest companies in the telecommunications industry.

These elements together create a situation where defendants might settle even suits they know lack merit in order to avoid crippling discovery costs and a nontrivial risk of an erroneous trial verdict. The *Twombly* Court recognizes this problem and rejects the conventional case management solution. The majority argues that judges operate under severe informational constraints impeding effective discovery management and cannot use summary judgment to weed out nonmeritorious cases when discovery costs pressure pre-summary-judgment settlements. In the Court’s opinion, these problems elevate the importance of case screening at the pleading stage.

To be sure, enhancing the pleading burden risks screening meritorious suits, but this is where the second factor—the nature of the *Twombly* allegations—comes into play. Based on those allegations alone, it is relatively unlikely, or at least highly uncertain, that the *Twombly* suit is in fact meritorious. The allegations describe a state of affairs that is not merely consistent with lawful competition, but fits neatly within the normal baseline

53 *Id.* at 1967 (noting that the class consists of at least 90 percent of all subscribers in the continental United States).
54 *Id.*
55 *Twombly*, 127 S.Ct. at 1967. The named defendants were Verizon Communications Inc., AT&T Inc., Qwest Communications International Inc., and BellSouth Corporation.
56 See infra notes 196-208 and accompanying text. The fact that defendants have a great deal of discoverable information and plaintiffs very little gives plaintiffs settlement leverage, especially when information about the merits is asymmetrically distributed. In addition, the class action aspect coupled with treble damages creates a potential award sizable enough to support a credible trial threat even for a meritless suit, assuming a large enough error risk at trial.
58 *Id.*
of conduct expected from a vigorously competitive telecommunications market.\textsuperscript{59} It is true that this conduct could also be the product of illegal agreement, but the complaint gives no reason to believe that it is, other than the plaintiffs’ conclusory statement to that effect.\textsuperscript{60} Thus, the complaint fails to support a probability of wrongdoing any greater than what exists under normal conditions in the market, and this background probability is not enough to support a lawsuit. If it were, anyone could file a suit about virtually anything merely by alleging perfectly ordinary conduct and arguing a possibility of wrongdoing.\textsuperscript{61}

It is tempting to conclude that there must be something amiss when competing firms stay out of one another’s markets and use common techniques to deter entry into their own. But this is an example of a baseline problem. Parallel conduct of this sort might seem odd when compared to the baseline of competitive behavior in general.\textsuperscript{62} But this is the wrong baseline for the \textit{Twombly} case. The correct baseline is competitive


\textsuperscript{60} \textit{Twombly}, 127 S.Ct. at 1970. Some readers might object to my calling this allegation “conclusory.” In deciding a 12(b)(6) motion to dismiss, the rule is that the judge must assume all plaintiff’s allegations are true except those that are conclusory. The question therefore is why the \textit{Twombly} plaintiffs’ allegations of agreement are conclusory. The Court’s answer is that a “fair reading” of the complaint shows that the plaintiffs themselves intended the allegations to be conclusions based on other allegations in their complaint. \textit{See id.} (“Although in form a few stray statements speak directly of agreement, on fair reading these are merely legal conclusions resting on the prior allegations.”). Had the plaintiffs good reason to suspect a conspiracy apart from the existence of parallel conduct, their allegations of agreement might have sufficed, although they might also have had to plead their additional reasons.

\textsuperscript{61} See \textit{Twombly}, 127 S.Ct. at 1969 (remarking that “Mr. Micawber’s optimism would be enough” to bring suit).

\textsuperscript{62} The \textit{Twombly} Court makes a point of noting that the allegations might be enough to support agreement in a different market setting. \textit{Twombly}, 127 S.Ct. at 1972 (“In a traditionally unregulated industry with low barriers to entry, sparse competition among large firms dominating separate geographical segments of the market could very well signify illegal agreement. . .”). This different setting would be associated with a different baseline.
behavior under the particular conditions of the telecommunications market. And compared to that baseline, there is nothing necessarily odd about what the defendants are doing.\textsuperscript{63}

By a “baseline,” I mean the normal state of affairs for situations of the same general type as those described in the complaint. The probability of wrongdoing for baseline conduct is not necessarily zero, but it should be very small, for otherwise the conduct in question would not be part of a socially acceptable baseline. Understood in these terms, what the \textit{Twombly} Court requires are allegations that differ in some significant way from what usually occurs and differ in a way that supports a higher probability of wrongdoing than is ordinarily associated with baseline conduct. For example, a negligence complaint might describe erratic behavior that deviates from what would normally happen under similar circumstances. A breach of contract claim might allege acts and events that ordinarily do not occur without a contractual agreement, and then allege a failure to do something that would normally be done if there were an agreement.

This baseline explanation for \textit{Twombly’s} holding helps to explain the more confusing aspects of the opinion. For example, some courts and commentators claim to be confused by the Court’s apparent approval of the very skeletal Form 9 complaint appended to the Federal Rules of Civil Procedure.\textsuperscript{64} The original Federal Rule drafters included a set of forms to illustrate the types of pleadings and motions that satisfy the

\textsuperscript{63} Id. at 1972 (concluding that there is a “natural explanation” for the parallel conduct alleged given the distinctive history of the telecommunications market).

\textsuperscript{64} See, e.g., Iqbal v. Hasty, 490 F.3d 143, 156 (2d Cir. 2007) (citing \textit{Twombly’s} approval of Form 9 as a conflicting signal); Dodson, \textit{supra} note 4, at 127 (noting tension between plausibility pleading and Form 9).
requirements of the Rules. Form 9 (now renumbered as Form 11\textsuperscript{65}) is a complaint for negligence arising from an automobile accident. The relevant allegations state: “On June 1, 1936, in a public highway called Boylston Street in Boston, Massachusetts, defendant negligently drove a motor vehicle against plaintiff who was then crossing said highway.”\textsuperscript{66} These allegations, while thin, are enough to support the existence of a duty of reasonable care (automobile drivers owe such a duty to pedestrians) and a substantial probability that the defendant breached that duty.\textsuperscript{67} After all, drivers do not usually strike pedestrians when driving with reasonable care, so the probability of negligence conditional on a pedestrian being struck should be quite high.

The baseline in the Form 9 case is the behavior of automobile drivers in general, which supports the inference of a breach of duty from a pedestrian being struck. This is the appropriate baseline because Form 9 does not suggest any other and there is no reason to think that driving on Boylston Street in Boston is materially different from driving on any other street.\textsuperscript{68}

\textsuperscript{65} See \textit{supra} note 47.
\textsuperscript{66} \textit{FED. R. CIV. P.} Form 9 ¶ 2. The complaint also alleges that the plaintiff suffered a broken leg and other losses as a result. The restyled Form 11 omits the particular date and place references. See \textit{supra} note 47.
\textsuperscript{67} To be sure, the plaintiff might have been negligent himself, but contributory negligence is for the defendant to allege.
\textsuperscript{68} The Court’s subsequent holding in \textit{Erickson v. Pardus}, 127 S.Ct. 2197 (2007), easily fits this framework. In \textit{Erickson}, a pro se plaintiff alleged that prison officials withheld his Hepatitis C treatment wrongfully and that he would suffer irreparable and life-threatening damage to his liver as a result, in violation of the Eighth Amendment’s prohibition on cruel and unusual punishment. \textit{Id.} at 2198–99. The defendants argued that the complaint failed to allege that the plaintiff would suffer “substantial harm” distinct from whatever harm he would have suffered anyway as a result of his condition. \textit{Id.} at 2199. The sensible baseline, given the prison regulations entitling inmates to treatment, is that the plaintiff would receive treatment for his serious illness. Withholding treatment deviates significantly from this baseline and in a way that correlates with substantial harm. After all, people suffering from serious illnesses usually get much worse when they do not receive medical treatment. Therefore, it is certainly plausible based on these allegations that the defendants’ decision to withhold treatment would cause substantial harm to the plaintiff.
There is another aspect that distinguishes *Twombly*. Not only is the likelihood of a meritorious suit much lower in *Twombly* than in the Form 9 negligence case, but the social cost of the litigation if suit turns out to be meritless is much higher. If meritless plaintiffs in a large treble-damages class action are able to get past the pleading stage and use the threat of discovery to leverage a large settlement, the result might disrupt competition in the telecommunications market, which would be directly contrary to antitrust goals. 69 By contrast, the impact of an unjustified settlement or trial verdict in a Form 9 automobile negligence case is likely to be much less serious.70

Defining the appropriate baseline will not always be easy, and in any event it involves a normative judgment. But it should be clear in many cases. For example, the

---

69 As Part IV.B explains, it is important to consider not only the likelihood that suit is meritless as opposed to meritorious, but also the cost of allowing a meritless suit to proceed compared to the cost of screening a meritorious suit. In *Twombly*, the costs are substantial on both sides. If the allegation of agreement is, in fact, correct, the dismissal would frustrate efforts to restore competition in a major market (although there are other options, such as action by the Federal Trade Commission). By the same token, if the allegation of agreement is incorrect, allowing the suit to proceed forward and secure an unjustified settlement might also disrupt competition in the telecommunications market.

70 I am not suggesting that the Court meant its plausibility standard to apply only to complex cases with a high risk of costly meritless suits. Some have suggested as much, but such an interpretation fits the language of *Twombly* rather poorly. See, e.g., Smith, *supra* note 4, at 18–20 (collecting sources suggesting a limitation to complex cases and criticizing this approach as a poor reading of *Twombly*). The Court condemns *Conley’s* “no set of facts” formulation for all lawsuits, not just for complex suits.
baseline in *Twombly* follows from an application of economic theory to the particular conditions of the telecommunications industry as alleged in the complaint and revealed through publicly available sources. And the baseline in the Form 9 example follows from ordinary experience with automobiles.71

My interpretation of *Twombly*’s plausibility standard differs from interpretations others have given. Some scholars, for example, read the Court to require factual allegations supporting each element of the legal claim.72 The problem with this formulation is that the *Twombly* plaintiffs did allege the element of agreement. Of course, the allegation was conclusory, but so is the allegation of negligence in the Form 9 complaint. What is different about the two situations is the relationship between the allegations and the ordinary baseline appropriate for the case.

Professor A. Benjamin Spencer interprets *Twombly*’s standard to require allegations of “objective facts” that raise “a presumption of impropriety.”73 He argues that factual allegations create such a presumption when they “present a scenario that, if true, suggests wrongdoing,” and he contrasts plausible allegations with those for which lawful and unlawful explanations are “equally possible.”74 Professor Spencer is correct to associate plausibility with allegations of fact that differ from an ordinary state of

71 The plaintiff’s description must not only differ from the baseline but also correlate with illegal conduct in more than a trivial way. Moreover, it is possible that the requisite degree of correlation might vary to a limited extent with the risks and costs of meritless litigation and the effectiveness of judicial case management for the particular type of suit.


74 See id. (also noting that “the facts presented in *Twombly* were equivocal and thus the presumption of propriety was not overcome”).
affairs, but his formulation of the requirement in terms of a “presumption of impropriety” is confusing. For example, he does not explain clearly how to tell whether a set of allegations “suggests wrongdoing” strongly enough to meet the standard.\footnote{He suggests that dismissal is appropriate when lawful and unlawful explanations are “equally possible,” which might mean that one should compare the relative strength of different inferences that the allegations support. However, he cannot mean that an inference of illegality must be more likely than an inference of legality. That standard is even stronger than the very strict pleading requirement of the Private Securities Litigation Reform Act. See \textit{Tellabs, Inc. v. Makor Issues & Rights, Ltd.}, 127 S.Ct. 2499, 2510 (2007) (holding that the PSLRA requires the plaintiff to allege sufficient facts so that “a reasonable person would deem the inference of [scienter] cogent and at least as compelling as any opposing inference one could draw from the facts alleged”).} My baseline idea clarifies the matter. What \textit{Twombly} requires are allegations describing a state of affairs that differs from a baseline of normality, and in a way that supports a stronger correlation to wrongdoing than for baseline conduct.\footnote{Professor Richard Epstein offers an interpretation of \textit{Twombly} that, while useful in some respects, also misses the mark. He argues that the result makes sense given the fact that the plaintiffs relied on “public information, easily assembled and widely available, and rebutted by other public information.” Epstein, \textit{supra} note 59, at 5, 13–15. I agree with Professor Epstein that the defendants’ conduct appears perfectly competitive when all the public information about the telecommunications market is considered, but I disagree that this turns the motion to dismiss into a “mini-summary judgment.” \textit{Id.} at 15. My baseline interpretation of \textit{Twombly} has the advantage of explaining how the Court’s holding extends beyond antitrust cases and cases involving public agency records or other publicly available information. I do agree with Professor Epstein, however, that pleading standards should balance error costs, a point I explore in more detail \textit{infra} in Part IV.}

Understood relative to a baseline, \textit{Twombly}’s plausibility standard should have only a minor impact on notice pleading as a practical matter.\footnote{Hannon, \textit{Note, supra} note 40, at 4 (finding no statistically significant effect on dismissal rates after \textit{Twombly} except in civil rights suits).} Most complaints that pass muster under notice pleading should also pass muster under plausibility pleading. They describe situations that are unusual relative to the most sensible baseline of normality and correlate significantly with unlawful conduct.\footnote{The Court distinguishes between “factually neutral” and “factually suggestive” allegations, a distinction that fits the idea of a legal baseline relative to which allegations could be neutral or suggestive. \textit{Twombly}, 127 S.Ct. at 1966 n.5.}
The major impact of *Twombly*—and I believe the reason critics are so concerned—is not so much what it says about the pleading standard, but rather what it says about discovery costs and settlement leverage and the ineffectiveness of case management more generally. This portion of the opinion, coupled with the Court’s explicit endorsement of case screening as an appropriate pleading function, might be interpreted by overburdened district judges as an invitation to use the vague “plausibility” standard aggressively, notwithstanding language in *Twombly* and later in *Erickson* to the contrary. In other words, critics fear that *Twombly* gives too much latitude to district judges, eager to screen cases and likely to read the opinion as granting permission to do so.

This fear is not unfounded given the judicial track record with strict pleading over the past twenty years. In the 1980s, federal judges tightened pleading requirements in antitrust, securities fraud, civil rights, and other cases that they believed had serious frivolous suit problems. The Supreme Court tried to stop this practice with its *Leatherman* decision in 1993 and again with its *Swierkiewicz* decision in 2002. The Court held in both cases that any departure from liberal notice pleading had to be

---

80 *Erickson*, 127 S.Ct. at 2200.
81 See, e.g., Hoffman, *supra* note 4, at 43 (expressing concern about what lower courts might do with *Twombley*). Another possible concern, which is relevant mostly for antitrust cases, is that judges untutored in economic theory will misapply economics at the pleading stage. This is certainly a risk, but it is also a risk at summary judgment and other stages of the litigation where the judge has power to influence the outcome. And it is arguably an even more serious problem when the jury deliberates. If we tolerate this risk at later stages, it is not clear why we should not tolerate it at the pleading stage as well.
accomplished through the formal rulemaking process or through legislation—and not by judge-made common law rules.\textsuperscript{85} However, lower court enthusiasm for heightened pleading persisted despite these holdings.\textsuperscript{86} Concerns about case backlog, high costs, and litigation delay were simply too strong for district judges to give up screening at the pleading stage, and as a result, many found ways to get around \textit{Leatherman} and \textit{Swierkiewicz}.\textsuperscript{87} Given this track record, it is not unreasonable to think that lower courts might push \textit{Twombly}’s vague plausibility standard as far as possible in the direction of more demanding pleading requirements. But aggressive screening through stiff pleading is not what the Supreme Court intended.

III. \textit{Twombly} in Historical Context

It is important to correct two exaggerated historical claims made by some of \textit{Twombly}’s critics.\textsuperscript{88} First, it is wrong to condemn \textit{Twombly}’s plausibility standard for being inconsistent with the language of Rule 8(a)(2) or the intent of the 1938 Advisory Committee. Second, one should be careful about any claim that \textit{Twombly} departs from

\textsuperscript{85} Both \textit{Leatherman} and \textit{Swierkiewicz} affirm \textit{Conley v. Gibson}’s liberal notice pleading standard and hold that the lower courts have no power to create heightened pleading rules on their own. \textit{Leatherman}, supra, at 168–69; \textit{Swierkiewicz}, supra, at 512–14.

\textsuperscript{86} See, e.g., Fairman, supra, note 8, at 1011–59.

\textsuperscript{87} \textit{Id.} at 1062-64. In fact, the Supreme Court itself sent mixed signals in its 1998 \textit{Crawford-El} opinion and its 2005 \textit{Dura Pharmaceuticals} opinion. Dura Pharm., Inc. v. Broudo, 544 U.S. 336, 347 (2005) (holding that Rule 8(a) requires securities fraud class action plaintiffs to allege loss causation with greater particularity); Crawford-El v. Britton, 523 U.S. 574, 598 (1998) (rejecting strict pleading for a Section 1983 complaint, but suggesting that courts might use the motion for a more definite statement or a Rule 7(a) reply to require more factual detail).

