

Boston University School of Law

Scholarly Commons at Boston University School of Law

Faculty Scholarship

2008

When Should a Case Be Dismissed? The Economics of Pleading and Summary Judgment Standards

Keith Hylton

Follow this and additional works at: https://scholarship.law.bu.edu/faculty_scholarship



Part of the Litigation Commons



BOSTON UNIVERSITY SCHOOL OF LAW

WORKING PAPER SERIES, LAW AND ECONOMICS
WORKING PAPER NO. 06-06



WHEN SHOULD A CASE BE DISMISSED? THE ECONOMICS OF PLEADING AND SUMMARY JUDGMENT STANDARDS

KEITH N. HYLTON

This paper can be downloaded without charge at:

The Boston University School of Law Working Paper Series Index:
<http://www.bu.edu/law/faculty/papers>

The Social Science Research Network Electronic Paper Collection:
http://ssrn.com/abstract_id=897486

When Should a Case Be Dismissed? The Economics of
Pleading and Summary Judgment Standards

Keith N. Hylton*

May 2007

Abstract. This paper applies a simple economic framework to the choice between pleading and summary judgment as points at which a claim can be dismissed. It concludes generally that pleading standards should vary with the evidentiary demands of the associated legal standards and the social costs of litigation. The common law's imposition of higher pleading standards for fraud claims is consistent with this proposition. The theory implies that the rigorous summary judgment standards that have been developed by antitrust courts should lead to a correspondingly rigorous assessment at the pleading stage.

* Professor of Law, Boston University, knhylton@bu.edu. Jon Baker, Bob Bone, Richard Epstein, Giuseppe Dari-Mattiaci, Bob Hahn, and workshop participants at AEI-Brookings, University of Amsterdam, and New York University provided helpful comments. I thank Boston University and Verizon for financial support. I am solely responsible for the views expressed, as well as errors and omissions.

I. Introduction

Pleading is the term lawyers apply to the claims plaintiffs assert when they enter court, and the specific language used to assert them. As every law student quickly learns, pleading under the common law was a high-stakes game. Under the common law writ system, a legal action began when the plaintiff obtained a writ ordering the defendant to appear and defend himself.¹ A plaintiff could lose if he chose the wrong writ, and each writ had its own pleading requirements.²

In the latter half of the 1800s, courts in England and in the United States began to reform the pleading process in an effort to simplify the requirements.³ The reforms aimed to make the pleadings serve the functions of providing notice to defendants and guidance to courts.⁴ The simplified pleading requirements were designed to make legal

¹ See, e.g., BENJAMIN J. SHIPMAN, HANDBOOK OF COMMON-LAW PLEADING 17-18 (Harry Winthrop Ballantine, ed., 3d ed. 1923) (“The original writ was a mandatory letter or executive order from the king to his officer, the sheriff, to compel the defendant to appear in court to answer the demand of the plaintiff”); see also JACK H. FRIEDENTHAL, MARY KAY KANE & ARTHUR R. MILLER, CIVIL PROCEDURE 252 (Thomson/West, 4th ed. 2005)

² Maitland notes that each writ or form of action had a particular procedure, which he refers to as a “procedural pigeon-hole,” including a “particular mode of pleading.” FREDERIC MAITLAND, THE FORMS OF ACTION AT COMMON LAW 3 (A.H. Chaytor and W.J. Whitaker, ed., 1936). As such, “[e]ach procedural pigeon-hole contain[ed] its own rules of substantive law” but “it [was] quite possible that a litigant [would] find that his case [would] fit some two or three of these pigeon-holes.” *Id.* at 3. If this was the case the litigant was forced to choose a single writ, but he could “make a bad choice, fail in his action, and take such comfort as he can from the hints of the judges that another form of action might have been more successful.” *Id.* at 3-4; see also FRIEDENTHAL, KANE & MILLER, *supra* note 1, at 252

³ Maitland notes several key changes to the writ system, including the Common Law Procedure Act of 1852, which provided that it was no longer necessary to mention any form or cause of action in any writ of summons, and the Judicature Act of 1873, which was the most crucial change as it abolished completely the forms of actions and provided that equity and law should be administered concurrently. MAITLAND, *supra* note 2, at 7; see also R. ROSS PERRY, COMMON-LAW PLEADING: ITS HISTORY AND PRINCIPLES 2 (1897) (noting that the Common-Law Procedure Act of 1852 abolished the common law forms of actions in favor of “the simplest possible statements of causes of demand”); FRIEDENTHAL, KANE & MILLER, *supra* note 1, at 252;

⁴ Charles Clark, *The New Federal Rules of Civil Procedure: The Last Phase – Underlying Philosophy Embodied in Some of the Basic Provision of the New Procedure*, 23 A.B.A. J. 976 (1937) [hereinafter Clark, *The New Federal Rules*] (discussing the reasoning behind the reformation of pleadings); FRIEDENTHAL, KANE & MILLER, *supra* note 1, at 252.

judgments turn on the underlying merits of the case rather than the skill of lawyers in satisfying arcane pleading rules.⁵

There are opposing views today on the degree to which courts should use pleading requirements to police the types of claims allowed to enter courts. One view holds that in order for pleadings to serve the purpose of guiding courts and providing notice, courts need to rigorously enforce pleading rules in order to bar claims that fail to meet them.⁶ The opposing view argues that the pleading rules should not be used to bar many claims because there are other devices that can serve this purpose; such as pretrial discovery, pretrial conference, and summary judgment.⁷

The paper presents an economic analysis of the choice between using pleading rules and other devices to screen claims.⁸ Specifically, I will focus on the choice between the pleading stage and summary judgment stage as points at which a claim can be dismissed.⁹ The analysis shows that dismissal rules enhance the average quality or merit of lawsuits, which in turn enhances incentives to comply with the law. Thus, early

⁵ Clark, *The New Federal Rules*, *supra* note 4, at 976-977; FRIEDENTHAL, KANE & MILLER, *supra* note 1, at 252.

⁶ James Alger Fee, *The Lost Horizon in Pleading Under the Federal Rules of Civil Procedure*, 48 COLUM. L. REV. 491 (1948) (Arguing that pleading rules should be strictly enforced to prevent litigation of issues that the parties did not intend to raise); *see also* FRIEDENTHAL, KANE & MILLER, *supra* note 1, at 255.

⁷ James A. Pike & John W. Willis, *The New Federal Deposition-Discovery Procedure*, 38 COLUM. L. REV. 1179 (1938) (discussing federal discovery and depositions procedures and how such procedures relate to notice pleading); *see also* FRIEDENTHAL, KANE & MILLER, *supra* note 1, at 255.

⁸ For a general discussion of the economics of pleading standards, see ROBERT G. BONE, *CIVIL PROCEDURE: THE ECONOMICS OF CIVIL PROCEDURE* 125-157 (New York: Foundation Press 2003). There is a related literature on the allocation of burdens of proof, see Bruce L. Hay & Kathryn E. Spier, *Burdens of Proof in Civil Litigation: An Economic Perspective*, 25 J. LEGAL STUD. 413 (1997). While this paper addresses in a general way the setting of optimal proof burdens, it does not address the allocation of proof burdens.

⁹ Typically, defendants challenge cases at the pleading stage by filing a motion to dismiss for failure to state a claim. The summary judgment motion typically occurs later in the litigation process and often challenges not only the sufficiency of the allegations but also the existence of factual support for the allegations. *See* FRIEDENTHAL, KANE & MILLER, *supra* note 1, at 465-469. Although a summary judgment is not the same as a dismissal (e.g., dismissal applies to the whole claim while summary judgment may apply to all or to only a portion, dismissal is for failure to state a valid claim while summary judgment implies the lack of any genuine issue of fact requiring trial), I treat the two as functionally equivalent in this paper.

dismissals, by eliminating low-merit claims before they become costly, offer benefits to society in comparison to late dismissals. I will apply the results of the analysis to the law on pleading standards, and especially the pleading standard for conspiracy claims under the Sherman Act.