\textsuperscript{88} See, e.g., Spencer, supra note 4, at 434–36, 461, 479–80 (making these criticisms); \textit{The Supreme Court 2006 Term, Leading Cases, Pleading Standards}, 121 HArv. L. Rev. 305, 311–12 (2007) (criticizing \textit{Twombly} for inconsistency with the original drafters’ intent). \textit{But see} Campbell, supra note 72 (arguing that a proper understanding of Rule 8(a) and 1938 Advisory Committee intent is consistent with \textit{Twombly}). Also, it is common for critics to describe \textit{Twombly}’s plausibility standard as a throwback to “fact pleading” under the codes, which is tantamount to criticizing it for inconsistency with the 1938 drafters’ intent.
the procedural vision of the Federal Rule drafters. At the core of that vision was a pragmatic commitment to making procedure an efficient means to enforce the substantive law accurately, and Twombly shares that pragmatism.

A. Rule 8(a)(2), Code Pleading, and 1938 Advisory Committee Intent

Charles Clark, Reporter for the original Advisory Committee and chief architect of the 1938 Federal Rules, was an authority on pleading, having published numerous articles and one of the most important books on the subject.89 Pleading reform was one of his key innovations for the new Federal Rules.90

Clark’s views on pleading and those of his contemporaries were developed in response to perceived deficiencies in common law and code pleading.91 These deficiencies had to do mainly with an insistence on technicality unrelated to any sensible pleading function.92 Although the code system improved on the common law,93 nineteenth century judges applied the code rules in a hyper-technical fashion, insisting on “strict and logical accuracy” and drawing hopeless distinctions among allegations of ultimate fact, legal conclusions, and evidentiary facts.94

89 CHARLES E. CLARK, HANDBOOK OF THE LAW OF CODE PLEADING (1928) (hereinafter cited as Clark, CODE PLEADING).
91 See 5A WRIGHT & MILLER, supra note 8, § 1202 (suggesting that the system of common law pleading was “excruciatingly slow, expensive, and unworkable. The system was better calculated to vindicate highly technical rules of pleading than it was to dispense justice.”).
92 See CLARK, CODE PLEADING, supra note 88.
93 Most code pleading provisions were similar to the New York Field Code section, which required that the complaint contain “a plain and concise statement of the facts constituting each cause of action without unnecessary repetition.” Id. at 138.
94 CLARK, CODE PLEADING, supra note 88, at 159. Clark gave an example of a bicyclist who was injured when she struck a hole in a negligently maintained street. The bicyclist’s suit was dismissed because, while she alleged the hole and the failure to maintain the street, she stated only that her bicycle “struck said defective, unsafe, and out of repair street,” not that it struck the hole in the street. Id. at 158.
In Clark’s view, a complaint, while it should include some facts beyond a mere conclusion of liability, need only state enough “as will isolate [the case] from all others, so that the parties and the court will know what is the matter in dispute, the case can be routed through the court processes to the proper method of trial and disposition and the judgment will be res adjudicata, so that the same matter cannot again be litigated.” Clark believed that merits screening should take place after discovery, at summary judgment in some cases and at trial in most.

There is no question that the Twombly complaint met Clark’s notice pleading standard. Still, one must be careful about accusing the Court of deviating radically from Rule 8(a)(2). The Rule refers to “a short and plain statement of the claim showing that the pleader is entitled to relief” only in order to avoid the code’s “facts constituting a cause of action” formulation. It does not refer to notice pleading explicitly. The term “notice pleading” was in common use at the time to refer to the most liberal pleading standard, so if notice pleading were intended, one might have expected the text of the Rule or the Committee Note to say so.

---

95 Charles E. Clark, The Complaint in Code Pleading, 35 Yale L. J. 259, 271 (1926) (noting that a complaint should be structured around a legal theory, for “otherwise it would be a meaningless jumble”); Smith, supra note 89, at 917–18 (noting that Clark demanded more than a conclusory allegation of negligence).
96 Clark, supra note 16, at 316; accord Charles E. Clark, Simplified Pleading, 2 F.R.D. 456, 460–61 (1942) (stating that the pleadings need only give notice “of the general nature of the case and the circumstances or events on which it is based”).
97 See Marcus, supra note 4, at 439–40.
98 As for the existence of an agreement, the defendants had the information to know whether that allegation was true and could easily deny it.
99 5A Wright & Miller, supra note 8, § 1202, at 87.
100 See Clark, Code Pleading, supra note 88, at 29–30 (distinguishing “issue pleading” at common law, “fact pleading” under the codes, and a new approach, “notice pleading,” that “is now urged for general adoption”); id. at 163 (observing that “notice pleading” requires only “a very general reference to the happening out of which the case arose”).
Moreover, not everyone agreed that the federal standard should be very liberal notice pleading. During the years leading up to adoption of the Federal Rules, and continuing for almost twenty years afterward, jurists and politicians sharply divided on the pleading issue, some insisting that specific pleading was essential to properly framing the lawsuit and rendering it manageable.\(^{101}\) It should be no surprise then that when Charles Clark explained the new pleading rule to an American Bar Association conference held in Cleveland, Ohio, he made a point of mentioning that different judges might vary to some extent in the detail they required,\(^{102}\) and that his very liberal views might not be shared by all Advisory Committee members.\(^{103}\)

What is clear is that the Committee intended to reject the code version of fact pleading, and \emph{Twombly}’s pleading standard is a far cry from that. The \emph{Twombly} Court does not insist that every fact essential to liability be alleged clearly and precisely, nor does it insist that the complaint contain only allegations of ultimate fact rather than legal


\(^{102}\) See \emph{AM. BAR ASS’N, RULES OF CIVIL PROCEDURE FOR THE DIST. COURTS OF THE UNITED STATES WITH NOTES AS PREPARED UNDER THE DIRECTION OF THE ADVISORY COMM. AND PROCEEDINGS OF THE INSTITUTE ON FED. RULES OF CIVIL PROCEDURE, CLEVELAND, OHIO JULY 21, 22, 23 1938, at 220 (1938):

\begin{quote}
 I think there is no question that the rules are based on the theory of a rather general form of pleading. Whether they go as far as I believe they do, you may, of course, question. I want to point out that in this field there is considerable room for variations between courts. After all, we made a generalized statement in the rules, a short and simple statement along the lines in which we believed, and to one judge that may require much more than it does to others.
\end{quote}

Clark also explained that a judge who believed more factual detail was desirable could grant a motion for a more definite statement under Rule 12(e), although he counseled against excessive use of this device. \emph{Id.} at 223.

\(^{103}\) \emph{Id.} at 220; see also Smith, \emph{supra} note 90, at 917.
conclusions or evidence. The plaintiffs’ lawyer is free to use whatever method works to explain why the existence of an agreement is plausible.

Rule 8(a) came to stand clearly for notice pleading mainly through judicial interpretations of the Rule. After Charles Clark was appointed to the Second Circuit Court of Appeals in 1939, he worked hard to establish Rule 8(a)(2) as a notice pleading rule and resisted efforts to construe it more strictly. But it is the Supreme Court’s 1957 opinion in *Conley v. Gibson* and its famous “no set of facts” language that firmly established notice pleading as the Rule 8(a) standard. With this opinion, the Court put to rest, at least for a time, the two decade controversy over the proper pleading standard.

If notice pleading is best understood as a judicial interpretation of Rule 8(a)(2), then it is hardly illegitimate for the Court to revisit this earlier interpretation and qualify

---

104 At one point, however, the Court does suggest distinctions between “the conclusory and the factual” and between “the factually neutral and the factually suggestive,” but these distinctions are very different from the code classifications. Bell Atlantic Corp. v. Twombly, 127 S.Ct. 1955, 1966 n.5 (2007).

105 See, e.g., Dioguardi v. Durning, 139 F.2d 774, 775 (2d Cir. 1944); Marcus, supra note 4, at 445.

106 *Conley v. Gibson*, 355 U.S. 41 (1957). The complaint in *Conley* stated lots of facts and would have satisfied any pleading standard. Justice Black wrote the opinion for the Court, and as a former trial lawyer he was a strong proponent of jury trial and disliked procedural niceties. It is very likely that his strong language endorsing notice pleading was meant to ensure that as many cases as possible made it to the jury. See Emily Sherwin, *The Story of Conley: Precedent by Accident*, in CIVIL PROCEDURE STORIES 295, 308 (2d ed. 2008).

107 The Court also stated that “the Federal Rules of Civil Procedure do not require a claimant to set out in detail the facts upon which he bases his claim. To the contrary, all the Rules require is a ‘short and plain statement of the claim’ that will give the defendant fair notice of what the plaintiff’s claim is and the grounds upon which it rests.” *Conley*, 355 U.S. at 45.

108 See Richard L. Marcus, *The Puzzling Persistence of Pleading Practice*, 76 TEX. L REV. 1749, 1750 (1998). That debate was to heat up again in the 1980s, when concerns about a litigation crisis in the federal courts grew more intense. See *supra* notes 82-87 and accompanying text.
or revise it. The Court must have a good reason to do so, of course, and that reason must fit the idea that the Court is applying a Federal Rule, even if rather loosely, rather than simply making common law. As I explain in the following section, the reason to revisit the interpretive question in *Twombly* has to do with responding to new litigation conditions, a reason that fits the principle of dynamic flexibility underlying the original Rules.

The doctrine of *stare decisis* is also relevant, especially in view of *Conley’s* fifty years of dominance. However, the values of predictability, reliance, equality, and stability that underlie *stare decisis* are not strongly implicated by pleading rules. It

---

109 Some might argue that events in 1955 indicate the Committee’s intent to adopt the most liberal version of notice pleading for Rule 8(a)(2), but that would be a mistake. In 1955, the Advisory Committee, with Charles Clark still acting as Reporter, rejected an effort, spearheaded by the Ninth Circuit Judicial Conference, to amend Rule 8(a)(2) by adding the phrase “facts constituting a cause of action,” so the amended Rule would read: “(2) a short and plain statement of the claim showing that the pleader is entitled to relief, which statement shall contain the facts constituting a cause of action.” See REPORT OF PROPOSED AMENDMENTS TO THE RULES OF CIVIL PROCEDURE FOR THE UNITED STATES DIST. COURTS, 18–19 (1955), available at http://www.uscourts.gov/rules/Reports/CV10-1955.pdf. The proposal was triggered by growing concern about the high costs of complex antitrust litigation in the 1950s. The problem with using this decision to infer committee endorsement of notice pleading is that the Committee’s two page explanation does not support the conclusion. Nowhere does it expressly refer to “notice pleading” by name. Rather, it defends Rule 8(a)(2) on the ground that it eliminates the technical skirmishing of code pleading by allowing only a general statement, yet still requires “the pleader to disclose adequate information as the basis for his claim for relief as distinguished from a bare averment that he wants relief and is entitled to it.” *Id.* at 19. Responding to criticisms of *Dioguardi, supra*, the Committee stressed that the *Dioguardi* decision “was not based on any holding that a pleader is not required to supply information disclosing a ground for relief.” *Id.*

110 Especially as plaintiffs are routinely given a chance to amend their complaints to meet the applicable pleading requirement. Professor Spencer makes much of the *stare decisis* constraint, likening interpretation of the Federal Rules to interpretation of a statute. See Spencer, *supra* note 4, at 461–63. However, the Federal Rules are not statutes and are generally interpreted more generously, even though the precise scope of interpretive flexibility is a matter of some dispute. See 4 WRIGHT & MILLER, *supra* note 8, § 1029 (stating that the FRCP are to be construed generously “to further the cause of justice” and noting that this liberal interpretive principle is endorsed explicitly by Rule 1). Compare Joseph P. Bauer, *Schiavone: An Un-Fortune-ate Illustration of the Supreme Court’s Role as Interpreter of the Federal Rules of Civil Procedure*, 63 NOTRE DAME L. REV. 720, 720 (1988) (arguing that the court’s role is more expansive when interpreting a FRCP than when interpreting a statute) with Catherine T. Struve, *The Paradox of Delegation: Interpreting the Federal Rules of Civil Procedure*, 150 U. PA. L. REV. 1099, 1119.
might not be wise policy for judges to require stricter pleading outside of the formal rulemaking process, a point I discuss later, but that is a very different objection than one charging a violation of *stare decisis*.

**B. The Vision of the Federal Rule Drafters**

Some critics of *Twombly* argue that the shift from notice to case screening marks a sharp break from the vision of the original Federal Rule drafters. That vision, they argue, is inconsistent with case screening at the pleading stage, since it favors decisions on a complete factual record developed through discovery and usually tested at trial.\(^{111}\) The problem with this argument is that it describes the drafters’ vision at the wrong level of generality.

It is true that Charles Clark eschewed case screening at the pleading stage and that the drafters contemplated decisions on the basis of evidence rather than allegations. However, the 1938 drafters were pragmatists first and foremost.\(^{112}\) They supported liberal pleading and evidence-based merits decisions as a pragmatic not a natural law ideal. They thought a procedural system with these elements would work much better than the common law and code systems they inherited.\(^{113}\) “Work much better” meant that the procedural system would enforce the substantive law more effectively and “without undue waste or friction or consumption of fuel.”\(^{114}\)

This pragmatic vision—that procedural rules should be designed instrumentally to enforce the substantive law fairly and efficiently given the way those rules actually work

---

\(^{111}\) This vision is sometimes referred to as the “liberal ethos.” *See* Marcus, *supra* note 4, at 439.

\(^{112}\) For a detailed account of the 1938 drafters’ pragmatism, *see* Bone, *supra* note 7, at 78–98.

\(^{113}\) *Id.*

in practice—operates at a deeper level than preferences for discovery and trial. The latter preferences just reflect structural choices. The drafters’ pragmatic vision supplies the underlying normative support.

The 1938 drafters’ choice of pleading rules closely fit their pragmatic vision.115 Simplified pleading accomplished the notice-giving function at minimum cost and allowed cases to proceed through discovery and on to trial where they could be decided based on what actually happened rather than on legal technicalities.116 The 1938 drafters’ choices also made pragmatic sense in light of the way litigation actually worked in the early twentieth century.117 Many cases were small, two-party affairs; the large, complex case of today was relatively unknown.118 Many lawyers were local practitioners who dealt with one another and with local judges repeatedly and had incentives to

115 See, e.g., Bone, supra note 7, at 101 n. 345; Subrin, supra note 101.

116 See id. at 508, 520; see also Daniel J. Meador, A Perspective on Change in the Litigation System, 49 ALA. L. REV. 7, 8–9 (1997).

117 See Bone, supra note 7, at 101 n. 345; Sherwin, supra, at 9–18.
maintain a reputation for reasonableness. The era of huge law firms with national practices and a much weaker connection to local communities had not yet arrived. To be sure, the decades of the 1920s and 1930s had their own litigation crisis. But that crisis involved an explosion of litigation produced by the increased use of automobiles, and automobile accident cases are rather simple affairs, usually involving only a few parties and fairly straightforward factual and legal issues.

It would have made sense in this earlier world to assume relatively manageable discovery and trial costs. Today’s world of litigation is very different. There are many more large and complex lawsuits with high stakes, large law firms and lawyers practicing nationwide who have much weaker incentives to build local reputations, and a much wider range of materials that can be targeted in discovery, including potentially massive electronic records. These changes have fueled a powerful perception of serious litigation cost and frivolous suit problems.

Long before the Twombly decision, lower federal courts in the 1980s responded to this sense of crisis by tightening up on pleading specificity requirements. And they continued in this vein despite Supreme Court decisions to the contrary in Leatherman and

---

119 When discussing the merits of a uniform code of pleading at the 1913 American Bar Association meeting, J. Hansell Merrill noted that few lawyers practiced interstate compared to the “much larger number whose practice is confined to a single state.” Report of the Committee on Uniform Judicial Procedure, 38 A.B.A. Rep. 44 (1913).

120 The automobile-litigation crisis was felt in many different states and even prompted proposals to create administrative solutions modeled on the then-recently enacted workmen’s compensation statutes. See Compensation for Automobile Accidents: A Symposium, 32 Colum. L. Rev. 785 (1932); see generally Robert G. Bone, Procedural Reform in a Local Context: The Massachusetts Supreme Judicial Court and the Federal Rule Model, in The History of the Law in Massachusetts: The Supreme Judicial Court 1692, 403–04 (R. Osgood ed. 1992) (describing the automobile-litigation crisis in Massachusetts).

121 See Meador, supra note 118, at 7, 8, 12–14.

122 See supra notes 82–87 and accompanying text.
This history is important because it shows just how strong the pressure is to rely on pleading specificity to address litigation problems. I am confident that the Federal Rule drafters, as procedural pragmatists, would also entertain the possibility of stricter pleading as a solution if they shared the same perceptions. Charles Clark emphasized the evolving nature of the Federal Rules and the importance of revising those Rules to keep pace with changing litigation conditions. Clark and his colleagues ultimately might reject strict pleading on pragmatic grounds, but they would not rule it out categorically any more than lawyers and judges do today.