The results suggest a positive economic theory of the common law on pleading standards and a normative theory for the developing antitrust pleading standards. In general, pleading standards should vary with the evidentiary demands of the associated legal standards and the social costs of litigation.¹⁰ This explains why the common law imposed higher pleading requirements for certain claims – e.g., claims of fraud, which were difficult to prove and imposed substantial social costs beyond litigation expenses.¹¹ The theory developed here implies that the rigorous summary judgment standards that have been developed recently by antitrust courts with respect to predatory pricing, resale price maintenance, and conspiracy claims based on parallel conduct should lead to a correspondingly rigorous assessment at the pleading stage. This provides an economic rationale for the Supreme Court’s decision in *Bell Atlantic v. Twombly*,¹² and suggests that it may have a broader application than to the circumstances of that case.

II. Economics of Civil Procedure: Dismissals and Pleading Standards

¹⁰ The evidentiary demands of the associated legal standard will often vary with the expected costs of error. In view of this, this paper’s thesis is equivalent to saying that the pleading standard should vary with the ratio of false conviction to false acquittal costs and the social costs of litigation.

¹¹ See FRIEDENTHAL, KANE & MILLER, *supra* note 1, at 272-273 (stating that the common law required fraud to be pleaded with specificity for various reasons, including the social cost of harm to the defendant’s reputation); CHARLES CLARK, CODE PLEADING 311-312 (2d ed. 1947)[hereinafter CLARK, CODE PLEADING] (stating the same).

¹² 127 S.Ct. 1955 (2007).

In order to gain some economic intuition for the role of pleading standards, we must first examine the function of dismissals in the litigation process. In this section I briefly set out the standard model of litigation and introduce dismissals into that model. The analysis that follows will tie together several building-block models from the literature on the economics of litigation. Since the models are familiar in the literature, I will use numerical assumptions largely instead of variable symbols representing factors of interest. The key factors of interest (and the only ones represented by variable symbols below) are the rate of compliance with the law and the average merit of lawsuits.

A. Standard Model of Litigation

The standard economic model of litigation is a one period model that focuses on the filing and settlement decisions. A suit is filed if the expected judgment exceeds the cost of litigation. So, for example, if the plaintiff thinks his likelihood of winning a judgment of \$100 is 75%, he will file claim as long as his legal expenses are less than \$75. Formally, let P_p represent the *plaintiff's prediction of prevailing*, which is the plaintiff's prediction of the likelihood of a verdict in his favor. Let the plaintiff's cost of litigation be \$30 and the loss suffered by the plaintiff (and the amount awarded if he wins his case) be \$100. A lawsuit will be filed, according to the standard model, when

$$P_p(\$100) > \$30 . \quad (1)$$

Condition (1) says that a suit will be filed when the plaintiff's expected judgment, which is the probability of victory multiplied by the damage award, exceeds his cost of litigation.

The factors that determine settlement in the standard model of litigation are the parties' expectations of success at trial, their litigation costs, and the anticipated damage award. Let P_d equal the defendant's prediction of the likelihood of a verdict in the plaintiff's favor. Let the defendant's cost of litigation be \$30, the same as the plaintiff's. Settlement in the standard model is described by the "Landes-Posner-Gould" condition,¹³ according to which settlement occurs if and only if¹⁴

$$(P_p - P_d) \$100 < \$30 + \$30 \quad (2)$$

In words, settlement occurs if the difference in the parties' expected awards is less than the sum of litigation costs.¹⁵

B. Optimal Civil Procedure and the Role of Dismissals

The standard model of litigation does not incorporate dismissal of lawsuits. In this part, I will augment the standard model to incorporate the dismissal decision. Determining whether society gains from the dismissal or maintenance of a lawsuit requires some examination of the goal of civil procedure.

¹³ See William M. Landes, *An Economic Analysis of the Courts*, 14 J. L. & ECON. 61 (1971); Richard A. Posner, *An Economic Approach to Legal Procedure and Judicial Administration*, 2 J. LEGAL STUD. 399 (1973); John P. Gould, *The Economics of Legal Conflicts*, 2 J. LEGAL STUD. 279 (1971).

¹⁴ The Landes-Posner-Gould condition can be derived from a straightforward consideration of the minimum a plaintiff is willing to accept in settlement and the maximum that a defendant would be willing to pay in settlement. The minimum a plaintiff would accept is the expected net gain from the suit: $P_p v - c_p$. The maximum the defendant will pay in settlement is his total expected cost from the lawsuit: $P_d v + c_d$. The parties have an incentive to settle, then, as long as $P_p v - c_p < P_d v + c_d$. The Landes-Posner-Gould condition (2) follows directly from this inequality. In a more complicated model that takes asymmetric information and strategic interactions into account, the Landes-Posner-Gould condition no longer provides a complete explanation of settlement outcomes, see Keith N. Hylton, *An Asymmetric Information Model of Litigation*, 22 INT'L REV. L & ECON. 153 (2002).

For example, suppose plaintiffs are uninformed and would find it unprofitable to sue an innocent defendant. If the inequality in (2) is reversed, the outcome is one in which some guilty defendants attempt to pool with innocents, and litigation occurs in some percentage of cases but not all, see *Id.*

¹⁵ This can be interpreted as a "bet" in which the difference in the stakes exceeds the cost of the gamble, see Keith N. Hylton, *Agreements to Waive or to Arbitrate Legal Claims: An Economic Analysis*, 8 SUP. CT. ECON. REV. 209, 228 (2000).

1. *Specifying the Goal of Procedure*

From a social welfare perspective, the goal of civil procedure should be to encourage litigation only when it is socially desirable. Following Shavell,¹⁶ the social desirability of litigation can be determined by comparing social costs when litigation occurs to social costs when litigation is prohibited.¹⁷ Let the probability of an injury when the potential defendant/injurer does not comply with the law (e.g., does not take care) be $\frac{3}{4}$. Let the probability of an injury when the potential defendant does comply with the law (takes care) be $\frac{1}{4}$. Let the cost of compliance (e.g., cost of taking care) be \$10. Finally, let the variable w represent *the percentage of potential injurers who comply with the law because of the threat of liability*.¹⁸ For simplicity I will assume that all injured victims file claims (i.e., condition (1) holds).

When lawsuits are prohibited, no one complies, so total social cost is simply the expected harm caused by non-complying actors:

$$\left(\frac{3}{4}\right)\$100 . \quad (3)$$

When lawsuits are permitted, the fraction w of potential injurers comply, so social cost is determined by the harms caused by compliers and non-compliers, their resultant litigation costs, and the cost of compliance. Since the harms and litigation costs generated by compliers and non-compliers occur at the rate $(1-w)\left(\frac{3}{4}\right) + w\left(\frac{1}{4}\right)$, total social cost under the regime in which lawsuits are permitted is

¹⁶ Steven Shavell, *The Social versus the Private Incentive to Bring Suit in a Costly Legal System*, 11 J. LEGAL STUD. 333 (1982).

¹⁷ The approach that follows is based in part on Keith N. Hylton, *The Influence of Litigation Costs on Deterrence Under Strict Liability and Under Negligence*, 10 INT'L REV. L. & ECON. 161, 165-166 (1990).

¹⁸ This model is general enough to apply to intentional harms as well. For example, in the antitrust setting, "the cost of compliance" could represent the profit forgone by the dominant firm if it forbears from some anticompetitive act.

$$[(1-w)^{3/4} + w^{1/4}]($100 + $30 + $30) + w($10) . \quad (4)$$

Suit is socially desirable, then, when social cost is lower in the regime in which lawsuits are permitted. Comparing (3) and (4), litigation is socially desirable when:

$$w[(3/4 - 1/4)$100 - $10] > [(1-w)^{3/4} + w^{1/4}]($30 + $30). \quad (5)$$

In words, this means that a suit is socially desirable when the expected “deterrence benefits” (injuries avoided net of avoidance costs multiplied by the probability of compliance) exceed the total litigation costs. An ideal system of civil procedure rules would seek to maximize the difference between deterrence benefits and litigation costs. Equivalently, an ideal system would minimize the sum of under-deterrence, over-deterrence, and litigation costs. Yet another way of saying the same thing is that an ideal system would minimize the sum of false-acquittal, false-conviction, and litigation costs.

It should be clear that this analysis implies that it might be socially optimal to prohibit all litigation (reverse the inequality in (5)). That would be equivalent to dismissing all lawsuits. But the question I will examine here is whether there is a policy that discriminates between lawsuits that should and should not be dismissed.

Finally, notice that the percentage of law compliers, w , plays an important role in determining the social desirability of litigation. Indeed, solving the inequality in (5), we arrive at the statement that litigation is socially desirable, under the present numerical assumptions, if and only if the percentage of law compliers is greater than the fraction $9/14$. This suggests that procedural rules that support compliance incentives can also enhance the economic benefits provided by litigation.

2. Determining an Optimal Dismissal Rule

Given that the goal of an ideal system of procedure is to minimize the sum of deterrence inefficiencies (under and over-deterrence) and litigation costs, the question I consider here is how that goal can be achieved through dismissal of lawsuits. Given the importance of the compliance rate in determining the social desirability of litigation, this will come down to choosing dismissal rules that support compliance incentives.

a. Merit of Claims and Social Costs of Litigation

The first step in linking dismissals to an optimal litigation procedure regime is to note that the percentage of potential injurers that comply with the law should, in general, be a function of the *merit* of lawsuits. Low merit lawsuits will target with equal or greater likelihood injurers who complied the law than injurers who did not. As a result, low merit lawsuits will provide relatively weak incentives for potential injurers to comply with the law. In general, the injurer's incentive to comply depends on the difference between his expected liability when he complies and when he does not. As this difference increases, the injurer's incentive to comply increases. Since the difference between the injurer's expected liability when he complies and when he does not declines as the merit level of lawsuits declines, so does the incentive to take care.¹⁹ This implies that the population percentage of law-compliers (represented in this model by the variable w) is an increasing function of the merit of the typical lawsuit.

¹⁹ This is a simple version of a more complicated model; see Keith N. Hylton, *Costly Litigation and Legal Error Under Negligence*, 6 J. L. ECON. & ORG. 433 (1990), which can be used to formalize this intuitive argument. Specifically, let r_1 be the multiple of liability imposed on a violator of the law and r_2 be the multiple imposed on a non-violator. If the court holds injurers liable only when they are found to have failed to comply with the law (e.g., take care), the potential injurer will take care whenever $x < [(r_1)^{3/4} - (r_2)^{1/4}]$. Since increasing merit causes r_2 to decline relative to r_1 , it also enhances incentives to take care and comply with the law.

In addition, low-merit claims widen the difference between the social and private gains from complying with the law. For litigation to have the potential to improve welfare, the incremental liability that results from failing to comply with the law should be at least as great as the incremental cost to society that results from failing to comply. However, if low merit lawsuits are capable of imposing substantial costs on defendants who complied with the law, the incremental liability for failing to comply may be miniscule. One way in which low merit lawsuit may impose substantial costs on defendants who complied with the law is if courts fail to use observable merit proxies to distinguish low merit lawsuits and dismiss them early.

A simple numerical example may clarify these points. Suppose that an injurer's failure to comply with the law imposes a cost of on a victim of \$100 with certainty. If the injurer complies with the law, no costs will be imposed on any victim. Suppose the injurer will be held liable (with certainty) for \$100 when he fails to comply. Suppose, also, that the injurer will be liable, because of low-merit lawsuits, for \$80 when he complies (this amount reflects litigation expenses and the expected value of an adverse judgment). Under these conditions, the incremental liability that results from failing to take care is only \$20. It should be clear that this weakens the injurer's incentive to comply in comparison to a system in which the incremental liability is greater. Moreover, since the social cost of failing to comply is \$100, the injurer should – from society's perspective – spend up to \$100 to comply. But if the incremental liability is only \$20, the injurer will spend up to \$20 for compliance and no more.

Thus, even if the final judgments issued by courts were perfectly aligned with the merit of the claims that are litigated, a failure to dismiss low merit claims leads to two

types of social cost. First, by permitting substantial litigation costs to be imposed on complying defendants, failures to dismiss low merit claims weaken incentives to comply with the law and to take socially desirable actions. Second, by directly increasing total litigation costs, failure to dismiss low merit claims reduces social welfare.

b. A Proxy for Merit

Since the plaintiff has the burden of proof,²⁰ the plaintiff's prediction of prevailing, P_p , is a natural choice as an observable proxy for the merit of his lawsuit. It is the plaintiff who must frame the case for the court, and set forth the facts and legal theories that justify a judgment in his favor. The plaintiff's work enables an objective observer to assess the plaintiff's likelihood of victory, based on publicly available information and that revealed by the plaintiff to support his claim. Of course, the plaintiff's prediction of prevailing is assumed in this analysis to be based on objective evidence that can be demonstrated to a court.²¹

Admittedly, the defendant has the burden to plead defenses, and these certainly have a substantial impact on the plaintiff's likelihood of victory. However, the plaintiff will anticipate defenses and adjust his prediction of prevailing to take their effect on the likely trial outcome into account. Given this, the plaintiff's prediction of prevailing, P_p , is the best candidate to serve as a measure of the merit of the typical lawsuit.

The foregoing implies that there is a critical level of merit below which lawsuits should be barred. Specifically, since the percentage of law-compliers, w , is in general an increasing function of the merit of the typical lawsuit, the percentage of law-compliers

²⁰ For an economic justification for the allocation of the burden of proof, see Hay & Spier, *supra* note 8.

²¹ If the plaintiff is rational, this is not a restrictive assumption. A rational plaintiff will base his prediction of victory on the evidence that can be demonstrated to a court, not on irrational prejudices or factual assertions that cannot be demonstrated (e.g., such as a belief that people who wear red shoes must always be guilty).

should increase as the merit proxy P_p increases, and conversely. If the merit level falls to the point at which the deterrence benefits are below the total costs of litigation, social welfare could be enhanced by barring lawsuits. In terms of the simple model presented earlier, if we allow \bar{P}_p to be the level of merit at which the deterrence benefit is just equal the total litigation costs (i.e., the left and right hand sides of (5) are equal), suit should be dismissed whenever the plaintiff's prediction of prevailing is below the critical threshold \bar{P}_p . When a plaintiff files a claim that appears to fall below that critical level of merit, the court should dismiss the claim.

Moreover, this argument implies that the critical merit threshold will increase as the costs of false convictions rise relative to the costs of false acquittals, and as the costs of litigation increase. In other words, the threshold level of merit below which a lawsuit should be barred increases with the ratio of false conviction to false acquittal costs and the costs of litigation.

C. Multi-stage Litigation and Abusive Suits

As Baxter noted,²² the standard model fails to capture the multi-stage nature of litigation. Litigation consists of several motions, some of which (e.g., the dismissal motion) can put an end to the case, while others alter its direction sharply.