In fact, the original rule drafters adopted stricter pleading for fraud in Rule 9(b), and they did so in part because of the harm to defendant’s reputation and the cost to settled transactions that ungrounded fraud allegations could create. It is true that these stricter pleading rules were well established at common law and in equity, but the drafters did not have to incorporate them if they believed that the traditional rules were as misguided as the code’s fact pleading approach more generally. That they thought it worthwhile to include 9(b) as an exception to 8(a) is a sign that they accepted as a

---

123 See id.
124 Clark, supra note 16, at 304.
125 In fact, Charles Clark seems to have tempered his more extreme notice pleading views in the late 1950s. See Marcus, supra note 4, at 451 & n.113.
126 Policy justifications for Rule 9(b)’s heightened pleading standard include preventing unnecessary harm to the defendant’s reputation, screening frivolous suits, protecting settled transactions, and preventing “fishing expeditions.” See 5A WRIGHT & MILLER, supra note 8, § 1296, at 31–47. Sometimes the rationale is expressed in terms of providing “fair notice” to defendants, see id. at 39–45, but Rule 9(b) is not required for this purpose since Rule 8(a) is supposed to do the notice-giving work. Furthermore, any adverse effect from a fraud allegation is a problem only if the allegation turns out to be wrong and the suit meritless. If the defendant is actually guilty of fraud, the reputation effect is deserved and any transaction based on the fraudulent representation should be rescinded. Thus, the stricter pleading standard for fraud must have to do with preventing unfounded allegations.
pleading objective the screening of cases when the frivolous suit risk was serious and harmful enough.\footnote{To be sure, the drafters, concerned about information access, included an exception in Rule 9(b) for allegations of “malice, intent knowledge, and other condition of mind.” Fed. R. Civ. P. 9(b). This exception indicates that the drafters, if they lived today, would take account of information access, as they should, when designing case-screening rules. How they would handle this factor depends on the pragmatic balance in light of the perceived severity of the frivolous suit problem. My point is simply that the drafters were alert to these problems and willing to use strict pleading to address them.}

I am not suggesting that procedural problems should be analyzed today using the same pragmatic approach that the original drafters employed. As Part IV demonstrates, I favor a more systematic and rigorous analysis attentive to the differences among normative theories and the availability of statistically reliable empirical support. Simply put, my point is that the 

Twombly
decision could have been endorsed by an early twentieth century pragmatist living today and concerned about the litigation problems of the twenty-first century.

IV. \textbf{Optimal Pleading and Case-Screening Rules: A Normative Analysis}

The previous discussion showed that 

Twombly’s plausibility standard, properly understood, marks only a small departure from existing pleading standards and one that is consistent with Rule 8(a)(2) and the pragmatic vision of the original Federal Rule drafters. Still, the question remains whether plausibility is justified on normative grounds, and more generally, whether stricter pleading is an optimal way to screen undesirable lawsuits. This Part discusses those issues.

These are particularly important issues to address today, especially given what the 

Twombly Court says about the efficacy of case-specific discretion. In what may be the most important part of the 

Twombly opinion—perhaps even more important than the discussion of \textit{Conley}—the Court expresses serious doubts about how well trial judges can
handle discovery problems and meritless suits using discretionary case management tools. The Court first notes that “discovery expense will push cost-conscious defendants to settle even anemic cases”\(^{128}\) and then points out that trial judges lack the information necessary to manage the problem effectively in complex antitrust cases,\(^{129}\) especially as weak and frivolous suits are likely to settle before the summary judgment stage.\(^{130}\)

This marks a significant change of course.\(^{131}\) Until *Twombly*, the Supreme Court had never openly and directly questioned the effectiveness of judicial discretion in managing litigation problems during the pre-trial phase.\(^{132}\) The Court in the past either approved trial judge discretion or deferred to it without evaluative comment.\(^{133}\)


\(^{129}\) *Id.* at 1967 & n.6.

\(^{130}\) *Id.*

\(^{131}\) Justice Stevens’s dissenting opinion is a good example of the attitude that prevailed before *Twombly*. See *id.* at 1987 n.13:

The Court vastly underestimates a district court’s case-management arsenal. . . . In short, the Federal Rules contemplate that pretrial matters will be settled through a flexible process of give and take, of proffers, stipulations, and stonewalls, not by having trial judges screen allegations for their plausibility *vel non* without requiring an answer from the defendant.

\(^{132}\) The Court has restricted trial judge discretion in some special areas unrelated to pleading and case management out of fear that judges might use their discretion to make decisions that expedite resolution but adversely affect the public interest. For example, in order to avoid arresting the development of constitutional law, the Court requires district judges in Section 1983 cases to decide the issue of constitutional violation before qualified immunity, even when a straightforward qualified immunity determination would easily dispose of the case. See Saucier v. Katz, 533 U.S. 194, 201, 210 (2001).

same is true for the federal courts of appeal. And the Advisory Committee on Civil Rules has operated for decades under the assumption that discretionary case management is an effective tool for dealing with litigation abuse.

The *Twombly* Court’s skepticism is in fact well justified. Serious litigation problems should not be left to trial judge discretion as much as they are today. Judges face information and other constraints that impair their ability to manage optimally, especially in the highly strategic environment of litigation. But if deferring to case-specific discretion is not the solution, then the policy and regulatory issues must be confronted directly at the rulemaking stage.

The following discussion addresses these issues for pleading and case screening. Section A first considers whether the *Twombly* Court’s thin plausibility standard can be justified by a process-based theory of fairness as reason-giving. Section B then analyzes pleading standards stricter than thin plausibility, designed to screen undesirable suits. Section B evaluates these stricter standards from an outcome-based perspective and

---

134 See, e.g., Mass. Inst. of Tech. v. Abacus Software, 462 F.3d 1344, 1367 (Fed. Cir. 2006) (deferring to the district judge’s case management decisions because “a judge's discretion is at its broadest on matters of trial management”); Sims v. ANR Freight Sys., 77 F.3d 846, 848 (5th Cir. 1996) (lauding the district judge’s “hard work and effective case management” skills even though his decision in the instant case caused a structural error); Jaquette v. Black Hawk County, 710 F.2d 455, 463 (8th Cir. 1983) (“In almost all cases the key to avoiding excessive costs and delay is early and stringent judicial management of the case.”). A caveat is in order. My claim that neither the Supreme Court nor the courts of appeals ever directly questioned the efficacy of trial judge discretion in managing cases before *Twombly* is based on extensive searches in the LEXIS and WESTLAW databases.


136 For a thorough analysis of these and other constraints, see Bone, *supra* note 135.
compares them to other screening tools. Finally, Section C sketches the outlines of an optimal case-screening approach.

A. *Twombly*'s Plausibility Standard from a Process-Based Perspective

Broadly speaking, there are two different normative approaches to analyzing procedural issues: process-based and outcome-based. A process-based approach evaluates a procedural rule by how it treats litigants independent of its consequences for outcome quality, while an outcome-based approach evaluates a rule by its effect on the quality of litigation outcomes.  

The following discussion outlines a process-based justification for *Twombly*'s plausibility standard in terms of a balance between fairness to defendants and fairness to plaintiffs. When proceduralists discuss pleading standards, they tend to assume that fairness applies just to plaintiffs and that any pleading standard stricter than liberal notice pleading can be justified only on efficiency grounds. This is a mistake. Fairness applies to both parties. As the following discussion demonstrates, fairness in the

---

137 For more on the distinction between process-based and outcome-based theories, see Robert G. Bone, *Agreeing to Fair Process: The Problem With Contractarian Theories of Procedural Fairness*, 83 B.U. L. REV. 485, 508–16 (2003). A process-based approach is nonconsequentialist insofar as it evaluates procedure without regard to its consequences, whereas an outcome-based approach is consequentialist. However, an outcome-based approach can be tied to a nonconsequentialist metric for evaluating outcomes, as is the case for an outcome-based approach that relies on a rights-based theory of the substantive law. See *infra* notes 178–184 and accompanying text. For the distinction between consequentialist and nonconsequentialist theories in general, see *Consequentialism and Its Critics* 1–13 (Samuel Scheffler ed. 1988). Another point is worth mentioning. The analysis in this section proceeds by deriving pleading and case-screening rules from broad normative theories of procedure. It is possible, however, to combine this analysis with a constructivist approach that develops procedural principles from an attractive normative account of settled practice. For an example of a constructivist approach, see Larry B. Solum, *Procedural Justice*, 78 S. CAL. L. REV. 181 (2004); see also Bone, *supra* note 9, at 943–47 (arguing that the Advisory Committee should use a constructivist approach and illustrating it with the example of the small claim class action).

138 See *infra* note 168.
pleading context has something to say not only about a plaintiff’s ability to sue, but also about when the defendant must respond to the plaintiff’s demands.

1. **Fairness to Defendants**

   We begin by asking whether fairness to the defendant requires some degree of pleading specificity. To get at this question, imagine a procedural system that has no pleadings at all. In this system, the plaintiff merely informs the defendant that he is being sued, perhaps by serving a summons, and then waits for the defendant to use discovery to find out why the suit has been brought. If the defendant does not respond within a specified period of time, he suffers a default judgment.

   One might object to such a system as inefficient because it invites too many weak and frivolous suits and increases the social costs of litigation. I discuss this outcome-based objection in Section B. below. But it is not clear that the prospect of bad outcomes exhausts all the grounds for objection. At least at first glance, there seems to be something unfair about a plaintiff forcing a defendant to shoulder the burden of litigation without giving the defendant any reason why he should. Explaining this intuition, however, is no easy matter.

   It is possible that the intuition is prompted by a concern about how easy it is for a plaintiff in a system without pleadings to file a frivolous suit simply to spite the defendant or pressure an unjustified settlement. If this is the crux of the problem, then the source of unfairness resides in the moral culpability of intentionally filing a meritless suit. An intentional meritless filing is similar to an intentional tort and morally blameworthy for many of the same reasons.

---

139 See infra Part IV.B.
Yet it is not clear that this explanation captures the root of the problem either.

The problem is not so much that a lawsuit might be frivolous, but rather that the defendant must bear the burden of response without being given any reason why he should do so.140 Still, the question remains: What makes this unfair when the defendant can uncover the reason for himself by conducting discovery?

The standard process-based theory of participation rights in not terribly helpful in this regard.141 According to this theory, respect for the dignity and autonomy of persons significantly affected by litigation requires that the state give each litigant a chance to tell her story, and perhaps also a chance to control the major incidents of litigation, regardless of the effect on outcome quality.142 Applying this theory to pleading, one might argue that a system without pleading does not give the defendant an adequate opportunity to

---

140 Unless, of course, there are extrinsic circumstances that shed light on the probable reason for suit.

141 See, e.g., JERRY L. MASHAW, DUE PROCESS IN THE ADMINISTRATIVE STATE 158–253 (1985). Another version of the process-based approach anchors the participation right in legitimacy rather than dignity. See Solum, supra note 137. It is worth noting that much of the discussion of intrinsic process value in civil procedure draws on the “procedural justice” literature. This literature demonstrates through psychological experiments that people are more likely to feel they have been treated fairly by the process and the outcome, even an adverse outcome, when they are given an opportunity to participate. See, e.g., E. ALLAN LIND & TOM R. TYLER, THE SOCIAL PSYCHOLOGY OF PROCEDURAL JUSTICE 26–40, 61-83, 93–127 (1988) (discussing these empirical studies); E. Allan Lind et al., In the Eye of the Beholder: Tort Litigants' Evaluations of Their Experiences in the Civil Justice System, 24 LAW & SOC'Y REV. 953, 967–87 (1990) (analyzing the relationship between tort litigants' fairness judgments and various objective and subjective factors). However, these psychological responses have no normative significance on their own. They might be relevant for a moral conventionalist if they counted as evidence of generally shared moral beliefs. However, psychological responses like these must be subjected to critical reflection before they can count as well-considered beliefs. In general, those who rely on procedural justice studies to support conclusions about fairness tend to assume that procedures are fair precisely because of the positive psychological reactions they elicit. But this is mistaken. If positive feelings are what count, then all positive feelings should be counted equally (even positive feelings about ice cream), but this quickly ends up sliding into some form of aggregate utilitarianism. See Bone, supra note 137, at 505–07.

142 I have argued elsewhere that the dignitary value underlies the strong right to a personal “day in court” that figures prominently in American civil procedure. Robert G. Bone, Rethinking the “Day in Court” Ideal and Nonparty Preclusion, 67 N.Y.U. L. REV. 193 (1992). I have also raised questions about the coherence of this theory. See id. at 279-85..
The problem with this argument, however, is that the defendant *does* have an opportunity to participate. While the plaintiff need not give reasons, she must notify the defendant that he is being sued, and once notified the defendant can take the plaintiff’s deposition, serve interrogatories, or use other discovery devices to ascertain why the suit has been brought.

Perhaps the need to use discovery burdens the defendant’s right to participate in a morally impermissible way. But that hardly seems sensible. To be sure, effective participation requires that the defendant incur litigation costs to find out why he is being sued, but the necessity of incurring litigation costs is not alone enough to render procedures morally objectionable. Moral concerns can arise when litigants lack the wealth necessary to finance litigation, but the sense of unfairness we are exploring is not confined to parties who have little wealth.

We come closer to the heart of the matter when we focus on the unchecked power a system without pleadings gives plaintiffs. This suggests that the objection might have to do with the unfairness of arbitrary state action in general. Many scholars of procedure and administrative law argue that respect for persons demands that the state act rationally and non-arbitrarily when adjudicating the entitlements or obligations of individuals.\(^\text{143}\)

\[\text{In constitutional due process terms, we might say that the notice is not “reasonable” or the participation not “meaningful.” See } \text{Armstrong v. Manzo, 380 U.S. 545, 552 (1965) (due process requires a hearing “at a meaningful time and in a meaningful manner”); Mullane v. Central Hanover Bank & Trust Co., 339 U.S. 306 (1950) (due process requires notice “reasonably calculated to apprise interested parties of the pendency of the action and afford them an opportunity to present their objections”).}\]

\[\text{See J. Mashaw, supra note 141, at 189, 198–99, 201–04 (deriving a requirement of comprehensibility or “thin rationality” from respect for dignity); Laurence H. Tribe, American Constitutional Law 663–64, 666 (2d ed. 1988) (arguing that Due Process provides an “institutional check on arbitrary government action” and identifying an intrinsic dignitary value to a hearing “as an expression of the rule of law, regarded here as the antithesis of power wielded without accountability to those on whom it focuses.”); Richard B. Saphire, Specifying Due Process Values: Toward a More...}\]
This means that the state must have good reasons for its actions.\textsuperscript{145} It is true that in our hypothetical system, the plaintiff rather than a state agent imposes the burden. But the state empowers the plaintiff, and state judges coerce the defendant with the threat of default.

Still, it is one thing for the plaintiff to have a good reason and quite a different thing for the plaintiff to inform the defendant of that reason. If fairness to the defendant has anything to say about pleading standards, it must be because there is a moral obligation not only to \textit{have} reasons, but also to \textit{provide} those reasons to the person affected. If such an obligation exists, it would follow that a procedural system should require some kind of initiating document, such as a complaint, that gives the defendant good reasons for bearing the burden of response. Moreover, the allegations of the complaint would have to do more than simply tell a story; they would have to tell a story that suggests in some way a legitimate ground for forcing the defendant into court. This means that the allegations should describe a state of affairs that differs from the ordinary baseline of acceptable conduct.\textsuperscript{146} Otherwise, plaintiffs could force defendants into court for engaging in activities that are well within the scope of perfectly appropriate conduct.

The question, however, is whether the state has a moral obligation to give reasons when it imposes substantial burdens on individuals. The answer to this question is


\textsuperscript{145} This is not the place to explore with care what constitutes a “non-arbitrary” decision. I employ the term in a loose way to refer to decisions that are backed by good reasons. For example, decision by random lottery is non-arbitrary by my definition as long as the decisionmaker has a good reason to use a lottery.

\textsuperscript{146} \textit{See supra} notes 59–71 and accompanying text (describing the baseline idea).
somewhat complicated. Many commentators argue that a general obligation exists, and they justify it in two different ways: either as a primary moral obligation rooted in the principle of respect for persons, or as a prudential obligation justified instrumentally for its value in helping to ensure that the state acts non-arbitrarily. The problem is that there are many situations in which the state does not have to give reasons. For example, juries render verdicts without giving reasons; judges rule on admissibility of evidence without giving reasons, and judges sometimes sentence criminal defendants without publicly justifying the sentences they choose.

But filing a lawsuit is different. The admissibility and sentencing examples involve decisions by an official (the judge) whose authority arguably derives from her official role rather than from the reason she gives, and who is more likely than a private party to be motivated by a concern for the public interest. Moreover, when the judge makes a decision without giving a reason, she often does so against the background of

---

147 See L. H. Tribe, supra note 144, at 666, 743-44 (noting that “the right to be heard from, and the right to be told why, … express the elementary idea that to be a person, rather than a thing, is at least to be consulted about what is done with one” and arguing that a due process government obligation to provide an explanation can be based on a dignitary value); Frank I. Michelman, Formal and Associational Aims, in DUE PROCESS, NOMOS XVIII, at 126, 126–27 (J. Pennock & J. Chapman eds. 1977) (arguing that “explanatory procedures,” including the giving of reasons for a decision, are “responsive to demands for revelation and participation” and can have value even when there is no opportunity to challenge the decision); Edmund Pincoffs, Due Process, Fraternity, and a Kantian Injunction, in DUE PROCESS, NOMOS XVIII, at 172, 175–79 (J. Pennock & J. Chapman eds. 1977) (deriving a moral right to reasons from the Kantian categorical imperative not to treat a person as a mere means to an end); Saphire, supra note 144, at 162-66 (arguing that due process protects inherent dignitary values of “revelation, explanation, and participation”); Summers, supra note 144, at 26-27 (arguing that the process value of procedural rationality includes giving reasons, an obligation that respects persons as rational beings).