The first model to show how the multi-stage nature of litigation changes the results of the standard one-period approach is that of Bebchuk.²³ In the Bebchuk model,

²² WILLIAM BAXTER, *THE POLITICAL ECONOMY OF ANTITRUST: PRINCIPAL PAPER 14-16* (Robert D. Tollison ed., 1980).

²³ Lucian Bebchuk, *A New Theory Concerning the Credibility and Success of Threats to Sue*, 25 J. LEGAL STUD. 1 (1996).

the multi-staged nature of litigation results in the prosecution of suits that appear to have a negative expected value.

The key reason negative-expected-value suits are prosecuted in the Bebchuk model is because early-stage litigation expenditures become sunk costs as litigation progresses.²⁴ Thus, once a plaintiff has reached a late stage of litigation, his incentive to continue depends on the prospect of winning and the prospective (forward looking) cost of litigation.

For example, suppose the cost of litigation to the plaintiff is \$60 in each of two periods (the total cost to the plaintiff is \$120). The expected value of the judgment to the plaintiff is \$100. This appears to be a negative expected value suit. However, at the start of the second period, the plaintiff's profit from continuing the litigation is \$40. Given this, the plaintiff has a credible threat of maintaining the lawsuit after the first stage. Suppose, in view of that credible threat, the defendant is willing to settle at the start of the second period for \$70. Given this settlement amount at the beginning of the second period, the plaintiff's threat to sue at the start of the first period becomes credible.

One lesson from the Bebchuk model is that the multi-stage nature of litigation can generate frivolous or abusive claims. For the present, I will loosely define an abusive claim as one with such a low probability of victory on the merits that the plaintiff's primary purpose for bringing the suit is to take advantage of the possibility of an erroneous decision in his favor.

D. Information and Sunk-Cost Effects in Litigation

²⁴ *Id.* at 6.

In addition to the “sunk-cost” effect identified by the Bebchuk model, multi-stage litigation also includes information effects.²⁵ As the litigants move from one stage to another, the information on the plaintiff’s probability of winning changes.

Consider a two-stage litigation model. The plaintiff’s loss (and damage award) is \$100, the first period litigation costs are \$30 for both plaintiff and defendant, and the second period litigation costs are \$35 for both parties. Given that first period costs are sunk, at the start of the second period the plaintiff’s claim remains credible if $P_p^2(\$100) > \35 . Again, given sunk costs, the parties will reach a settlement if $(P_p^2 - P_d^2)(\$100) < \$35 + \$35$. Suppose that settlement amount is S^2 . Following Bebchuk, suit will be filed at the beginning of the first period if $S^2 > \$30$. Given that a credible suit will be filed at the beginning of the first period, the defendant has an incentive to settle at the same time. The settlement incentive for the defendant will increase as the defendant’s first period litigation cost increases. These are straightforward implications of the Bebchuk model.

However, there is one difference between the case just explored and the Bebchuk model. In this version, the plaintiff’s probability of victory changes over time. It is only the second stage probability of victory that matters to the plaintiff’s incentives to file. Thus, a plaintiff could file a suit with a virtually zero chance of victory based on information available at the start of stage 1. Such a suit would appear to be abusive in the sense loosely defined above.

This suggests a different approach to assessing abusiveness. Abusiveness should be determined based on information available at the start of the second period of litigation. Thus, if the plaintiff’s prediction of prevailing at the start of the second period, P_p^2 , is below some critical threshold, the suit should be considered abusive, even if the

²⁵ See BONE, *supra* note 8, at 38.

plaintiff's prediction of prevailing at the start of the suit, P_p^1 , is above that threshold. In a setting in which the likelihood of victory changes over time, credibility should be determined by the prediction of victory at the final stage.

In addition, the notion of a positive-expected-value lawsuit should also reflect the final period likelihood of victory, based on information available at the start. The expected value of the judgment should be assessed on the basis of the end-period expectation of the plaintiff's prediction of prevailing given information available at the start of the first period – in formal terms $E(P_p^2 | P_p^1)$. Thus, if we consider the expected likelihood of victory at the end of the second period, based on information available at the start of the first, the lawsuit just examined has a positive expected value when $E(P_p^2 | P_p^1) \$100 > \$30 + \$35$. Credibility, which should also be determined by the end-period expectation of victory, may exist even when this positive-expected-value condition is not satisfied.²⁶

E. Pleading Standards and Dismissal: Pleading Versus Summary Judgment

In this part I will focus on dismissal in the multi-stage model. Assume now that the first stage of litigation is the pleading stage, and the beginning of the second stage is the summary judgment moment. The analysis in this part builds on the preceding models and delivers several rules for pleading-stage dismissals.²⁷

²⁶ Indeed, in the two period case just discussed, credibility exists when $S^2 > \$30$. S^2 will be determined by final period predictions and costs. For example, if the defendant moves first with a take-it-or-leave-it settlement offer, his rational offer is $S^2 = P_p^2(\$100) - \35 .

²⁷ The reader who prefers to avoid the technical parts of the model may prefer to skip directly to the four rules at the end of this part.

Let us now define an abusive suit as one in which the plaintiff's prediction of prevailing at the start of the second period is below some critical threshold level.

Formally, an abusive suit is one in which

$$P_p^2 < \tau, \quad (6)$$

where τ is the threshold level of merit below which a suit should not be permitted to survive a motion for summary judgment. The threshold merit level should be based on the objective of minimizing social costs,²⁸ or, equivalently, on the goal of minimizing the sum of false-acquittal, false-conviction, and litigation costs. It follows that the threshold merit level should increase as litigation becomes less productive as a deterrent and more costly to society.

The question examined in this section is the dismissal standard a court should apply to the pleading stage. Given the merit threshold necessary at the summary judgment stage, a court should dismiss at the pleading stage if

$$E(P_p^2 | P_p^1, \Psi) < \tau, \quad (7)$$

where Ψ is the additional information supporting the plaintiff's claim that, based on objective evidence, appears likely to be introduced into the trial during the first period (i.e., discovery). In other words, if, given information available at the pleading stage, the expected merit level of the lawsuit at the summary judgment stage is below the threshold at which dismissal should occur, the plaintiff's lawsuit should be dismissed at the pleading stage. Pleading stage dismissals are socially desirable under this rule because they enable courts to reduce overall social costs without having to further increase the

²⁸ See *supra* Part II.B.

merit threshold. In plain terms, they permit courts to throw out some bad claims early in order to permit more good claims to survive the end-stage dismissal test.²⁹

To gain useful insights from the pleading stage dismissal criterion, consider a case in which there are two merit levels possible at the final (summary judgment) stage. One is the same level of merit at the pleading stage, P_p^1 . This is the pessimistic scenario in which the plaintiff finds nothing to enhance the merit of his claim after engaging in discovery. The other merit level is the *discovery-enhanced merit level* ψ , where $\psi > P_p^1$. Suppose the likelihood of reaching the discovery-enhanced merit level ψ is α (i.e., using expression (7), $\Psi = \{\alpha, \psi\}$). Since $E(P_p^2 | P_p^1, \Psi) = \alpha\psi + (1 - \alpha)P_p^1$, a case should be dismissed at the pleading stage whenever

$$\alpha < \frac{\tau - P_p^1}{\psi - P_p^1}. \quad (8)$$

This implies several rules for dismissal at the pleading stage.

1. *Dismissal at the pleading stage should occur only in those cases in which the claims and evidence asserted at the pleading stage are insufficient to meet the merit requirement at the summary judgment stage.*³⁰

2. *Given that the pleadings are insufficient to meet the summary judgment merit requirement,³¹ a claim should be allowed to pass the pleading stage only if the discovery-*

²⁹ One might argue that some bad claims will get better if they are allowed to survive in court longer with more evidence coming in to support them. That possibility is taken into account in the “discovery-enhancement” term Ψ in expression (7). However, it is also possible that bad claims could get worse as more information enters. If, for example, a jury is more likely to err once a certain type of evidence is put before them – e.g., evidence that the plaintiff is a particularly sympathetic victim of bad luck – then allowing a claim to survive beyond the first stage results in greater litigation costs and a potential *reduction* in the accuracy of the trial.

³⁰ In terms of the model, this is equivalent to requiring $P_p^1 < \tau$.

³¹ That is, given $P_p^1 < \tau$.

*enhanced merit level is unambiguously greater than the summary judgment merit requirement.*³²

*3. Again, assume the pleadings are insufficient to meet the summary judgment merit requirement. If the discovery-enhanced merit level is barely sufficient to satisfy the summary judgment merit threshold, the plaintiff's pleadings must indicate that it is virtually certain that he will produce that enhanced level of merit in order to be allowed past the pleading stage.*³³

4. Since the summary judgment merit requirement increases with the social cost of the type of litigation initiated by the plaintiff, pleading-stage dismissals should occur more often for more costly claims. In particular, as the ratio of false conviction to false acquittal costs increases, and as the costs associated with litigation increase, dismissals should occur more frequently at the pleading stage. For example, if the plaintiff's claim imposes relatively high costs on the defendant, say by severely damaging his business or by imposing exorbitant discovery costs, the threshold level of merit should be correspondingly high.

This analysis of pleading-stage dismissal implies a surprisingly stingy approach on the part of courts. In any case in which there is considerable doubt as to whether the plaintiff will be able to survive summary judgment, because the potential quantum of evidence in support of the plaintiff will at best approximate the minimum needed to survive summary judgment, the courts should dismiss at the pleading stage.

³² In formal terms, given that $P_p^I < \tau$, a claim should pass beyond the pleading stage only if $\psi > \tau$.

³³ In formal terms, where ψ is only slightly greater than τ , α must be close to one for the claim to be allowed beyond the pleading stage.

E. Circumstantial and Direct Evidence

To further explore the implications of this analysis of pleading-stage dismissal, suppose the level of merit at the summary judgment phase is a function of two types of evidence, circumstantial and direct. Circumstantial evidence, as the name implies, is evidence that creates a strong inference in favor of the proposition it supports, while falling short of direct support.³⁴ For example, footprints of the same size as the defendant's provide circumstantial evidence that the defendant was at the scene of an accident or a crime. Direct evidence, on the other hand, is evidence that more or less demonstrates the proposition it supports.³⁵ For example, a videotape showing the defendant at the scene of an accident or a crime provides direct evidence in support of the plaintiff's claim.

The foregoing analysis implies that the court's quickness to dismiss at the pleading stage, for any given summary judgment dismissal standard, should vary with the relative probabilities of direct and circumstantial evidence entering the case. If the probability of obtaining direct evidence is high, a court should be more willing to allow the claim to proceed beyond the pleading stage. For example, suppose the plaintiff is likely to obtain through discovery a document or eyewitness testimony that conclusively

³⁴ See, e.g. *Department of Economic Development v. Arthur Andersen & Co. (U.S.A.)*, 924 F.Supp 449, 474 (S.D.N.Y. 1996) ("Circumstantial evidence is evidence that tends to prove a disputed fact whose existence follows inferentially from the existence of evidentiary facts..."); MARK REUTLINGER, EVIDENCE: ESSENTIAL TERMS AND CONCEPTS 28 (1996) (Circumstantial evidence is "[e]vidence that tends to prove a fact not directly, but by proof of some intermediate fact, from which the fact ultimately to be proved can be inferred").

³⁵ See, e.g. *Burell v. Board of Trustees of Georgia Military College*, 125 F.3d 1390, 1393 (M.D.Ga. 1997) ("Direct evidence is evidence, which if believed, proves existence of fact in issue without inference or presumption.") (internal citation and quotation marks omitted); REUTLINGER, *supra* note 34, at 28 ("Direct evidence...is evidence that tends to show the existence of a fact without the intervention of proof of any other facts, or the need to draw any inferences from such intermediate facts.")

links the defendant to an accident or a crime. This analysis suggests the court should be relatively lenient in allowing such claims to pass the pleading stage.

On the other hand, if the likelihood of direct evidence is low, and the level of merit likely to result from circumstantial evidence is just below the level needed to meet the summary judgment threshold, the court should dismiss at the pleading stage. Under these assumptions, the foregoing pleading-stage dismissal analysis implies that the plaintiff should be almost certain to meet the circumstantial evidence requirement at the summary judgment stage in order to avoid dismissal at the pleading stage. However, if the plaintiff were almost certain to avoid dismissal at the summary judgment stage, he should be able to indicate that level of merit at the pleading stage. Given this, a claim that fails to demonstrate that level of merit at the pleading stage should be dismissed.

F. Facts and Law

The pleadings could be insufficient to meet the summary judgment merit requirement on legal or factual grounds. The pleadings might be insufficient on factual (evidence) grounds for several reasons. The plaintiff may fail to plead facts that are necessary to infer a violation of the law. Alternatively, the plaintiff may have asserted facts that could serve as circumstantial evidence of a violation, but the facts may be insufficient in light of the summary judgment standard.

Under the predictive approach of this model, pleading facts necessary to infer a violation of the law should be a minimal requirement at the pleading stage, in order to determine that the claim will satisfy the summary judgment threshold. For most legal claims, there are facts that must be found in order to determine that a violation has

occurred. If a plaintiff fails to plead such facts, the court has no basis on which to assess the likelihood that the plaintiff will satisfy the summary judgment threshold. Consider the general differences between negligent, intent-based, and conspiratorial tort claims. A plaintiff in a negligence suit based on a car accident may fail to plead any facts indicating substandard driving on the part of the defendant. A plaintiff in an intentional tort suit may fail to plead facts indicating deliberate conduct on the part of the defendant. A plaintiff in a conspiracy claim may fail to plead facts indicating that the defendants' actions were not coincidental or independent. In each of these examples, the claim should be dismissed at the pleading stage because it fails to meet minimal predictive requirements.

The pleadings might be insufficient on legal grounds for several reasons. For example, the plaintiff might assert all of the relevant facts and legal theories necessary to infer a violation of the law. However, the plaintiff may be mistaken on the law – as in the case of a nuisance plaintiff who asserts a claim based on an interference with his right to light and air.³⁶ Alternatively, the plaintiff may assert a detailed set of facts that fail to indicate by themselves a violation of the law and without presenting legal theories that would allow a court to predict whether there is a valid legal claim.

The approach of this model, which uses the summary judgment standard as an anchor, is most useful for assessing evidentiary insufficiency at the pleading stage. Legal insufficiency is uncontroversial as a basis for dismissing a claim at the pleading stage.

III. Application to Pleading Law

³⁶ *E.g., Fontainebleau Hotel Corp. v. Forty-Five Twenty-Five, Inc.*, 114 So. 2d 357 (Fla. App. 1959).

This analysis suggests that there should be a close relationship between the summary judgment dismissal standard and the pleading stage dismissal standard. The summary judgment dismissal standard is a function of the legal standard for liability and the facts offered to support the plaintiff's claim that the legal standard will be satisfied. This has immediate implications for existing views on the proper standard for dismissal at the pleading stage.

A. Liberal Versus Conservative Views of Pleading Requirements

On one hand there are liberal views, such as the “aggregate of operative facts” and “primary right” theories.³⁷ Under the operative-facts theory, proponents assert that plaintiffs should be able to pass the pleading stage if they have asserted all of the facts that might lead to a valid claim for liability, even if they have not linked those facts to specific legal theories.³⁸ This view appears to be too liberal under this paper's analysis because if the facts asserted fail to allow one to predict what will happen at the summary judgment stage, then the claim should be dismissed at the pleading stage.