148 See, e.g., Goldberg v. Kelly, 397 U.S. 254 (1970). This is, of course, based on the assumption that it is more difficult for an actor to act arbitrarily when she must justify her actions publicly. It is also worth mentioning that non-arbitrariness and reason-giving can be justified on outcome-based grounds, for example, as instrumental to the accurate enforcement of rights.


150 See id. at 658.
arguments pro and con or under conditions where a plausible reason or reasons can be inferred. And when the decision is completely crazy (say sentencing a petty thief to life imprisonment), it is usually subject to review and reversal.

None of these fairness-promoting features is present when a plaintiff files a lawsuit. For one thing, the filing decision is delegated to a private party, who does not act in an official role and whose motivations are personal. Without a complaint requirement, there is absolutely nothing to signal the defendant what the reason for suit might be or anything to check an impermissible filing.  

To be sure, jurors render verdicts without giving reasons, but jurors do so as agents of the court. Moreover, a jury verdict is preceded by evidentiary submissions and arguments that make it possible to infer plausible reasons, and when there is no plausible reason, the verdict can be overturned through a motion for a judgment notwithstanding the verdict or new trial, or by appeal. In short, a procedural system without pleadings is tantamount to a naked delegation of state power to private individuals.

---

151 This last point, the existence of checks on impermissible decisionmaking opens up the possibility that the fairness of a procedural system that allows filings without reasons might depend on how effectively the system screens impermissible filings in other ways.

152 See, e.g., Barzelis v. Kulikowski, 418 F.2d 869 (9th Cir. 1969) (“A jury, however, does not have to give reasons for what it does.”).

153 Another difference is that in the examples of no reason giving there are sound institutional grounds for doing without reasons. One can argue, for example, that the practice of jury verdicts, evidentiary rulings, or criminal sentencing works much better on balance without an obligation to give reasons. Therefore, even if there is a fairness value at stake in the abstract, it should not count because giving reasons is inconsistent with the effective functioning of the practice in question or it is outweighed by the benefits of decisionmaking without an obligation to state reasons publicly. By contrast, it is not clear why the practice of filing lawsuits would work better in general without an obligation to give good reasons for forcing the defendant to respond. Cf. Schauer, supra note 149, at 657–58 (arguing that reason-giving as an element of institutional design should be seen as contingent rather than necessary).

154 It is worth mentioning, however, that the United States Supreme Court has held that the Fourteenth Amendment’s Due Process Clause does not prohibit a state from commencing a criminal prosecution without probable cause. Albright v. Oliver, 510 U.S. 266 (1994). At first glance, this decision
A fairness rationale based on an obligation to give reasons is different from the usual justification for notice pleading. The usual justification evaluates the sufficiency of notice in terms of what is needed to make other elements of the system work effectively. For example, notice is deemed adequate if it gives the defendant enough information to prepare an Answer or otherwise respond.\textsuperscript{155} By contrast, the fairness argument treats notice as a matter of political morality not contingent on other elements of the system. It is required either as a primary moral obligation or as a prophylactic measure to ensure nonarbitrary decisionmaking.

The idea of fairness as reason-giving has important implications for \textit{Twombly}. If the argument is persuasive, it means that the Court’s rejection of \textit{Conley}’s possibility standard is supported by a principle of fairness, not just a policy of case screening.\textsuperscript{156} Of course, our litigation system does have a complaint requirement, in contrast to the

\textsuperscript{155} See 5A \textsc{Wright \& Miller}, \textit{supra} note 8, § 1215, at 190–93.

\textsuperscript{156} The \textit{Twombly} Court itself is not absolutely clear about the basis for its plausibility standard. On the one hand, the Court discusses the need for case screening at the pleading stage, suggesting that it might have in mind a case-screening rather than a fairness justification. Bell Atlantic Corp. v. \textit{Twombly}, 127 S.Ct. 1955, 1966-67 (2007). On the other hand, the Court introduces plausibility not in the course of discussing case screening, but rather in the course of dismissing the “possibility” standard as an improper interpretation of \textit{Conley} and wrongheaded. \textit{Id.} at 1968–69 Notably, the Court does not argue that the possibility standard is undesirable because of a high risk of meritless suits or high discovery costs. Instead, it argues that possibility simply is not a sensible standard in general because it tolerates “a wholly conclusory statement of claim” and allegations that do not even support “a ‘reasonably founded hope’ that a plaintiff would be able to make a case.” \textit{Id.} at 1968–69. While the Court does not refer here to fairness, a concern about fairness to defendants certainly could explain the force and generality of its criticism.
hypothetical system without pleadings that we have been discussing. Moreover, Conley’s possibility standard at least gives the defendant notice of what the dispute is generally about. The problem, however, is that a possibility standard tolerates complaints that do no more than describe conduct within the ordinary baseline of acceptable behavior. A complaint of this sort gives the defendant no better reason to defend than no complaint at all. In other words, if possibility pleading is sufficient, a defendant acting perfectly appropriately could be forced to bear the burden of a defense without the plaintiff offering a good reason why he should do so.

Although no court has gone as far as to construe Twombly in precisely these fairness terms, at least one court has equated the plausibility standard explicitly with fairness, observing that “Rule 8 is born out of a need to ensure fundamental fairness for defendants.” Moreover, there are other courts that assume plausibility is included in what it means to give “fair notice.”

One additional point deserves special mention. The argument in this section relies on a process-based theory of procedural fairness, but as I have explained elsewhere, there are reasons to doubt whether process-based theories make sense in civil

---

157 United States ex rel. Snapp, Inc. v. Ford Motor Co., 532 F.3d 496, 503 (6th Cir. 2008) (also stating “a complaint need not provide an exhaustive roadmap of a plaintiff's claims, but it must be sufficient to ‘give the defendant fair notice of what the . . . claim is and the grounds upon which it rests.’”). It is important to note, however, that the court does not explain what it means by “fundamental fairness.”

158 See, e.g., Phillips v. County of Allegheny, 515 F.3d 224, 233–34 (3d Cir. 2008) (“[W]e understand the Court to instruct that a situation may arise where, at some point, the factual detail in a complaint is so undeveloped that it does not provide a defendant the type of notice of claim which is contemplated by Rule 8.”); Airborne Beepers Video, Inc. v. AT&T Mobility LLC, 499 F.3d 663, 668 (7th Cir. 2007) (“[A]t some point the factual detail in a complaint may be so sketchy that the complaint does not provide the type of notice of the claim to which the defendant is entitled under Rule 8”); cf. Robbins v. Oklahoma, 519 F.3d 1242, 1246 (10th Cir. 2008) (“This requirement of plausibility serves not only to weed out claims that do not (in the absence of additional allegations) have a reasonable prospect of success, but also to inform the defendants of the actual grounds of the claim against them.”).
procedure. However, these doubts are less relevant to the pleading stage, because they rest on the assumption that all parties are legitimately before the court and thus fairly subject to the rules that make courts work most effectively. For defendants who are not properly brought into court, being required to defend a lawsuit is similar to being harmed in other ways. Insofar as the state’s harming persons without good reason fails to accord those persons the respect to which they are entitled outside litigation, forcing a person to defend without good reason might do the same within litigation.

In sum, the fairness argument based on reason-giving has force in the pleading context. It needs a more rigorous development than I have space to give it here. But

---

159 See, e.g., Bone, supra note 137, at 509-10 (critiquing process-based theories of procedure); Other scholars have expressed similar doubts, although on different grounds. See Larry Alexander, Are Procedural Rights Derivative Substantive Rights?, 17 LAW & PHIL. 19, 31–36 (1998) (questioning the existence of substance-independent procedural values); Larry Alexander, The Relationship Between Procedural Due Process and Substantive Constitutional Rights, 39 U. FLA. L. REV. 323, 325–26, 341–43 (1987) (“[B]ecause the procedure for applying a rule can always be viewed as part of the substance of the rule itself, a concern for procedure apart from substance verges on incoherence.”); Ronald A. Dworkin, Principle, Policy, Procedure, in A MATTER OF PRINCIPLE 72, 101–03 (1985) (questioning whether there can be any moral harm from denying participation apart from the moral harm of an erroneous outcome).

160 To elaborate a bit, the question that a process-based proponent must answer is why respect for persons is not fully satisfied by an outcome for each individual litigant that is the best it can be given institutional constraints and the equal rights of other litigants. See Bone, supra note 137. In other words, since adjudication is primarily about producing good outcomes, why are parties entitled to anything more than those procedures that are justified on outcome-based grounds? Of course, there are background rights that all institutions must respect, such as the right against torture, but the case for process-based procedural rights extends further to encompass institution-specific rights as well, such as the right to control one’s own lawsuit.

161 A skeptical reader might wonder whether the burden of discovery is substantial enough to trigger a moral obligation to give reasons. I believe the answer is yes, at least in some cases. Discovery can be very costly, especially when the plaintiff acts strategically to resist disclosure. Moreover, filing can also impose serious reputation and psychological harms. Another issue has to do with whether respect for persons extends to artificial persons like corporations. While I cannot explore this issue here, it is worth mentioning that the principle against the arbitrary exercise of state power extends generally, suggesting that its support is general too. I am indebted to Ian Farrell for alerting me to this point. Finally, one might object that the fairness argument fails to take account of important social interests. Suppose, for example, that plaintiffs suspect the defendant is using an unsafe method to transport nuclear waste, but the suspicions are not based on any concrete facts or evidence. Also, suppose that an accident could expose millions of people to serious levels of radioactivity. The possibility standard would permit this lawsuit, but the
the analysis in this section is sufficient to show that the argument must be taken seriously. In particular, if fairness demands plausibility not just possibility, Twombly’s critics have a much more challenging task than they realize. They must be able to point to an opposing principle with moral weight that supports a more relaxed standard than plausibility. As the following section shows, this is no easy matter.

2. *Fairness to Plaintiffs*

Sometimes the information needed to make allegations is in the possession of the defendant and inaccessible to the plaintiff even with a reasonable investigation. Critics of Twombly, and more generally of rules stricter than liberal notice pleading, argue that dismissal of a lawsuit is unfair when the plaintiff cannot obtain the information necessary to meet the applicable pleading standard. The question, however, is why. More precisely, why is it *unfair* to impose pleading requirements that apply to all plaintiffs and are justified in general terms simply because those requirements happen to make it difficult for some plaintiffs to sue?

plausibility standard would bar it. One might argue that the powerful social interests at stake should allow plaintiffs at least to take some depositions. However, it is not at all clear that this is the right thing to do. Litigation is not the only institution for handling problems of this sort. Nuclear waste disposal, for example, is heavily regulated and there are public agencies capable of taking action. More generally, the hypothetical raises the complicated question of the proper use of adjudication as an investigative tool. Adjudication is most centrally about determining legal rights. It also has investigative value, of course, but this value is not enough alone to justify suit without some colorable legal claim. Furthermore, if the need for depositions is compelling enough on particular facts, it might be possible to make an exception. The fairness argument supports only a *prima facie* right to reasons. If litigation is somehow the only viable option to handle a serious social problem, the social interests at stake might outweigh the prima facie right.

I am grateful to Patrick Woolley for prompting me to think about this issue.

\(^{162}\) See, *e.g.*, Hoffman, *supra* note 4; Spencer, *supra* note 4.
One might argue that a strict pleading requirement is unfair because it deprives meritorious plaintiffs of a *right*—the right of access to court.\(^{163}\) This argument, however, begs a difficult question: Is there a *moral* right to access for individuals with meritorious civil claims? The right must be moral because a purely legal right provides no ground for objecting to conditions that the law imposes on the right.

One might try to anchor this right in respect for individual dignity by invoking the dignitary theory of process value.\(^{164}\) But obviously any right supported in this way must make room for reasonable procedural limitations; otherwise no procedural system could operate at all effectively. Therefore, it is not enough to criticize stricter pleading simply on the ground that it denies access. The critic must be prepared to explain why the standard is not a reasonable limitation on the right of access if it makes the procedural system work more effectively overall by screening meritless suits.

In addition, respect for individual dignity applies equally to defendants and plaintiffs. This insight has several important implications. For one, if respect for the defendant’s dignity requires a plaintiff to give legitimate reasons, as the previous section argued, then plaintiff’s right to access must be balanced against defendant’s right to reasons, and it is possible that *Twombly*’s plausibility standard strikes the morally appropriate balance.

---

\(^{163}\) The Supreme Court has recognized a “fundamental constitutional right of access to the courts.” Bounds v. Smith, 430 U.S. 817, 828. However, this constitutional right is limited and usually requires that the lawsuit involve a “fundamental interest.” See * Tribe*, *supra* note 6; Michelman, *supra* note 6, at 1169.

\(^{164}\) See *supra* notes 141-143 and accompanying text. Or perhaps anchor it in the conditions for political legitimacy. But legitimacy in turn can be grounded in respect for persons. See, e.g., Frank I. Michelman, *The Supreme Court and Litigation Access Fees: The Right to Protect One’s Rights—Part II*, 1974 DUKE L.J. 527, 539, 552–53, 556–57 (1974) (arguing that “all citizens by virtue of their citizenship have a ‘fundamental interest’ in access” to courts no matter what substantive right they allege and regardless of whether extrajudicial remedies exist).
Moreover, even pleading requirements stricter than plausibility might be justified on balancing grounds, but with a somewhat different moral balance. Intentionally filing a meritless suit out of spite or in order to force an unjustified settlement offends a defendant’s dignity in much the same way an intentional tort does: the plaintiff intentionally harms the defendant. This means that a procedural system must balance two moral rights: the defendant’s right to be free from intentionally filed meritless suits and the plaintiff’s right of access to file a meritorious suit. A strict pleading rule designed to screen meritless suits might be a morally acceptable way to strike this balance.

To be sure, stricter pleading treats plaintiffs who do not have access to necessary information less favorably than plaintiffs who do have access. However, the disadvantaged plaintiffs are not targeted for special treatment. The burden results from the general application of a rule itself justified on procedural grounds. Many procedural rules share this feature. When limits on discovery create errors, for example, it is because the plaintiff who is denied additional discovery is unable to obtain information in the possession of the defendant, and when summary judgment terminates a lawsuit that should go to trial, it is often because the plaintiff lacks evidence in the possession of the defendant that is needed to fill a gap in the prima facie case.\footnote{One might argue that strict pleading is different from these examples because it denies access altogether and thus deprives a meritorious plaintiff of any opportunity to obtain relief. But this is incorrect. As I explain later, plaintiffs who cannot sue because of strict pleading can still use the threat of filing as leverage for a pre-filing settlement. \textit{See infra} notes 223-226 and accompanying text. To be sure, the pre-filing settlement is likely to fall short of full compensation, but that is true for post-filing settlements as well.}

In sum, two major points emerge from this process-based fairness analysis. First, there is a serious argument that \textit{Twombly}’s plausibility standard can be justified by fairness as reason-giving, consistent with a proper balance of fairness principles. It is
important to see, however, that this process-based fairness argument supports only a thin plausibility standard like the one the *Twombly* Court intended.\textsuperscript{166} All the defendant is entitled to receive is some reason why his situation is special enough to require a defense. This is satisfied by allegations of conduct outside the normal baseline and supporting a higher-than-background probability of wrongdoing.

Second, it follows that fairness does not have strong or obvious implications for pleading standards that require more factual detail than thin plausibility. These stricter pleading standards aim to screen undesirable lawsuits in order to prevent unjustified settlements and reduce litigation costs, and as such they must be evaluated on outcome-based grounds.

B. Stricter Pleading and Case Screening from an Outcome-Based Perspective

This section analyzes the use of pleading standards to screen undesirable suits. These standards include those that demand greater factual specificity than thin plausibility, and they also include plausibility itself, insofar as plausibility cannot be defended on process-based fairness grounds and therefore must be justified by its efficacy in achieving outcome-based case-screening goals.

The following discussion develops an outcome-based analysis. It first focuses on the choice of a suitable metric for evaluating outcome quality and distinguishes between utilitarian and rights-based approaches. It then addresses the difficult, but fundamental, question of which lawsuits to screen; in other words, the problem of defining the “undesirable” suit. With this background in place, the discussion turns to an analysis of

\textsuperscript{166} See supra Part II.B.
the costs and benefits of different case-screening approaches, comparing strict pleading with other options.

1. **Choosing an Outcome-Based Metric**

Outcome-based theories evaluate procedural rules according to the quality of the outcomes they make possible. Strict pleading furthers outcome quality by screening undesirable lawsuits. (In this section, I refer to suits that should be screened as “undesirable lawsuits” to pitch the discussion at an appropriate level of generality, and I address the definition of “undesirable” in Section 2B. below.167) Strict pleading, however, is not the only screening method; there are others such as penalties, fee-shifting, summary judgment, and pre-filing merits screening. All these methods must be evaluated and compared under an outcome-based approach, and the metric for evaluation depends on whether the outcome-based theory is utilitarian or rights-based.

a. **Utilitarian**

The utilitarian theory—or more precisely the law-and-economics version of utilitarianism—evaluates procedural rules by how well they balance social benefits and costs in the aggregate. The optimal rule from among a set of feasible alternatives is the rule that maximizes expected social benefit net of costs, or what is equivalent, minimizes the total of expected social costs.168 The social costs that matter for a procedural rule are

---

167 *See infra* notes 186–194 and accompanying text.