The other liberal approach to pleading focuses on the primary right of the plaintiff: for example, whether there has been a violation of the plaintiff's right to exclusive use of his property.³⁹ This approach also appears to be too liberal. Asserting the violation of a primary right at the pleading stage may not be enough to allow a court to predict the outcome of a summary judgment motion.

³⁷ FRIEDENTHAL, KANE & MILLER, *supra* note 1, at 259-260. *See also* Silas A Harris, *What is a Cause of Action?*, 16 CAL. L. REV. 459 (1928).

³⁸ Charles E. Clark, *The Code Cause of Action*, 33 YALE L. J. 817 (1923).

³⁹ JOHN NORTON POMEROY, *REMEDIES AND REMEDIAL RIGHTS BY THE CIVIL ACTION ACCORDING TO THE REFORMED AMERICAN PROCEDURE* 460-461 (4th ed. 1904).

The conservative position is the “theory of the pleadings” doctrine, which requires the plaintiff to tailor his assertions at the pleading stage to satisfy a specific legal theory.⁴⁰ This is closer to the approach suggested by this paper. Of course, it might be too conservative, if applied in a manner that does not take into account the possibility that there could be more than one legal theory.

The standard suggested by this paper is a predictive test that asks whether the facts and theories presented at the pleading stage are likely, in light of the evidence that will enter, to satisfy the summary judgment requirements – where close calls are assessed in light of the policies behind the summary judgment requirements. This suggests that pleading stage requirements should vary with the demands of summary judgment stage requirements. Certainly, specific legal theories will be required at the summary judgment stage, as well as facts to support those legal theories. Where the summary judgment standard is relatively high, in terms of the factual support required, the pleading stage requirements should be relatively high. Conversely, where the summary judgment stage requirements are relatively low, the courts should be liberal at the pleading stage. Moreover, if courts erect a demanding summary judgment standard, the implicit standard of proof applied to the pleading stage claims should also be demanding, and conversely.⁴¹

This does not, at least in theory, lead to the “man-traps” that Bentham once complained had riddled the common law,⁴² causing plaintiffs to lose valid claims because

⁴⁰ FRIDENTHAL, KANE & MILLER, *supra* note 1, at 261.

⁴¹ Suppose, for example, the courts have adopted a demanding summary judgment test, in response to a valid perception that false conviction costs are high relative to false acquittal costs. It follows under this paper’s approach that assessments of evidence should be influenced by the same policies behind the summary judgment standard. Thus, if false conviction costs are high relative to false acquittal costs, doubts should be resolved against the plaintiff. Conversely, if false acquittal costs are high relative to false conviction costs, doubts should be resolved against the defendant.

⁴² JEREMY BENTHAM, A COMMENT ON THE COMMENTARIES: A CRITICISM OF WILLIAM BLACKSTONE’S COMMENTARIES ON THE LAWS OF ENGLAND 163-164 (Charles W. Everett ed., 1928).

they had failed to assert some special combination of words at the pleading stage. The theory implied here suggests that an objective observer, presumably the court, should attempt to forecast the legal theories and summary judgment standards associated with those theories. This approach would not penalize a plaintiff for failing to assert a claim that was clearly implied by the facts set out in the pleadings.

B. Understanding the Law on Pleading

We can understand the law on pleading standards in light of the foregoing analysis. And by “the law,” I refer to the way courts have dealt with pleadings more than the particular rules set out in the civil procedure codes. The procedural codes reflect broad judgments about the desirability of litigation that may not be reflected in some of the narrow, case-specific decisions on pleading.⁴³

Discrimination claims appear to have relatively light pleading requirements.⁴⁴ This is consistent with the theory of this paper because the summary judgment stage requirements are relatively low. At the summary judgment stage, a discrimination plaintiff will either have direct evidence or circumstantial evidence of discrimination.

If the plaintiff has direct evidence of discrimination, such as a memo in which an employer asserts a plan to discriminate against a certain class of employees, he will most likely win his case and clearly have enough to survive a summary judgment motion. Some courts have defended the relatively low pleading threshold in discrimination cases on the basis of this possibility, but that is unpersuasive. Direct evidence is unlikely to

⁴³ For a discussion of the conflict between the broad judgments behind procedural code reform and case-specific decisions, see Richard L. Marcus, *The Revival of Fact Pleading Under the Federal Rules of Civil Procedure*, 86 COLUM. L. REV. 433 (1986).

⁴⁴ See *Swierkiewicz v. Sorema*, 534 U.S. 506 (2002); see also Michael Zimmer, *The New Discrimination Law: Price Waterhouse is Dead, Whither McDonnell Douglas*, 53 EMORY L.J. 1887, 1933-1934 (2004).

appear, and the prospect of discovering it would provide a basis for allowing every claim to continue beyond the pleading stage.

The more likely scenario is the case in which the plaintiff will have only circumstantial evidence at the summary judgment stage. However, the circumstantial evidence required in discrimination cases is not particularly difficult to amass and relatively easy to interpret.⁴⁵ For example, if the plaintiff can show that similar employees not within the same protected class were treated differently (better) by the employer, that may be sufficient to survive a summary judgment motion.

However, the discrimination claim is a relatively new species within the population of legal claims. The best evidence to support the theory of this paper is likely to be found in the common law treatment of pleading requirements. Two cases in which the common law imposed special pleading requirements are fraud and defamation.

In the case of fraud, the common law required pleading of specific facts detailing the fraudulent act.⁴⁶ One justification typically provided is that fraud amounts to an allegation of immorality.⁴⁷ The other traditional justification is that an allegation of fraud could serve as the basis for avoiding contracts, deeds, and similar documents.⁴⁸ For example, a contracting party might refuse to pay after a delivery of widgets. When the widget supplier sues for payment, the purchaser could assert a claim of fraud against the

⁴⁵ Zimmer argues that there are many “different types of evidence that can help form the basis for a reasonable factfinder to draw [an] inference of discrimination” such as unequal treatment, subjection to stereotyping, failure on the part of the employer to follow its own policies, or statistical evidence regarding an employer’s employment policies and practices and their relation to diversity in the workplace. Zimmer, *supra* note 44, at 1934-1937. In such cases “if there is sufficient evidence in an individual discrimination case that would support a reasonable jury’s decision to draw an inference of discrimination, the defendant’s motion [for summary judgment] should be denied.” *Id.* at 1938.

⁴⁶ FRIEDENTHAL, KANE & MILLER, *supra* note 1, at 272. *See also* CLARK, CODE PLEADING, *supra* note 11, at 311-312, 617; JOSEPH H. KOFFLER & ALISON REPPY, COMMON LAW PLEADING 194 (1969).

⁴⁷ FRIEDENTHAL, KANE & MILLER, *supra* note 1, at 273; CLARK, CODE PLEADING, *supra* note 11, at 311-312.

⁴⁸ FRIEDENTHAL, KANE & MILLER, *supra* note 1, at 273.

supplier. Given the potential damage to commerce if claims of this sort dragged on in court without any serious effort to weed frivolous from valid charges, the specificity requirement developed in the common law governing fraud actions seems desirable.

In terms of this paper's theory, the specificity requirement adopted under the common law for fraud actions is an example of courts raising the merit threshold in order: (1) to ensure that a claim admitted into court was likely to meet the evidentiary requirements of the legal standard and (2) to reduce the frequency of socially wasteful litigation. As legal hornbooks note, the law imposed high evidentiary burdens on plaintiffs in fraud actions, requiring them to prove misrepresentation of fact, knowledge of falsity, intent to deceive, reliance, and substantial injury.⁴⁹ Where the legal standard imposes a heavy burden on plaintiffs, the pleading standard should impose a relatively high burden.