168 This is an application of the Kaldor-Hicks efficiency criterion. *See* Allan M. Feldman, *Kaldor-Hicks compensation*, in 2 *THE NEW PALGRAVE DICTIONARY OF ECONOMICS AND THE LAW* 417 (Peter Newman ed. 1998). It is all too common in the procedure literature to associate “efficiency” with reducing litigation costs and deterring frivolous suits, and “fairness” and “justice” with facilitating court access and expanding litigation opportunities for meritorious plaintiffs. *See, e.g.*, Spencer, *supra* note 73, at 22; cf. Hoffman, *supra* note 4, at 1–2 (juxtaposing efficiency and access to court). This way of framing the analysis is misleading. An efficiency metric is as concerned for plaintiffs as it is for defendants. It aims, after all, to achieve a socially optimal balance among all the competing objectives. In particular,
the costs of error operating with the rule (expected error costs) and the costs of administering the rule (expected process costs).  

As for error, any case screening device is inevitably imperfect; it lets some undesirable lawsuits pass through and screens some desirable lawsuits. Thus, any analysis of error costs must consider two different types of error: failing to screen undesirable suits (a false positive), and screening desirable suits (a false negative). Moreover, what matters is the expected cost of error, which is the cost of an error discounted by the probability it will occur. To complicate matters a bit further, the cost of one type of error need not be the same as the cost of the other. A well-known example is from the criminal law, where the cost of convicting the innocent (a false positive) is considered much more serious than the cost of acquitting the guilty (a false negative).

Thus, the overall objective of a utilitarian approach is to select the case screening rule or combination of rules that minimizes the total sum of expected social costs, efficiency, like fairness, is inconsistent with a single-minded commitment to reducing litigation costs and frivolous suits. Such a commitment gives insufficient weight to the error costs of screening meritorious suits. Efficiency and fairness/justice support different metrics, of course, but the difference does not have to do with running roughshod over plaintiffs.

For a more detailed account of expected error and process costs and how to combine them to evaluate pleading rules, see ROBERT G. BONE, THE ECONOMICS OF CIVIL PROCEDURE 125–49 (2003). For example, both false positive and false negative errors dilute the deterrent effect of the substantive law, which increases social costs. Moreover, a very defective screening mechanism that allows lots of meritless suits in and keeps lots of meritorious suits out might lead to a loss of confidence in the court system, which could produce serious long-term costs if people resort to less peaceful means to settle disputes.

This is just an application of the general idea of expected utility or expected value. Formally, let \( r \) be the probability of an error and let \( c \) be the cost of the error. The expected error cost is \( r \times c \). For a more extensive explanation of expected value, see BONE, supra note 169, at 20–36.

including false positive error costs, false negative error costs, and process costs. Each component of the sum is an *expected cost*, meaning that it is the cost of the particular thing if it happens (i.e., a false negative error, a false positive error, or a motion and hearing applying the rule) discounted by the probability it will happen.\(^{173}\)

Obviously it makes no sense on this view to object to a strict pleading or other screening rule on the sole ground that it screens desirable lawsuits, including suits in which the defendant actually violated the law and the plaintiff could prove it with access to discovery. It depends on how many desirable and undesirable lawsuits are screened, the relative costs of the two types of error, and the expected process costs of administering the rule. At least in theory, a strict pleading rule might be optimal even if it screened a large number of desirable suits.\(^{174}\)

\(^{173}\) The objective function to be minimized can be expressed mathematically as follows. Let \(r_{\text{FN}}\) be the probability of a false negative error (screening of a desirable suit) and let \(c_{\text{FN}}\) be the cost of a false negative error. Let \(r_{\text{FP}}\) be the probability of a false positive error (filing and litigation of an undesirable suit) and let \(c_{\text{FP}}\) be the cost of a false positive error. Let \(s\) be the probability that whatever rule is in place will be invoked, thereby necessitating a motion, hearing, and the rest, and let \(c_{\text{PC}}\) be the process cost expended to apply the rule when the rule is invoked. The goal of an economic analysis is to choose the rule or approach that minimizes the following expression: \(r_{\text{FN}} \times c_{\text{FN}} + r_{\text{FP}} \times c_{\text{FP}} + s \times c_{\text{PC}}\).

\(^{174}\) For a numerical example, suppose 30% of all potential lawsuits are undesirable (e.g., meritless) and 70% are desirable (e.g., meritorious). Suppose that a strict pleading rule screens 60% of the undesirable suits and 20% of the desirable suits, and assume the cost of allowing an undesirable suit (false positive) is 20 and the cost of screening a desirable suit (false negative) is 10. Finally, suppose defendants file motions to dismiss in 30% of the cases and that it costs 5 to deal with a motion when one is filed. With these assumptions, one calculates the expected cost of the strict pleading rule as follows. The probability of a false negative error is: \(0.20 \times 0.70 = 0.14\). Similarly, the probability of a false positive error is: \(0.40 \times 0.30 = 0.12\). Therefore, \(r_{\text{FN}} = 0.14, c_{\text{FN}} = 10, r_{\text{FP}} = 0.12, c_{\text{FP}} = 20, s = 0.3,\) and \(c_{\text{PC}} = 5\). The expected cost of the strict pleading rule is: \(0.14 \times 10 + 0.12 \times 20 + 0.3 \times 5 = 5.3\). Compare this to the expected cost without strict pleading. Assuming no desirable suits are screened with very liberal pleading, the expected false negative error costs are zero, and assuming all undesirable suits are filed and get past the pleading stage, the expected false positive error costs are \(0.3 \times 20 = 6.0\). Therefore, if the only choice is between a strict pleading rule and no strict pleading rule, the strict pleading rule wins, even though it screens 20% of desirable suits.
b. Rights-Based

Those commentators who reject the law-and-economics approach rely, if only implicitly, on a rights-based theory to condemn strict pleading.\(^{175}\) A rights-based theory focuses on protecting individual rights rather than minimizing social costs in the aggregate.\(^{176}\) The rights-based objection to strict pleading is that it prevents meritorious plaintiffs from vindicating their substantive legal rights. Notice that this objection differs from its process-based counterpart by locating the violated right in the substantive law rather than a general right of access to court and defining the violation in terms of the outcome rather than the way the process itself treats litigants.

The fact that plaintiffs have trouble vindicating their substantive rights, however, is not by itself enough to condemn strict pleading or any other case screening device on rights-based grounds. For one thing, the plaintiff is not the only party with a legal right. The defendant also has a right to be free from liability when the substantive law so provides. It follows, therefore, that if a moral wrong occurs when a meritorious plaintiffs’ legal right is not vindicated, a moral wrong must also occur when a defendant is forced into court by a meritless plaintiff and held liable or pressured to settle

\(^{175}\) Unfortunately proceduralists rarely explain their normative premises carefully, and this is true for those who criticize \textit{Twombly} and strict pleading. Some \textit{Twombly} critics frame the problem as a tension between achieving efficiency and guaranteeing access for meritorious plaintiffs to vindicate substantive rights. \textit{See, e.g.}, Hoffman, \textit{supra} note 4, at 1–2, 7; Spencer, \textit{supra} note 4, at 433, 479–83. Others simply note the risk of screening meritorious suits, apparently assuming that this alone is a sufficient ground for objection. \textit{See, e.g.}, Ward, \textit{supra} note 11, at 911. To make sense of these arguments, it is best to interpret them as invoking a right on plaintiffs’ side capable of pulling against economic efficiency and supporting a priority concern for plaintiffs.

\(^{176}\) For a developed argument defending a rights-based theory of procedure, see Ronald A. Dworkin, \textit{Principle, Policy, Procedure, in A MATTER OF PRINCIPLE} 72 (1985).
If these two wrongs have the same moral weight, it is not appropriate to protect the plaintiff’s right exclusively at the expense of tolerating meritless suits that infringe the defendant’s.\footnote{A settlement can be unfair or unjust to a defendant even though the defendant gave his consent. Consent cannot validate a settlement when the background conditions that produce the consent are themselves normatively flawed. For more on this point, see Bone, supra note 135, at 1981–85.}

Furthermore, the argument proves too much. If a moral wrong occurs whenever a substantive right is not vindicated, all procedural systems would be morally defective because all systems are prone to at least some error.\footnote{There is a difference, however, between how a false negative and a false positive work in this argument, which might complicate the point a bit. A false negative denies access to an entitled plaintiff and thus absolutely prevents her from vindicating her substantive right through adjudication. A false positive, on the other hand, simply imposes a risk of unjustified liability; the defendant can still win in the end. To be sure, the innocent defendant is forced to defend a lawsuit, but it is not clear that this is a moral wrong if his legal entitlement is a right to be free from liability as opposed to a right to be free from suit. Cf. Mitchell v. Forsyth, 472 U.S. 511, 526 (1985) (distinguishing an immunity from a substantive defense, noting that “[qualified immunity] is an immunity from suit rather than a mere defense to liability”). Nevertheless, the defendant is also pressured to settle, and if the pressure is intense enough, it is not clear why paying the unjustified settlement is morally different from failing to gain access on plaintiff’s side. See Dworkin, supra note 176 (making this point about procedural rights).} And society would be morally committed to preventing all errors no matter how high the cost—an obviously absurd result.

Therefore, both types of error—screening desirable suits and inviting undesirable suits—matter to a rights-based approach, just as they matter to a utilitarian approach. However, the two types of error need not be balanced in the same way. The key is to probe behind the substantive right to its underlying justification. If a substantive legal right aims to protect interests that are valued in moral terms, the failure to vindicate the particular right is a cost that must be measured in moral terms as well.
Suppose, for example, that a plaintiff files a Section 1983 claim against state officials for violation of her First Amendment right to freedom of speech. Many people assign substantial moral weight to constitutionally protected interests, including, and perhaps especially, the interests protected by the First Amendment. Moral interests have the property that they are not easily outweighed by ordinary costs lacking moral weight. The First Amendment right, for example, is generally understood to prevent a state legislature from banning speech even if the ban would make so many people better off that aggregate welfare would increase substantially as a result.

It follows that a false negative—the erroneous screening of plaintiff’s meritorious First Amendment-based claim—creates a cost measured in moral terms, and this cost cannot be easily outweighed, in the usual utilitarian sense, by ordinary false positive and process costs even if the latter are substantial. Unless a false positive implicates substantive interests on the defendant’s side with comparable moral weight, the error risk balance should favor the plaintiff.

Rights-based balancing might generate the same result as utilitarian balancing, but it also might not. For example, many courts and commentators believe that meritless Section 1983 suits chill aggressive, yet lawful, action by state officials and discourages talented individuals from seeking government positions. If these costs are substantial

---

180 Section 1983 here refers, of course, to 42 U.S.C. § 1983, which provides the remedial vehicle for individuals to sue state officials who violate their constitutional rights.

181 This is captured doctrinally by the strict scrutiny test that requires a compelling state interest. See CHEMERINSKY, supra note 6, at 541–42.

182 See Anderson v. Creighton, 483 U.S. 635, 638 (1987) (noting that “permitting damages suits against government officials can entail substantial social costs, including the risk that fear of personal monetary liability and harassing litigation will unduly inhibit officials in the discharge of their duties”); Peter H. Schuck, Suing Our Servants: The Court, Congress, and the Liability of Public Officials for Damages, 1980 SUP. CT. REV. 281, 305–15 (finding that the risk of personal liability can chill government
enough and the risk of meritless filings large enough, utilitarian balancing could support a strict pleading rule for constitutional claims.

The matter looks quite different, however, from a rights-based perspective. The moral weight behind the plaintiff’s constitutionally protected interest creates a prima facie right to litigate a meritorious section 1983 claim. To be sure, this prima facie right can be rebutted. But a rights-based analysis, unlike its utilitarian counterpart, requires more than a modest improvement in aggregate welfare to do so.  

This brief account leaves many philosophical questions unanswered, and this Article is not the place to explore them with care. The discussion in this section is sufficient to highlight four points that will be needed for the analysis of regulatory options in Sections B and C below. First, even a rights-based approach must consider both types of error. Second, rights-based balancing is likely to support a weaker screen that tolerates more undesirable suits than utilitarian balancing, at least when moral interests are at stake on the plaintiff’s side. Third, rights-based balancing is not necessarily appropriate for all lawsuits; it depends on whether there is a substantive interest at stake that has moral weight.

Fourth, rights-based balancing has implications for an equality constraint. Just as one should compare moral stakes on both sides of the party line, one should also compare

---

183 This is one way to understand the Court’s reasoning in Goldberg v. Kelly, 397 U.S. 254 (1970) (a procedural due process, not a First Amendment case). See also Fuentes v. Shevin, 407 U.S. 67, 90 n.22 (1972) (noting in a procedural due process case that “ordinary costs cannot outweigh a constitutional right” to a hearing and that “procedural due process is not intended to promote efficiency or accommodate all possible interests: it is intended to protect the particular interests of the person whose possessions are about to be taken”).

But see Jack M. Beermann, A Critical Approach to Section 1983 with Special Attention to Sources of Law, 42 STAN. L. REV. 51, 95 (1989) (questioning the empirical basis for concerns about chilling local government officials).
moral stakes across different lawsuits. This means that scarce process resources should be allocated to produce an error risk distribution that is roughly proportional to the importance of the different moral interests at stake.\footnote{See Dworkin, supra note 176. This equality constraint needs more careful elaboration, but the simple formulation in the text is sufficient for our purposes. The general idea is that a procedural system should care more about outcome errors when the moral harm from those errors is more serious, and therefore invest more in preventing errors associated with more serious moral harms. However, it would be absurd if the equality constraint demanded that scarce process resources be funneled exclusively to those lawsuits that have moral stakes and away from those that do not. Professor Dworkin’s theory, which derives ultimately from a background right to treatment as an equal, \textit{id.} at 84–85, avoids this problem by assuming that all outcome errors produce moral harm no matter what substantive law is involved. The moral harm inheres in the denial of a substantive legal right itself (plaintiffs and defendants have mirror image rights in Dworkin’s theory). \textit{Id.} at 80–81. The seriousness of this moral harm, however, depends on how the legal system as a whole values the type of substantive interest at stake, which in turn depends on how much the legal system is willing to invest to protect that interest. \textit{Id.} at 93–98. The result is a requirement similar to the one I describe in the text: every procedural system must distribute the risk of error consistent with the best theory of moral harm embedded in the legal system as a whole. This means that the error risk distribution must respect the relative weight of moral harm as between litigants in the same lawsuit and among litigants in different types of suits.} If the moral importance of freedom of speech, for example, greatly exceeds the moral importance of being free from property damage, procedures should be designed to produce a substantially lower risk of false negative error in free speech cases than in property damage cases. Thus, the optimal mix of cases entering the litigation system is the mix that results from giving proper weight to each of the moral interests at stake.

2. \textit{Defining the “Undesirable” Suit}

To recap the argument so far, I began by showing that a thin plausibility standard of the sort the \textit{Twombly} Court endorses might be justified on process-based fairness grounds, but that any rule requiring more factual detail must be justified, if at all, on outcome-based grounds, as a device to screen undesirable suits. From an outcome-based perspective, however, there is no reason to focus exclusively on strict pleading. In particular, a utilitarian should choose whatever regulatory approach achieves the
minimum of expected error costs plus expected process costs. And a rights-based proponent should choose whatever regulatory approach achieves the best distribution of error risk in light of the moral nature of the substantive interests at stake.

The remainder of this Article applies these two metrics to chart a roadmap for designing optimal case-screening rules, and along the way identifies the most salient factors to consider.\textsuperscript{185} This analysis has a lesson for district judges who are tempted to interpret \textit{Twombly}’s plausibility standard strictly. Whether strict pleading should be used as a screening tool involves complex normative, empirical, and analytical issues. For this reason, there is a danger that district judges who use strict pleading to handle case backlog and frivolous suits will only make matters worse. It is much better to leave the policy analysis to the Advisory Committee and the formal rulemaking process than to try to fashion rules on a case-by-case basis.

Before delving into the analysis in depth, it is essential first to define as precisely as possible what constitutes an “undesirable lawsuit” that should be screened.\textsuperscript{186} The following addresses this threshold question.

Virtually everyone can agree that truly meritless suits are undesirable, but what constitutes a truly meritless suit is far from obvious. The clearest example is a case in


\textsuperscript{186} Unfortunately, this definitional question is all too often ignored in discussions of case screening, despite its pivotal importance. \textit{See} Warren F. Schwartz & C. Frederick Beckner, \textit{Toward a Theory of the “Meritorious” Case: Legal Uncertainty as a Social Choice Problem}, 6 GEO. MASON L. REV. 801, 801 (1998) (noting the lack of careful attention to the question). Indeed, there are commentators who even assume that some suits with substantial merit should be screened. \textit{See, e.g.,} \textit{id.} at 816–19 (defining a frivolous suit as any lawsuit with less than a fifty percent chance of success); David Rosenberg & Steven Shavell, \textit{A Model in Which Suits are Brought for Their Nuisance Value}, 5 INT’L REV. L. & ECON. 3, 3 (1985) (including in the category of nuisance suits some that have merit but that the plaintiff would not take to trial because the litigation costs exceed the expected recovery).
which the true state of affairs plainly shows that the defendant is not legally responsible.