In addition to the difficulty of meeting the legal standard, the social cost of fraud litigation provides another reason for imposing a heightened pleading standard. There are several costs clearly connected to fraud claims. One is the cost of litigation, which is associated with any legal claim. Another is the reputation cost imposed on the defendant. In markets in which reputation matters greatly, an allegation of fraud could severely damage a business. And another cost is the disruption in commerce. If courts refused to screen out frivolous claims of fraud, businesses would find it more difficult to enforce valid contracts, forcing them to rely on alternatives such as long-term relationships or demanding property to be held as security.

⁴⁹ See, e.g., JOSEPH M. PERILLO, CALAMARI AND PERILLO ON CONTRACTS 333-346 (5th ed. 2003).

The common law also required specificity in pleading for defamation actions.⁵⁰ This is also consistent with this paper's analysis. Here, the problem is not the burden imposed on the class of potential defendants. In the case of defamation claims, as Holmes noted long ago,⁵¹ society has an interest in making sure that they do not discourage speech. Defamation claims have had to meet relatively high proof standards. The rigorous pleading standard is a reflection of the relatively high proof standard and the perceived social cost of cheap defamation litigation.

C. Pleading in Antitrust

Like discrimination claims, antitrust is also a relatively new area of litigation. Because of its relatively recent birth, one cannot examine the common law treatment of pleading in antitrust claims in order to test whether the theory set out earlier in this paper provides a justificatory account of long-settled law. Moreover, the dismissal standards at summary judgment and pleading stages are still in the process of development in antitrust.⁵²

Recognizing the social costs of false convictions, antitrust courts have raised the evidentiary requirements to survive a summary judgment motion with respect to many claims. *Brooke Group* establishes a recoupment test for predatory pricing claims that

⁵⁰ FRIEDENTHAL, KANE & MILLER, *supra* note 1, at 299-300; *see also* CLARK, CODE PLEADING, *supra* note 11, at 315-316; KOFFLER & REPPY, *supra* note 46, at 195-196.

⁵¹ OLIVER WENDELL HOLMES, JR., THE COMMON LAW 138-140 (Little, Brown, and Company 1881).

⁵² For a recent independent analysis of antitrust dismissal standards that reaches conclusions similar to those in this part, see Richard A. Epstein, *Motions to Dismiss Antitrust Cases: Separating Fact From Fantasy*, AEI-Brookings Joint Center Working Paper (March 2006), <http://aei-brookings.org/publications/abstract.php?pid=1059>. On the general point that antitrust dismissal standards sometimes reflect general perceptions of the social costs of litigation, see Stephen Calkins, *Equilibrating Tendencies in the Antitrust System, with Special Attention to Summary Judgment and Motions to Dismiss*, in PRIVATE ANTITRUST LITIGATION: NEW EVIDENCE, NEW LEARNING 205 (L. White, ed., 1988).

plaintiffs find extremely difficult to satisfy at the summary judgment stage.⁵³ The courts have recognized the high costs of permitting erroneous claims of predation to work their way all the way to a jury.⁵⁴ The resale-price-maintenance cases have established high proof standards for claims that reach the summary judgment stage.⁵⁵ Here the law reflects the increasing recognition that resale price maintenance is not, as a general matter, socially harmful, and is in many instances socially beneficial.⁵⁶ The conspiracy case law has imposed a “plus factors” requirement at the summary judgment stage that requires plaintiffs to provide evidence suggesting that the defendants’ conduct could not be explained by independently-motivated action.⁵⁷

In many of these areas of antitrust litigation, courts have not worked out precisely what should be required at the pleading stage. In other words, summary judgment tests have imposed increasingly high burdens on plaintiffs in antitrust, while pleading standards have remained unclear in many jurisdictions. This opens the possibility of high merit thresholds required at the summary judgment stage coupled with low merit thresholds at the pleading stage.

⁵³ *Brooke Group Ltd. v. Brown & Williamson Tobacco Corp.*, 509 U.S. 209 (1993).

⁵⁴ *Matsushita Elec. Indus. Co. v. Zenith Radio*, 475 U.S. 574, 594 (1986) (internal citations omitted).

⁵⁵ *Monsanto Co. v. Spray-Rite Service Corp.*, 465 U.S. 752 (1984).

⁵⁶ F.M. SCHERER & DAVID ROSS, *INDUSTRIAL MARKET STRUCTURE AND ECONOMIC PERFORMANCE* 548-558 (3d ed. 1990). Keith N. Hylton, *Antitrust Law: Economic Theory & Common Law Evolution* 257-260 (2003); Lester G. Telser, *Why Should Manufacturers Want Fair Trade?* 3 J. L. & Econ. 86 (October 1960); Victor Goldberg, *The Free Rider Problem, Imperfect Pricing and the Economics of Retailing Services*, 79 Nw. U. L. Rev. 736 (1984); J.R. Gould & L.E. Preston, *Resale Price Maintenance and Retail Outlets*, 32 *Economica* 302 (1965); Patricia B. Reagan, *Resale Price Maintenance: A Reexamination of the Outlet Hypothesis*, 9 Res. Law & Econ. 1 (1986).

⁵⁷ See, e.g., ERNEST GELLHORN, WILLIAM E. KOVACIC & STEPHEN CALKINS, *ANTITRUST LAW AND ECONOMICS* 277-282 (Thomson/West, 5th ed. 2004). For an alternative to the “plus factors” view, see Judge Posner’s opinion in *In re High Fructose Corn Syrup Antitrust Litigation*, 295 F.3d 651 (7th Cir. 2002). Citing his view that Section 1 of the Sherman Act prohibits tacit collusion, Judge Posner would permit courts to follow a totality-of-the-circumstances approach to assessing circumstantial evidence at the summary judgment stage. One implication of this paper’s analysis is that as leniency programs bring in more direct evidence in conspiracy cases, fewer cases will require an examination under the plus-factors test.

One case in which this uncertainty was resolved is *Bell Atlantic v. Twombly*.⁵⁸ The plaintiffs asserted that the incumbent “baby-bell” telephone monopolies, also known as Incumbent Local Exchange Carriers (ILECs), conspired to: (1) prevent competitive local exchange carriers (CLECs) from entering into their markets and competing against them, as encouraged by access requirements of the Telecommunications Act of 1996, and (2) avoid entering each other’s markets in order to compete.⁵⁹

The district court examined the plaintiff’s complaint for evidence that it was likely to satisfy the summary judgment standard, which requires the provision of evidence of plus factors (evidence that the alleged conspirators were acting against individual self-interest). The district court found that all of the evidence asserted by the plaintiff’s complaint was entirely consistent with individually-motivated profit-seeking conduct. The district court dismissed the complaint at the pleading stage.⁶⁰

The Second Circuit reversed on the ground that the Federal Rules of Civil Procedure do not impose heightened pleading standards for antitrust claims.⁶¹ The court noted that heightened standards exist for fraud, mistake, and defamation claims; and asserted that outside of those narrow categories all that is required is that the complaint be sufficient to give notice of the claims the plaintiff intends to prove in court.⁶²

⁵⁸ 127 S.Ct. 1955 (2007).

⁵⁹ *Id.* at 1962.

⁶⁰ *Twombly v. Bell Atlantic Corp.*, 313 F.Supp.2d 174 (S.D.N.Y. 2003).

⁶¹ *Twombly v. Bell Atlantic Corp.* 425 F.3d 99, 112 (2d Cir. 2005).