In these cases, the liability gap is created by a missing fact, the existence of which is decidable on purely descriptive grounds. For example, suppose the plaintiff claims she slipped on a puddle of water that the defendant carelessly failed to remove, when in fact there was no puddle of water and she tripped on her shoelaces instead. The true state of affairs in this example is missing a fact necessary to liability. For another example, suppose the plaintiff alleges that the defendant failed to perform an express agreement when in fact the defendant had absolutely no dealings with the plaintiff that could possibly create an agreement. Again, the defendant’s innocence turns on a fact that, in theory at least, can be verified objectively with sufficient information.

Cases turning on the application of standards such as reasonable care are not as easy to classify. For example, suppose the plaintiff claims that the defendant failed to notice an icy patch while driving. Whether the defendant exercised reasonable care cannot be determined categorically in the same way that factual matters can. The decision requires judgment even when all relevant information is available to the decisionmaker. It is still possible, however, to define a meritless suit in this situation. A suit is meritless if it is clear that no reasonable (or perhaps rational) decisionmaker considering all the available information could possibly find that the defendant failed to meet the standard. In the icy roadway hypothetical, for example, the evidence might clearly show that the defendant driver was attentive and the icy patch hidden completely from view.

---

187 This is, of course, the test for summary judgment, but assessed objectively on the assumption that all the information is available to the decisionmaker, not just whatever information the parties have been able to obtain.
This leaves cases in which the evidence is not as clear, so the determination of liability involves a range of reasonable judgment. None of these cases should be classified as meritless. Yet some of them will present a stronger case for violating the applicable negligence standard than others when all of the existing information is considered. In the icy roadway example, the case for a violation would be very strong if the evidence showed that the icy patch was plainly visible and the driver failed to notice it because he was adjusting the car radio. The case would be much weaker (but not meritless) if the evidence showed that the icy patch was somewhat obscured from view but still visible with sufficient effort.

If one adopts a utilitarian perspective, it is possible to justify treating some of these cases as undesirable even though they are not meritless. An undesirable lawsuit, according to this view, is a lawsuit where the probability that a fully informed decisionmaker would find a violation of the standard falls below some threshold level. A utilitarian can support a threshold significantly above zero because of the high costs of the litigation system.\footnote{As the threshold is set higher, the cases that enter the system have a greater chance of liability and thus a stronger claim on scarce private and public litigation resources. See, e.g., Keith N. Hylton, *When Should a Case Be Dismissed? The Economics of Pleading and Summary Judgment Standards*, available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=897486. See also Schwartz & Beckner, supra note 186, at 816–19 (using social choice theory to justify setting the threshold at fifty percent).} A rights-based proponent, however, should have considerable difficulty defending a threshold above zero when the meritorious plaintiff’s substantive right implicates weighty moral interests that are not matched on the defendant’s side.\footnote{I do not mean to suggest that a rights-based proponent would never set a threshold above zero. Process costs and false positive error costs matter in a rights-based theory, and if these costs are sufficiently serious, a rights-based proponent can support a strictly positive screening threshold too. Before doing so, however, she will demand a high risk of very serious costs and reliable information confirming that risk.}

Moreover, any system of adjudication like ours that relies on judges to develop new
substantive law must be careful about screening meritorious suits. A case might appear weak at first but end up producing a novel legal precedent.\textsuperscript{190}

The Twombly opinion illustrates the importance of defining undesirable lawsuits clearly. The Court refers to a “largely groundless claim,” “anemic cases,” and cases with “no reasonably founded hope that the discovery process will reveal relevant evidence.”\textsuperscript{191} These references can be construed as encompassing not only meritless suits but also meritorious suits where the merits are “anemic” and the claim “largely groundless.”\textsuperscript{192} If this is the correct interpretation, the plausibility standard should be construed as a merits threshold; in other words, the complaint would have to show a plausible chance of success at trial.

If this is what the Court has in mind, it has entered treacherous territory. The normative stakes involved in screening \textit{meritorious} suits are complex and highly contested. This favors a decisionmaking process that is more open to broad-based public input and more conducive to debate than case-by-case adjudication. In addition, the Court is not well suited to set an optimal merits threshold. The task requires global information on social costs and benefits, statistical and economic expertise, and a systemic approach not available through individual case adjudication. As a result, any effort to screen meritorious cases is better left to formal rulemaking or to the legislative process.

\textsuperscript{190} One might cite \textit{Brown v. Board of Education}, 347 U.S. 483 (1954) as an example. Of course, the judge has a chance to develop new law at the motion to dismiss stage when she evaluates the strength of the case in light of the legal standards. However, she will not have this chance if a screening device prevents filing, and even if she does, her decision might benefit from exposure to more information than is available at the pleading stage.


\textsuperscript{192} \textit{Id.} (emphasis supplied).
One can, however, interpret Twombly’s plausibility standard as focusing exclusively on meritless suits. Understood in this way, the standard is not a merits threshold, but rather an epistemological threshold. Plausibility is the Court’s way of expressing the level of confidence that a judge must have in deciding that a suit is meritless. This interpretation better fits the Court’s opinion. Recall that the Court meant its plausibility standard to be quite generous and made a point of stating that heightened pleading was for the formal rulemaking process or for Congress.\footnote{Id. at 1973 n.14. It also fits what the Court says about the importance of judges not making predictions at the pleading stage of plaintiff’s likely success at trial. \textit{Id.} at 1965 (“a well-pleaded complaint may proceed even if it strikes a savvy judge that actual proof of those facts is improbable, and ‘that a recovery is very remote and unlikely’").} Moreover, screening meritorious suits on the ground of high litigation costs is hard to square with the Court’s rights-based views in other areas.

In keeping with the previous discussion, I shall define an “undesirable suit” as a meritless suit, and I shall define a meritless suit as one in which the defendant is clearly not liable as an objective matter.\footnote{\textit{Bone}, supra note 185, at 533.} This definition should be relatively uncontroversial, and it is consistent with what I believe is the best interpretation of \textit{Twombly}. It is worth bearing in mind, however, that much of the following analysis can be applied equally well to undesirable lawsuits defined to include meritorious suits that fall below a merits threshold.

A caveat is in order before proceeding further. We actually know very little about meritless litigation. There is an intense and widespread perception among judges, lawyers, and scholars that it is a serious problem, but reliable empirical studies are hard...
to come by.\textsuperscript{195} The following discussion assumes that a problem exists and examines the regulatory alternatives for dealing with it. But if there is not a serious problem, obviously there is no reason to adopt costly regulations to fix it. Also, the paucity of empirical evidence affects the following discussion in another way. It means that the analysis must rely heavily on theory, especially game theoretic tools, to predict causes and effects.

3. Analyzing the Regulatory Options

There are two general approaches to dealing with the problem of meritless filings. One approach targets the causes of the problem, while the other targets the symptoms. The following discussion first explains why targeting causes is likely to be unsuccessful and then examines approaches like strict pleading that target symptoms by screening meritless suits.

a. Targeting Causes

The \textit{Twombly} Court assumes that the cause of meritless filings is asymmetry of discovery costs and the settlement leverage it confers.\textsuperscript{196} When the defendant has a great deal of information relevant to liability and the plaintiff has very little—as was true in the \textit{Twombly} case and is true for many of the cases that prompt meritless suit concerns—the plaintiff can threaten to impose high discovery costs without having to worry that the defendant will reciprocate. Faced with a credible threat of incurring substantial discovery costs, the defendant will choose to settle rather than bear the expense.\textsuperscript{197}

\begin{footnotesize}
\begin{enumerate}
\item See Randy J. Kozel \& David Rosenberg, \textit{Solving the Nuisance-Value Settlement Problem: Mandatory Summary Judgment}, 90 VA. L. REV. 1849, 1851 n.3 (2004) (citing empirical studies and noting a “general consensus that a problem exists” but also emphasizing the “paucity of empirical research substantiating its extent”).
\item \textit{Twombly}, 127 S.Ct. at 1967.
\item See e.g. Chris Guthrie, \textit{Framing Frivolous Litigation: A Psychological Approach}, 67 U. CHI. L. REV. 163 (2000); Rosenberg \& Shavell, \textit{supra} note 186.
\end{enumerate}
\end{footnotesize}
One or another version of this explanation is widely shared, and it has a superficial plausibility. If it were true, one obvious way to treat the problem would be to reduce discovery cost asymmetry by limiting discovery—although this is difficult to do effectively as the *Twombly* Court itself recognized.\textsuperscript{198} The fact is, however, that this account explains only part of the meritless suit problem, and probably not a very substantial part at that. There is strong reason to believe that informational asymmetry is a much more important cause.

To understand why, let us consider a case with symmetric information and see if it is possible to explain meritless filings on the basis of asymmetric costs. Symmetric information means that both the plaintiff and the defendant know the suit is meritless. Because the suit is meritless, it can succeed at trial only if the factfinder makes a mistake. Assuming the chance of a mistake is small, the likelihood of recovery should be too small in most cases for a rational plaintiff to be willing to incur the cost of litigating through trial.\textsuperscript{199} Thus, the defendant should, if rational, simply refuse to settle, confident that the plaintiff will drop. Anticipating this response from the defendant, the plaintiff will not file in the first place. The fact that discovery costs are substantial and asymmetrically

\textsuperscript{198} *Id.* The limits might be imposed by general rule, but that approach has problems as well. See Bone, *supra* note 135, at 2006–11.

\textsuperscript{199} Throughout the rest of the analysis, I use the standard economic model of litigation. See, e.g., ROBERT COOTER & THOMAS ULEN, LAW AND ECONOMICS 481 (1988); Bone, *supra* note 169, at 30. That model assumes that the parties are rational and make decisions by considering their expected benefits and costs. An expected benefit or expected cost is just the benefit or cost discounted by the probability it will materialize. See *supra* note 171 (explaining expected value). For example, the plaintiff’s expected benefit from filing suit is the likely trial award discounted by the probability of success at trial and her expected cost is the cost of litigating. A plaintiff will file suit if the expected benefit exceeds the expected cost. This can be expressed mathematically as follows. Let $p$ be the plaintiff’s probability of trial success, $w$ be the likely trial award if the plaintiff is successful, and $c$ the plaintiff’s anticipated cost of litigating through trial. The plaintiff will file suit if $p \times w > c$. When this condition holds, the lawsuit has positive expected value and is known as a PEV suit. When the inequality is reversed, the lawsuit has negative expected value and is known as a NEV suit.
distributed makes no difference because a rational plaintiff will never go through the discovery stage.\footnote{See Bone, supra note 169.}

Defenders of the cost-asymmetry explanation respond to this line of reasoning by arguing that a defendant will settle for any amount less than the cost of answering the plaintiff’s complaint because any such settlement will make him better off than actually paying to file an answer.\footnote{See, e.g., Rosenberg & Shavell, supra note 186.} A plaintiff with a meritless suit, therefore, will file as long as what she expects in settlement exceeds her costs of preparing and filing a complaint. As I have explained elsewhere, this argument has a number of flaws, one of which is that it predicts at most nuisance settlements probably too small to justify the expense of regulatory intervention.\footnote{See Bone, supra note 185, at 537–41. Also, in cases where both sides know suit is meritless, repeat-play defendants can benefit from building a reputation for fighting meritless suits, which deters future meritless filings. One possible exception to these predictions deserves special mention. If there are enough separate stages of a lawsuit and total litigation costs are divided among the different stages, it is possible for the plaintiff in a negative expected value suit to make a credible threat to take the case all the way through trial. See Lucian Bebchuk, A New Theory Concerning the Credibility and Success of Threats to Sue, 25 J. LEGAL STUD. 1 (1996). The analysis relies on backward induction and is somewhat complicated. There is no need to explore it, however, because it is very unlikely to apply to a truly meritless suit or to produce substantial settlements even when it does. See Bone, supra note 185, at 539 n.73; see generally Joseph A. Grundfest & Peter H. Huang, The Unexpected Value of Litigation: A Real Options Perspective, 58 STAN. L. REV. 1267, 1299–1305 (2006) (describing an option model of litigation that can explain negative expected value suits but does not clearly explain truly meritless suits that have tiny variance and that both parties know have no trial value).}

The result is different when the plaintiff can make a credible threat to take the case through discovery and trial because of a very large risk of error or extremely high stakes. To illustrate, consider a large class action, like the *Twombly* case, and assume that it is meritless. Also assume that the total expected classwide recovery if the defendants are held liable is $50 million and that it would cost $2 million to litigate the
case through trial. If the class attorney makes litigation decisions in the best interests of the class as a whole, the attorney should be willing to take the case to trial whenever the risk of error at both summary judgment and trial exceeds 4%.  

In this example, the filing incentive does not depend on asymmetric costs; it depends on unusually high error or large stakes. No doubt some meritless filings can be explained in this way. Still, the conditions that make this explanation viable are not easy to satisfy for a truly meritless suit, especially when judges are strict about insisting that the plaintiff submit admissible evidence to get past the summary judgment stage. Given the definition of a meritless suit, it is hard to imagine how a plaintiff could find enough evidence to support a prima facie case.

---

203 The numbers work out as follows. With an error rate of 4%, the expected classwide recovery is 0.04 x $50 million, or $2 million, exactly equal to the expected litigation costs. Thus, anything in excess of 4% yields a positive expected value from taking the case to summary judgment and then on to trial. The result changes if the class attorney makes the decisions in her self-interest, because the attorney expects only a fraction of the total recovery. The typical class attorney fee is 25% of class recovery. See Theodore Eisenberg & Geoffrey Miller, Attorney Fees in Class Action Settlements: An Empirical Study, 1 J. EMPIR. LEGAL STUD. 27 (2004). When the attorney expects 25% of the recovery and litigation costs are $2 million, litigation is cost-justified for any error rate in excess of 4% only if the expected recovery exceeds 200 million. (0.04 x 0.25 x $200 million = $2 million.) Alternatively, litigation is cost-justified when expected recovery is $50 million and litigation costs are $2 million only if the error rate exceeds 16%. (0.16 x 0.25 x $50 million = $2 million.)

204 Filing might also occur when the plaintiff is motivated by revenge or her attorney is overly optimistic about her ability to trick a jury. But it is difficult to imagine a level of optimism sufficiently high or a desire for revenge sufficiently strong to make it worthwhile to try a case that the attorney knows is meritless. Another way in which trial can become a credible threat for a meritless suit is if the plaintiff’s attorney develops a reputation for taking meritless suits to trial. See Amy Farmer & Paul Pecorino, A Reputation for Being a Nuisance: Frivolous Lawsuits and Fee Shifting in a Repeated Play Game, 18 INT’L REV. L. & ECON. 147 (1998). But the conditions that make reputation work in this way are difficult to sustain for meritless suits.

205 In general, for a suit to be meritless, the plaintiff must have no admissible evidence of an essential element of her prima facie case, or the defendant must have a clearly decisive affirmative defense. In the former situation, it is difficult to see how the plaintiff could get past summary judgment without admissible evidence. In the latter situation, the defendant should be able to prevail on summary judgment if its evidence supports only one reasonable conclusion. In an antitrust conspiracy case like Twombly, for example, the plaintiff must offer evidence of conspiracy above and beyond parallel conduct in order to avoid summary judgment, and it is not clear what evidence a meritless plaintiff could provide. The plaintiff

More generally, it is difficult to construct a convincing account of filing and litigation incentives that points to asymmetric discovery costs as the cause of most meritless filings. Targeting asymmetric costs therefore is not likely to be productive.

The key to explaining meritless litigation is to introduce asymmetric information about the merits at the pleading stage. The analysis so far assumed that plaintiffs and defendants both know the suit is meritless, and this assumption is critical to the results. Since the plaintiff knows, she is not willing to go to trial (unless the case is one with unusually high error or stakes), and since the defendant knows, he can call the plaintiff’s bluff.206 By contrast, when at least one party does not know the suit is meritless, filing becomes much more attractive because then meritless suits sometimes obtain substantial settlements.207

It is extremely difficult, however, to cure informational asymmetries. The obvious method is to use discovery, but the problem, as the Twombly Court recognized, is that

---

206 Professor Paul Stancil has developed an interesting model that he claims shows how discovery cost asymmetries can produce meritless filings even when both parties know that the suit is meritless. Paul J. Stancil, *The Practical Economics of Pleading*, available at http://papers.ssrn.com/pape.tar?abstract_id=1144850. Professor Stancil seems to assume a cost-free stage, which makes his model similar to other models developed by Professors Rosenberg and Shavell and by Professor Bebchuk. See Rosenberg & Shavell, *supra* note 186; see also Bebchuk, *supra* note 202, at 20–22 (generalizing the Rosenberg-Shavell model to any cost-free stage). The problem with using cost-free-stage models is that there is not likely to be a cost-free stage sufficiently late in the litigation to support a substantial settlement. Professor Stancil assumes that the case will proceed directly to discovery after the complaint is filed, but this ignores the possibility that the defendant could file a motion to dismiss or a motion for a more definite statement, and also ignores the mandatory scheduling conference before formal discovery begins. Furthermore, discovery itself is not a completely one-sided affair. It costs the plaintiff something to propound discovery requests. And the defendant can increase the plaintiff’s costs by refusing to comply and forcing a motion to compel.

207 See *infra* note 212-213 and accompanying text.
defendants are likely to settle before much discovery takes place.208 And the more
discovery one allows, the higher the costs and the greater the pressure to settle. Thus,
targeting causes is not likely to be a very successful strategy.

b. Targeting Symptoms

Targeting symptoms means screening meritless suits, and there are two general ways to do this. One is to employ a gatekeeping rule, like strict pleading, that conditions access to court directly.209 The other is to employ an incentive-shaping rule, like Rule 11 penalties or fee-shifting, that applies later in the litigation and works by reducing the expected payoff from a meritless suit.210

(1) Gatekeeping Rules

Information about the merits can be asymmetrically distributed in two different ways. The plaintiff might know the suit lacks merit while the defendant does not, or the defendant might know while the plaintiff does not.211 The strategic dynamics are different between the two, but the reason meritless plaintiffs sue is the same. They

---

209 For a classic analysis of gatekeeping, although not in the context of litigation, see Reiner Kraakman, Gatekeepers: The Anatomy of a Third-Party Enforcement Strategy, 2 J. L. ECON. & ORG. 53 (1986). In our situation, the gatekeeper is the judge, who prevents a meritless filing by checking for it at the pleading stage, and gatekeeping rules assist the judge in this task. See id. at 54 (defining a gatekeeper).
210 In a sense, all screening rules shape incentives. The difference is that a gatekeeping rule screens cases directly at the pleading stage while an incentive-shaping rule works indirectly by affecting the payoff.
211 Of course, a plaintiff who knows suit is meritorious might try to persuade the uninformed defendant of that fact, and an informed defendant who knows suit is meritless might do the same for the uninformed plaintiff. However, these efforts are not likely to succeed because the other type—i.e., informed plaintiffs in meritless suits and informed defendants in meritorious suits—will mimic in an effort to trick the uninformed party into believing that they are the same and should receive a large settlement as well. Anticipating this, the uninformed party will discount the truth of any statements or representations made by the opposing side. See, e.g., Lucian Arye Bebchuk, Suing Solely to Extract a Settlement Offer, 17 J. LEG. STUD. 437, 442 n.8 (1988); Steven Shavell, Sharing of Information Prior to Settlement or Litigation, 20 RAND J. ECON. 183 (1989).
believe that there is significant chance they will receive the same settlement meritorious plaintiffs receive. The reason is simple. When the defendant decides to settle but does not know who is meritorious and who is meritless, he cannot adjust his offer to the type of suit, so he makes the same offer to everyone and meritless as well as meritorious plaintiffs receive it. When the information asymmetry is reversed and the plaintiff does not know whether her suit is meritless, she expects the settlement due a meritorious plaintiff if her suit turns out to be meritorious.

The first type of informational asymmetry—informed plaintiffs and uninformed defendants—can be difficult to handle with gatekeeping rules. For example, any device that relies on input from the defendant, such as a merits review at the filing stage, is not likely to work well, since the defendant lacks the information necessary to mount a successful challenge. Devices that depend on increasing the cost of filing, such as increasing filing fees, will not discourage meritless suits as long as meritless plaintiffs

---

212 Developing a fighting strategy by refusing to settle does not work as well when information is asymmetrically distributed. Cf. supra note 202 (mentioning the possibility of a fighting strategy for symmetric information cases). If the defendant does not know whether suit is meritless, he cannot confine his fighting strategy to meritless suits, so he will sometimes end up fighting a meritorious suit by mistake and pay litigation costs as a result. If the plaintiff does not know whether suit is meritless, the defendant has to worry that fighting might backfire if the meritless plaintiff is sufficiently confident that her suit is meritorious. Defendants sometimes fight in these situations but not always and fighting generates additional litigation costs.

213 In fact, a settling defendant has no incentive to offer anything less than a meritorious plaintiff would accept because he knows that only meritless plaintiffs would accept the lower amount. See Avery Katz, The Effect of Frivolous Suits on the Settlement of Litigation, 10 INT’L REV. L. & ECON. 3 (1990).

214 For example, some states have created medical malpractice panels to screen medical malpractice suits immediately upon filing. See Jean A. Macchiaroli, Medical Malpractice Screening Panels: Proposed Model Legislation to Cure Judicial Ills, 58 GEO. WASH. L. REV. 181, 188–97, 239–51(1990). It is significant, however, that medical malpractice suits usually involve informed defendants and uninformed plaintiffs.
expect settlements greater than the filing cost.\textsuperscript{215} Moreover, substantial filing costs can create undesirable distributional effects if they disproportionately affect meritorious plaintiffs with little wealth, a result that is particularly troubling from a rights-based perspective committed to distributing error risk proportional to relative moral harm.\textsuperscript{216}

Strict pleading is also a poor option. If the meritless plaintiff and her attorney are willing to file a suit they know lacks merit, they are probably unethical enough to concoct allegations to get past the pleading stage. Since the defendant is unaware the suit lacks merit, he will have trouble verifying the truth of these allegations and arguing convincingly at a dismissal hearing.

The second type of informational asymmetry—informed defendants and uninformed plaintiffs—is more suitable for gatekeeping rules. \textit{Twombly} is an example of this type of case, since the defendants are likely to have private information bearing on the existence of a conspiracy. So too are other cases that prompt some of the most serious concerns about frivolous suits, including securities fraud, medical malpractice, and some civil rights litigation.\textsuperscript{217}

A plaintiff might be uninformed in these cases because the cost of investigating before filing is too high. Even if the cost is reasonable, the plaintiff’s attorney might file

\textsuperscript{215} The analysis is more complicated. One must compare the filing cost with the \textit{expected} gain from a meritless filing, which is the amount of the settlement discounted by the probability an offer is made. \textit{See} Katz, \textit{supra} note 213, at 13–15.

\textsuperscript{216} \textit{See supra} note 184 and accompanying text.

\textsuperscript{217} In fact, most of the empirical studies indicating a possible problem focus on substantive fields where the defendants are likely to be the informed parties. \textit{See}, \textit{e.g.}, Kozel & Rosenberg, \textit{supra} note 195, at 1851 n.3 (collecting studies, including ones done for securities fraud and medical malpractice cases). For example, securities fraud plaintiffs are likely to have difficulty obtaining the information necessary to allege scienter. Antitrust plaintiffs suffer from similar problems, as \textit{Twombly} itself illustrates. Also, in medical malpractice cases, the doctor is the one with knowledge about what he did or did not do. And those civil rights suits involving constitutional rights that feature defendant’s intent as an element also fit this profile. \textit{See}, \textit{e.g.}, Crawford-El v. Britton, 93 F.3d 813 (D.C. Cir. 1996).
without investigating in the hope of settling before the discovery stage or learning about the merits more cheaply from the informed defendant’s response. Of course, the defendant has an incentive not to reveal useful information when the suit is meritorious in the hope that the ignorant plaintiff will conclude (incorrectly) that her suit is meritless, but there is a limit to this strategy’s effectiveness when the plaintiff anticipates it in advance.218

Not all gatekeeping approaches are viable for this category,219 but strict pleading is more promising in these cases than it is in the informed-plaintiff cases, even though it also has problems. Strict pleading screens meritless suits not just by giving the judge more detailed allegations to assess the merits, but also, and more importantly, by encouraging attorneys to investigate before filing. When the plaintiff’s attorney investigates and discovers the suit lacks merit (and knows that the defendant also knows), the informational asymmetry is removed, and along with it, the information-based incentive to file a meritless suit.220 It is possible that the attorney will simply fabricate the allegations, but in these cases, unlike those with informed plaintiffs and uninformed defendants, the defendant knows the truth and can make it difficult for the plaintiff to

---

218 I have analyzed these incentives in some detail elsewhere. See Bone, supra note 185, at 550–63. The strategic dynamics are complex but the intuition is as described in this paragraph.  
219 A merits review at the filing stage benefits from the fact that the defendant is informed and can provide input, but the fact that the plaintiff is uninformed makes it more difficult to test the defendant’s assertions. Many medical malpractice suits, for example, involve an informed doctor and an uninformed patient, and the informed doctor can provide information to a medical malpractice screening panel. See supra note 214. A professional panel has the expertise and incentive to challenge the doctor’s arguments as long as the panel is neutral (which is at least a questionable assumption), and this might offset to some extent the plaintiff’s informational disadvantage.  
220 Or if she does file, she is likely to obtain only a small nuisance settlement. See supra note 202 and accompanying text.
trick the judge. Also, an attorney who fails to investigate when she should is not necessarily a liar and should be more reluctant to make up allegations.

Any benefit from strict pleading in screening meritless suits must be balanced against the cost of screening meritorious suits. The result of this balance depends on whether a utilitarian or rights-based metric is employed. But for both metrics, two factors are likely to reduce the weight on the false negative side of the balance.

First, in many cases where the defendant has private information, the plaintiff can still obtain clues from reasonably accessible sources to support an inference of wrongdoing.\textsuperscript{221} Provided the focus of the strict pleading rule is limited to determining whether the suit is truly meritless, and provided the requisite strength of the inference is not set too high, at least some meritorious plaintiffs should be able to gather enough clues by investigating to make satisfactory allegations. Admittedly, the expense of investigating increases the cost of filing and this, like filing fees, can have undesirable distributional effects.\textsuperscript{222} However, unlike filing fees, the investment in a pre-filing investigation can pay dividends later in the litigation if the results help guide a more efficient discovery process.

Second, a meritorious plaintiff who cannot file suit because she cannot obtain enough information before filing still might be able to obtain a pre-filing settlement. To be sure, a rational defendant would not settle if he was certain the meritorious plaintiff could not file. But certainty is elusive. A large corporate defendant cannot keep track of

\textsuperscript{221} See, e.g., Epstein, supra note 59 (discussing publicly available information to support an inference of conspiracy in antitrust cases); Geoffrey P. Miller, \textit{Pleading After Tellabs}, 11–19 (April 2008), \textit{available at} http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1121396 (discussing the types of information that can support an inference of scienter in a securities fraud suit).

\textsuperscript{222} See \textit{supra} note 216 and accompanying text.
all the different ways that clues to its wrongdoing might be disclosed. Corporate employees might leak information, for example, or a creditor or other third party might behave in ways that signal something is amiss. Furthermore, if the plaintiff’s attorney hires a private investigator, as is common in securities fraud suits after the Private Securities Litigation Reform Act imposed strict pleading, the defendant will not necessarily know which investigator the plaintiff hired and how skillful the investigator is.

When both the plaintiff and the defendant are uncertain whether a reasonable investigation will reveal enough to enable suit, they each have an incentive to settle before any investigation is performed. By settling, the defendant saves the cost of litigation in the event the plaintiff’s investigation reveals enough information to plead successfully, and the plaintiff saves the cost of investigating as well as the cost of litigating to a post-filing settlement in the event an investigation proves fruitful.

---

223 See Miller, supra note 221, at 17.

224 There is precedent for a practice of extra-adjudicative settlement emerging in an environment of long delays and high litigation costs. In the late nineteenth and early twentieth century, repeat-play employers cooperated with brokers for injured employees to create a private scheme of aggregative settlement for workplace injuries. See Samuel Issacharoff & John Fabian Witt, The Inevitability of Aggregate Settlement: An Institutional Account of American Tort Law, 57 VAND. L. REV. 1571, 1584–96 (2004). Also, in the first half of the twentieth century, liability insurers developed a private scheme of scheduled settlement for automobile accident claims. Id. at 1603–10. At least some of these settlements appear to have been negotiated without the filing of a lawsuit. See id. at 1582–83. These developments are associated with special factors, such as correlated conduct and reputation markets, which might not exist in all cases. They are relevant nevertheless because they arose during periods of intense case congestion (a sharp spike in tort litigation in the late 1800s and a huge influx of automobile cases in the early 1900s) that should have greatly weakened any credible threat by plaintiffs to litigate their claims.

225 Consider the following simple and somewhat more formal analysis. Let R be the expected recovery by settlement or trial judgment if a meritorious plaintiff files suit and gets past the pleading stage. Let c be the plaintiff’s cost of litigating to the point of recovery and let d be the defendant’s cost. Assume that there are two types of cases. In type 1, all the information about wrongdoing remains hidden. In type 2, enough information is available from reasonably accessible sources to allow an investigating plaintiff to meet the pleading burden. Neither party knows for sure what type of case he has, but both estimate a
These settlements will not be as large as a meritorious plaintiff would receive if she did not have to worry about meeting a strict pleading burden, since they reflect a discount for the probability the investigation proves fruitless.\textsuperscript{226} Still, from a utilitarian perspective, paying something in settlement adds to deterrence and reduces the cost of false negatives, making strict pleading more attractive on balance. From a rights-based perspective, the significance of the compensation shortfall depends on the moral theory underlying the substantive law, but at least the defendant is held accountable to some extent—a morally more acceptable result than letting the defendant off scot free.

Nevertheless, a strict pleading approach has another flaw. Its beneficial effects follow only if the pleading standard is not set too high. If the standard is too onerous, few, if any, meritorious plaintiffs will be able to satisfy it even with a reasonable investigation. In that case, the defendant will offer nothing or only a small settlement, confident that there can be very few suits. Given the difficulty of distinguishing among different pleading standards linguistically, any strict pleading rule must rely heavily on trial judge discretion. But discretionary judgments are likely to vary, making it difficult for lawyers to predict the outcome in advance and thus weakening the rule’s beneficial effect in discouraging meritless filings. Moreover, discretionary judgments increase process costs by inviting more motions to dismiss, more hotly contested dismissal hearings, and more extensive deliberation on the part of the judge.

\textsuperscript{226} See supra note 225.
Still, one should not exaggerate these problems. Discretionary judgments should be more reliable and exclude fewer meritorious suits if they focus on the narrow question whether the suit is or is not truly meritless.\textsuperscript{227} Moreover, this determination can be improved if it is informed by more knowledge about the case, and Section C below outlines an approach along those lines.

(2) \textit{Incentive-Shaping Rules}

The other way to screen meritless suits is to use incentive-shaping rules. The rules I have in mind operate after the pleading stage and work indirectly by affecting incentives to file and litigate meritless suits. The two chief examples are penalties for meritless filings and fee-shifting.\textsuperscript{228} As we saw in the previous section, gatekeeping rules, and especially strict pleading, are likely to work better when the informational asymmetry favors the defendant than when it favors the plaintiff. The result is less clear for incentive-shaping rules.\textsuperscript{229}

To illustrate, consider a perfectly accurate penalty system that targets only meritless suits. The informed plaintiff knows she will be penalized should the defendant

\begin{itemize}
  \item \textsuperscript{227} This is as compared with an inquiry that ranges more widely to engage the broad question whether the suit meets some threshold level of merit.
  \item \textsuperscript{228} Other creative approaches have been proposed, such as making summary judgment mandatory as a condition for a judicially enforceable settlement agreement and giving the defendant an option to bar settlement. See Kozel & Rosenberg, \textit{supra} note 195 (proposing mandatory summary judgment); David Rosenberg & Steven Shavell, \textit{A Solution to the Problem of Nuisance Suits: The Option to Have the Court Bar Settlement}, 26 INT’L REV. L. & ECON. 42 (2006) (proposing an option to bar settlement). In addition, purely substantive reforms, such as caps on pain and suffering and elimination of punitive damages, have been proposed to reduce filing incentives for frivolous suits, but I do not address these alternatives here.
  \item \textsuperscript{229} Fee-shifting conditional on a trial loss is not a good choice to deter meritless filings. It does generate deterrence benefits for some cases. See Katz, \textit{supra} note 213, at 17–19 (showing that the effect of adopting the English Rule depends on whether it increases expected recovery for meritorious plaintiffs); Farmer & Pecorino, \textit{supra} note 204 (arguing that fee shifting can deter meritless filings when reputation is the driving force). But it is also likely to deter a large number of weak but meritorious suits as well. See Steven Shavell, \textit{Suit, Settlement and Trial: A Theoretical Analysis Under Alternative Methods for the Allocation of Legal Costs}, 11 J. LEG. STUD. 55 (1982).
\end{itemize}
seek penalties, and this discourages meritless filings.\textsuperscript{230} Moreover, with perfect accuracy, the penalty threat has no effect on meritorious suits. By contrast, when the plaintiff is uninformed, as she is in the informed-defendant cases, she must worry about paying the penalty in the event her suit turns out to be meritless, so even a perfectly accurate penalty deters some meritorious suits.

One can handle the over-deterrence problem for the informed-defendant cases to some extent by imposing the penalty only if the plaintiff does not do a reasonable pre-filing investigation.\textsuperscript{231} But this modification introduces new problems. The legal definition of what constitutes a reasonable pre-filing investigation is necessarily imprecise so even meritorious plaintiffs must worry about whether their investigation will qualify, and this new source of uncertainty deters meritorious suits.\textsuperscript{232}

Of course, no penalty system is perfectly accurate and errors at the penalty stage are likely to deter some informed plaintiffs from filing meritorious suits, especially if they are risk-averse.\textsuperscript{233} Moreover, if the issues at the penalty stage are hotly contested, as one might expect when the plaintiff’s attorney is targeted and her reputation is at stake,

\textsuperscript{230} The defendant does not know at the outset but presumably will find out and seek penalties if the case goes through discovery. Therefore, the plaintiff must factor the penalty risk into her settlement calculations. \textit{See} Katz, \textit{supra} note 213, at 19–20.

\textsuperscript{231} This is the approach used by current Rule 11. \textit{See} FED. R. CIV. P. 11(b). I have argued elsewhere that a penalty equal to the cost of discovery might encourage all plaintiffs to conduct reasonable pre-filing investigations, but this assumes reasonableness can be defined clearly enough to minimize errors and uncertainty at the enforcement stage. \textit{See} Bone, \textit{supra} note 185, at 592.

\textsuperscript{232} There is another reason why a penalty system that targets only meritless suits might work better in cases where the plaintiff is the informed party. As a practical matter, trial judges are much more willing to impose penalties when the plaintiff and her attorney file knowing that the suit is meritless (or with reckless disregard for the obvious possibility). Penalties smack of punishment for a moral wrong and knowingly filing is much more culpable than a negligent failure to conduct a reasonable pre-filing investigation.

\textsuperscript{233} To be sure, meritorious plaintiffs know they are likely to settle and not actually face a penalty proceeding, but the penalty risk reduces the settlement amount and thereby deters meritorious filings at the margin. \textit{See} Katz, \textit{supra} note 213.
the process cost of litigating a penalty could be substantial.\textsuperscript{234} These high process costs, combined with uncertain success, can also discourage defendants from seeking penalties.\textsuperscript{235} This disincentive can be offset by increasing the penalty amount, but given the inevitability of error at the penalty stage, a larger penalty is likely to deter more plaintiffs with meritorious suits.\textsuperscript{236}

These costs might be substantial enough to rule out penalties. However, a system of penalties limited to extreme cases of intentional meritless filings might work if the risk of erroneously penalizing a meritorious suit can be limited. This option is discussed in the following section.

C. Toward An Optimal Approach

The previous analysis, though brief, points to the desirability of a hybrid approach, one sensitive to the different reasons for meritless filings. Moreover, since we know very little about the extent of the meritless suit problem, it is wise to proceed cautiously, especially when measures aimed at deterring meritless suits can have serious

\textsuperscript{234} This does not necessarily imply that the expected process costs of a penalty system will be high as well. While the cost of any particular penalty proceeding is likely to be large, the number of proceedings should be small if penalties are working properly. In that case, the defendant should expect few frivolous suits and therefore a low likelihood of success, which should produce few penalty motions and thus few proceedings. Still, expected process costs might be lower with strict pleading, because pleading-based dismissals do not implicate the attorney’s reputation as directly and conspicuously and therefore are not likely to trigger the same degree of contentiousness.


\textsuperscript{236} This is another reason why a penalty approach is better suited to informed-plaintiff than informed-defendant cases. The risk of chilling meritorious suits might be even greater when the plaintiff is uninformed and penalized for failing to conduct a reasonable pre-filing investigation if the vagueness of the reasonableness standard creates a large enough error risk.
adverse effects on meritorious suits and add substantially to process costs. One additional point is worth emphasizing. The analysis here is exclusively outcome-based. If process-based fairness principles require a thin plausibility standard, then that standard must be applied regardless whether any of the following measures are also implemented.

It is possible that the only serious meritless suit problem involves large and complex class actions, especially those in which the expected recovery and error risk support a credible threat by the class attorney to go to trial. The extent of meritless class action litigation is a matter of some dispute, but assuming it is large enough to warrant regulatory attention, the appropriate response is to focus exclusively on the class action and target the solution to the particular features of that problem. Various approaches have been suggested, including merits review at the certification stage and mandatory summary judgment. This is not the place to analyze these different options. My point is simply that if the problem is with the class action, the solution should be confined to the class action.

If the meritless suit problem extends beyond the class action, as the popular view holds and the strategic analysis of the previous sections suggests, then the regulatory solution should be tailored to the nature of the broader problem. In particular, since

---

237 See 5A WRIGHT & MILLER, supra note 8, § 1332 (discussing criticism of the 1983 amendments to Rule 11 citing the increase in satellite litigation adding substantial process costs as well as excessive screening of meritless suits).
238 See supra notes 203–205 and accompanying text.
239 For a skeptical view, see Charles A. Silver, We’re Scared to Death: Does Class Certification Subject Defendants to Blackmail?, 78 N.Y.U. L. REV. 1357 (2003).
241 See Kozel & Rosenberg, supra note 195.
information asymmetry is likely to be a major cause, regulation should be tailored to the information structure of the typical lawsuit.\textsuperscript{242}

For example, a penalty system might be appropriate for intentional meritless filings.\textsuperscript{243} The defendant should be required to prove that the suit is meritless and the plaintiff knew it. In order to maximize deterrence, sanctions should be mandatory rather than discretionary, as they are in current Federal Rule 11.\textsuperscript{244} The size of the penalty should also be set in advance rather than left to trial judge discretion, as it is under Rule 11.\textsuperscript{245} One possibility is to set the penalty equal to the defendant’s reasonable costs, including attorney fees, of responding to the complaint.\textsuperscript{246} Given that deterrence is the goal, certainty and predictability are particularly important. Moreover, case-specific discretion is a poor method for selecting an optimal penalty because trial judges lack the

\textsuperscript{242} It is worth mentioning that some of the proposals offered by others, such as mandatory summary judgment, are designed for cases of symmetric information with asymmetric costs, and are not likely to work as well for cases of asymmetric information. \textit{See, e.g.}, Kozel & Rosenberg, supra note 195. \textit{See also} Ted M. Sichelman, \textit{Why Barring Settlement Bars Legitimate Suits – A Reply to Rosenberg and Shavell}, __ CORN. J. L & PUB. POLICY __, available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=986234 (criticizing Rosenberg’s and Shavell’s proposal to give defendants an option to bar settlements).

\textsuperscript{243} It might also be appropriate to impose penalties on a plaintiff or her attorney who does \textit{absolutely no} pre-filing investigation without good reason. This conduct fails to show respect for the defendant and the court and implicates moral culpability. \textit{See, e.g.}, Johnson v. A.W. Chesterton Co., 18 F.3d 1362, 1366 (7th Cir. 1994). However, extending penalties beyond this extreme case to an attorney who does some but not a reasonable investigation is more problematic, given the vagueness of the reasonableness standard. For this reason, I favor narrowing the current scope of subdivisions (b)(2) and (b)(3) of Rule 11.

\textsuperscript{244} \textit{See} FED. R. CIV. P. 11(c)(1). Also, Rule 11’s safe harbor provision should be eliminated. \textit{See} FED. R. CIV. P. 11(c)(2). And, of course, the penalties should be limited to the category of suits covered by current Rule 11(b)(1).

\textsuperscript{245} \textit{See} FED. R. CIV. P. 11(c)(4).

\textsuperscript{246} A fee-shifting penalty forces the plaintiff to internalize the private costs of filing and litigating.
global information necessary to figure out what general deterrence requires in particular cases.  

A penalty rule structured in this way should limit the risk of penalizing meritorious suits. With the burden on the defendant and meritlessness limited to extreme cases, the chance of penalizing a meritorious plaintiff is reduced. Proving lack of merit and plaintiff’s knowledge is bound to be costly, but defendant’s costs can be reimbursed as part of the penalty, thereby reducing the disincentive to seek penalties. As for expected process costs in general, if the penalty works effectively, there should be few penalty proceedings.

Special provision should be made for cases, such as civil rights suits, where a deserving plaintiff’s substantive interest is valued in moral terms and a rights-based balance is appropriate. In these cases, the cost of screening meritorious suits has moral weight. At the same time, there is a moral harm to defendants when a plaintiff knowingly files a meritless suit to harass or force an unjustified settlement. With moral harms on both sides of the balance, some screening of meritorious suits should be permissible. Thus, a penalty approach could fit this situation too, but the penalty might be set at a lower amount.

A penalty system narrowly structured in this way might not deter enough meritless plaintiffs from filing, especially as the plaintiff can gamble on the defendant not knowing the suit is meritless. However, it is not clear how concerned one should be about these cases since it is questionable how much of the frivolous suit problem is due to

\[247\] Current Rule 11 instructs the judge to choose the penalty that achieves general deterrence, but it is difficult to see how trial judges can do that effectively on a case-by-case basis.
filings by informed plaintiffs. As noted previously, the cases triggering the strongest complaints, like antitrust and securities fraud suits, normally involve the reverse situation; they are informed-defendant cases where the plaintiffs are the ones who are uninformed.248

As for these informed-defendant cases, the previous analysis suggests that strict pleading might work but that it also has serious shortcomings. There is precedent, however, for a more promising approach based on supplementing strict pleading with limited pre-dismissal discovery,249 and recently at least one trial judge has indicated a willingness to use this approach to address the information-access problems raised by Twombly’s plausibility standard.250 The idea is simple: give the plaintiff a chance to

248 See supra note 217 and accompanying text.
249 Judges routinely allow targeted discovery before deciding a motion to dismiss for lack of subject matter or personal jurisdiction, but rarely before deciding a 12(b)(6) motion. Nevertheless, a few courts before Twombly permitted discovery in cases subject to heightened pleading standards and with information in the exclusive control of the defendant, usually after finding the original complaint deficient and before the plaintiff filed an amended complaint. See, e.g., Cordero-Hernandez v. Hernandez-Ballesteros, 449 F.3d 240, 244 (1st Cir. 2006) (quoting Becher for limiting discovery to those cases where the allegations make it likely wrongdoing occurred and the information needed to plead with particularity is in the exclusive control of the defendant); New England Data Serv., Inc. v. Becher, 829 F.2d 286, 290 (1st Cir. 1987); Parish v. Beneficial Ill., 1996 U.S. Dist. LEXIS 4453, *1–2 (N.D. Ill. 1996) (permitting a plaintiff to conduct some discovery before re-filing an insufficiently pled RICO claim); see also Reints v. Sheppard, 90 F.R.D. 346, 347 (M.D. Pa. 1981) (agreeing with the policy that dismissal before discovery may be inappropriate for cases of heightened pleading specificity if the information is held exclusively by the defendant). It is worth mentioning that Federal Rule 27 allows a prospective party to take depositions before a lawsuit is filed. However, the prevailing view is that the right is limited to perpetuating testimony that might be lost before trial (such as a dying witness), although there are a few decisions to the contrary. See Bay City Middlegrounds Landfill Site v. Kuhlman Electric Corp., 171 F.3d 1044, 1047 (6th Cir. 1999). But see Reints v. Sheppard, 90 F.R.D. 346, 347–48 (M.D. Pa. 1981); Bowles v. Pure Oil Co., 5 F.R.D. 300, 302–03, 303 n.2 (D. Pa. 1946). Also, some states have special rules allowing pre-filing discovery, but these rules apply only in state court and are available only under limited conditions. See generally Lonnie Sheinkopf Hoffman, Access to Information, Access to Justice: The Role of Presuit Investigatory Discovery, 40 U. MICH. J.L. REFORM 217 (2006) (discussing the range of pre-suit discovery tools).
250 See, e.g., In re Graphics Processing Units Antitrust Litig., 527 F.Supp.2d 1011, 1032–33 (N. D. Calif. 2007) (indicating that narrowly tailored discovery might be allowed before the plaintiff filed an amended complaint provided the need for it is well established). But see In re Flash Memory Antitrust...
conduct limited discovery before deciding a motion to dismiss for failure to state a claim or a motion for a more definite statement. 251

An approach allowing limited discovery must be designed carefully to limit error and process costs. As a threshold matter, the plaintiff should be required to file an affidavit with her complaint describing in detail all the steps she took to investigate the merits before filing and stating what she learned. This requirement will help assure that the plaintiff does not substitute discovery for a pre-filing investigation and impose costs on the defendant without good reason.

In addition, it is important that the discovery be limited in a clear way. One could leave the question of how much discovery to allow to the trial judge’s discretion, but the trial judge is likely to have trouble determining the optimal amount in a complex case because of information and strategic constraints. 252 Therefore, the amount of discovery should be defined by general rule. One possibility is to give the plaintiff the option to take one deposition of each defendant and perhaps serve a narrowly tailored request for documents. 253 The objective is to ensure that discovery is not so costly that it pressures the defendant to settle before it even takes place. 254


251 Another possibility is to give discretion to the trial judge to stage discovery in increments, evaluating the strength of the case after each stage. To proceed to the next stage, the plaintiff would have to convince the judge that the case is strong enough to warrant further discovery given what has already been obtained. I am skeptical, however, that trial judges can make good decisions about discovery’s benefits to implement a staged approach like this in a complex case. Moreover, additional discovery increases costs and adds to the plaintiff’s settlement leverage.

252 See supra notes 128–136 and accompanying text.

253 Justice Stevens proposed something similar to this in his Twombly dissent. See Bell Atlantic Corp. v. Twombly, 127 S.Ct. 1955, 1986–87 (2007).

254 The rule also could require that the discovery be confined to those elements that the plaintiff has trouble alleging because of the informational asymmetry. To facilitate this, the plaintiff might be
After the discovery is complete, the rule might require the plaintiff to file an amended complaint satisfying a strict pleading standard. Alternatively, it might bypass the amendment stage altogether and let the judge determine directly whether the suit is meritless by evaluating whatever information the plaintiff provides (perhaps supplemented by information from the defendant). This approach adds process costs, but the additional costs are limited by the narrow scope of the inquiry, which should be confined to whether the suit is truly meritless.

If the judge determines based on the available information that the suit is probably meritless, the rule might mandate dismissal. Alternatively, the plaintiff might be allowed to proceed on the condition that she pay all the defendant’s costs, including reasonable fees, up to summary judgment if the case is dismissed at that stage. This alternative reduces the risk of false negatives by allowing a plaintiff who has trouble verifying the merits to continue litigating if she is sufficiently convinced to assume the risk of paying the defendant’s fees. Moreover, its cost-shifting condition relieves the defendant from pressure to settle.

This proposal needs more careful analysis than I can give it here. One thing is clear, however. If pleading stage discovery is a good way to deal with the uninformed

---

255 The judge would be required to meet whatever level of confidence the strict pleading rule requires.

256 For one thing, defendants have incentives to conceal information or delay its disclosure when they know that the discovery tools available to plaintiffs are extremely limited and the time period for using them is short. The plaintiff’s lawyer can anticipate these incentives and will try to counter them by strategically deploying the limited discovery available to her. Given this strategic interaction, it is difficult to predict exactly how effective pleading stage discovery will be. However, it is promising enough to give it a try and see how it works.
plaintiff, the Federal Rules should be revised to authorize it explicitly. Allowing
pleading stage discovery fits the current Rules awkwardly at best.\textsuperscript{257} Moreover, with a
new rule, the procedure can be designed optimally and the provisions applied uniformly
to all district courts.

V. Conclusion

This Article set out to analyze Twombly in the broader context of court access
regulation and to chart a roadmap for designing optimal case-screening rules. It began
with a critical examination of Twombly itself. That analysis concluded that the Court’s
plausibility standard represents only a modest departure from traditional notice pleading.
Moreover, Twombly’s interpretation of Rule 8(a)(2) is consistent with the text and history
of the Rule, compatible with principles of stare decisis, and in line with the pragmatic
vision of the original Federal Rule drafters.

The Article then turned in Part IV to the broader normative issues implicated in
regulating court access through strict pleading. It argued that a pleading standard along
the lines of Twombly’s thin plausibility requirement might be justified by a process-based
principle of fairness as reason-giving, but that anything stronger had to be evaluated on
outcome-based grounds. The analysis explored the implications of utilitarian and rights-
based theories for various screening methods, along the way highlighting the crucial

\textsuperscript{257} For example, formal discovery is not supposed to begin until after the parties meet and confer
pursuant to Rule 26(f). See FED. R. CIV. P. 26(d)(1). The trial judge can grant an exception, but judges
might be reluctant to do so without express authorization if they view routine exceptions as inconsistent
with the purposes of the Federal Rules’ elaborate discovery scheme. See, e.g., In re Flash Memory Antitrust
Litig., 2007 U.S. Dist. LEXIS 95869 (N.D. Calif. 2007) (reviewing all the possible bases for pleading stage
discovery and concluding that allowing discovery “before an operative complaint is filed” is inconsistent
with the timing and protective measures that constitute the Federal Rules’ discovery scheme). Rule 27
allows pre-filing depositions but only to perpetuate testimony that might be lost before trial. See supra note
249.
threshold task of defining an undesirable lawsuit, the necessity of correctly identifying the cause of the problem, and the importance of proceeding cautiously in the absence of empirical information and tailoring solutions to probable causes. It concluded by outlining a promising hybrid approach targeting informational asymmetry, one that addresses the informed-plaintiff cases with a system of penalties and the informed-defendant cases with a system of limited discovery prior to a merits review.

Whatever screening approach is adopted, however, two general points should guide its design. First, the project should be handled by formal rulemaking or the legislative process, not by the common law method or case-by-case interpretation of the Federal Rules. Because it requires a controversial choice of normative metric and a coordinated analysis of different rule options, this project is most suitable for a process open to public input, able to generate and properly consider relevant empirical information, and capable of addressing the issues from a global and systemic perspective. Second, because the screening approach should be tailored to the types of cases that involve meritless filings most seriously, any set of rules should be substance-specific. General trans-substantive rules cannot properly take account of the different cost-benefit tradeoffs for different types of cases or the different implications of utilitarian versus rights-based metrics.

Our institution of civil adjudication reflects and protects society’s most basic values, and it is essential that access be regulated in a careful and deliberate way. The hope of this Article is that the framework outlined here will guide a careful analysis and assure a more just and efficient set of pleading and merits-based screening rules.