⁶² The Second Circuit reasoned that the Federal Rules of Civil Procedure “set forth a pleading standard under which plaintiffs are required to give the defendant fair notice of what the...claim is and the grounds upon which it rests. *Id.* at 107 (internal citation and quotation marks omitted). The court defined fair notice as “that which will enable the adverse party to prepare for trial, allow the application of res judicata, and identify the nature of the case so that it may be assigned the proper form of trial.” *Id.* at 108 (internal citation and quotation marks omitted). The Second Circuit then stated that “the Supreme Court noted that it had previously declined to extend heightened pleading requirements to other contexts beyond fraud or mistake. *Id.* at 107 (internal citation and quotation marks omitted) Thus, the Second Circuit found only the liberal requirements of notice pleading must be satisfied in antitrust cases and “[t]o survive a motion to

The Supreme Court reversed the Second Circuit and held that the federal procedure rules require allegations sufficient to raise the “right to relief above the speculative level,”⁶³ and that it was therefore necessary for the plaintiffs to provide evidence of plus factors at the pleading stage.⁶⁴

This paper’s theory supports the standard adopted by the Supreme Court (and the district court). There are several reasons. First, the merit threshold has already been set high for parallel action claims at the summary judgment stage in order to reduce the frequency and costs of false convictions. Failing to adjust the pleading standard is inconsistent with this policy. Maintaining a low merit standard at the pleading stage undermines the restrictive policy embodied in the high summary judgment standard by encouraging the very lawsuits that the high summary judgment standard aimed to discourage.

Second, a low pleading standard permits the plaintiff to impose discovery costs on the defendants, which are generally understood to be a heavy burden in antitrust cases. The discovery costs are large enough to encourage defendants to settle claims even when the plaintiff’s claim is highly likely to fail. In the case of large class action for treble damages, high discovery costs coupled with only a one percent chance of victory at trial could force defendants to pay substantial settlements.

What, precisely, are the false-conviction costs associated with conspiracy claims based on circumstantial evidence? Suppose firms in an oligopolistic industry are not colluding, but they have reduced prices in parallel fashion (i.e., roughly simultaneous and

dismiss...an antitrust claimant must allege only the existence of a conspiracy and a sufficient supporting factual predicate on which the allegation is based.” *Id.* at 114.

⁶³ *Twombly*, 127 S.Ct. at 1965.

⁶⁴ *Id.* at 1965-1966.

similar conduct). A low proof standard for pleading encourages plaintiffs to file suit whenever a case, however weak, can be made that harmful price-fixing may have occurred. Realizing this, the potential defendant would be well advised to avoid increasing price in a parallel fashion, since that would generate price-fixing claims. But if firms are aware that lawsuits will follow parallel price increases, then they will also know that it is risky to engage in price competition, since upward price movements, which inevitably must occur if price competition is really vigorous, will be followed by lawsuits. The result is that the litigation threat discourages the vigorous competition that the law aims to encourage.

The third reason supporting the Court's decision requiring the plaintiffs to present evidence at the pleading stage that indicates that they will be able to satisfy the summary judgment standard is that the best that they could hope for was to satisfy the summary judgment standard. Given this, the plaintiffs should have been required to show that they are highly likely to satisfy that standard. The likelihood of finding direct evidence of a conspiracy is low, and would require combing through warehouses full of internal documents. There is no obvious upside to the proof that the plaintiffs might bring in these cases, which distinguishes them from most employment discrimination cases. In parallel-action conspiracy cases the standard types of circumstantial evidence are insufficient to meet the summary judgment standard, while the standard types of circumstantial evidence are sufficient in discrimination cases. Parallel-action conspiracy claims therefore depend heavily on the hope of finding direct evidence, which is highly unlikely.

The pleading standard demanded by the Supreme Court in *Twombly* recognizes what common law courts have recognized for many years; that pleading standards should be adjusted to take into account the requirements of the legal standard and the social costs of litigation. Pleading standards for fraud were set relatively high under the common law for these reasons. The economic model in this paper suggests that pleading standards in antitrust should be shaped by the same factors.

IV. Conclusion

This paper applies a simple economic framework to the choice between pleading and summary judgment as points at which a claim can be dismissed. It concludes generally that pleading standards should vary with the evidentiary demands of the associated legal standards and the social costs of litigation. The common law's imposition of higher pleading standards for fraud claims is consistent with this proposition. It also implies that the rigorous summary judgment standards developed in antitrust courts with respect to claims of predation, resale price maintenance, and conspiracy should be paired with rigorous pleading requirements.

Appendix

This appendix presents a more formal version of some of the arguments of the text. I will focus on the single-period litigation model here. The results for the two-period model are straightforward extensions.

The model below puts a great deal of weight on the plaintiff's prediction of success in a lawsuit. It assumes that this prediction is based largely on case-specific facts rather than general information about population compliance.

Standard Model of Litigation

The standard model of litigation is a one period model that focuses on the filing and settlement decisions. Let P_p equal the plaintiff's prediction of success. A lawsuit will be filed when

$$P_p > \gamma_p, \quad (\text{A1})$$

where γ_p is the ratio of the plaintiff's litigation cost to the damage award. Condition (A1) can be understood as saying that a suit will be filed when the plaintiff's prediction of success exceeds the ratio of his cost of litigation to his damages.

Let P_d equal the defendant's prediction of the likelihood of a verdict in the plaintiff's favor. Settlement in the standard model is described by the "Landes-Posner-Gould" condition, according to which settlement occurs if and only if

$$(P_p - P_d) < \gamma_p + \gamma_d \quad (\text{A2})$$

Specifying the Goal of Procedure

Let θ_{nc} equal the probability of an injury when the potential defendant/injurer does not comply with the law (e.g., does not take care). Let θ_c equal the probability of an injury when the potential defendant does comply with the law (takes care). Let x equal the cost of care per dollar of damages. Assume x is a random variable distributed according to the cumulative distribution function G . Let w equal the percentage of potential injurers who take care because of the threat of liability.

When lawsuits are prohibited, no one complies, so total social cost (recall that the damage amount is equal to 1) is

$$\theta_{nc} \tag{A3}$$

When lawsuits are permitted, the fraction w of potential injurers comply, so social cost is

$$(1-w)\theta_{nc} + w(\theta_c + x) + [(1-w)\theta_{nc} + w\theta_c](\gamma_p + \gamma_d) \tag{A4}$$

Suit is socially desirable, then, when⁶⁵

$$w[(\theta_{nc} - \theta_c) - x] > [(1-w)\theta_{nc} + w\theta_c](\gamma_p + \gamma_d). \tag{A5}$$

This means that a suit is socially desirable when the deterrence benefits (injuries avoided net of avoidance costs) exceed the litigation costs. Since γ_p , γ_d , θ_{nc} , θ_c , and x are fixed in this model, procedural rules can impact the social benefit from litigation most directly by supporting the compliance probability w . Specifically, (A5) implies that procedural rules should be structured to support an outcome in which

$$w > \frac{\theta_{nc}\gamma}{((\theta_{nc} - \theta_c)(1 + \gamma) - x)}, \tag{A6}$$

⁶⁵ See Steven Shavell, The Social Versus the Private Incentive to Bring Suit in a Costly Legal System, 11 J. Legal Stud. 333 (1982).

where $\gamma = \gamma_p + \gamma_d$. (A6) defines a threshold level of w below which lawsuits no longer enhance social welfare.

Determining an Optimal Dismissal Policy

The private incentive to comply is determined by a comparison of the cost of compliance with the reduction in total liability that results from compliance. Let r_{nc} and r_c be the probability of being held liable in the noncompliance and compliance states. Let s_{nc} and s_c be the share of litigation expense borne in the noncompliance and compliance states. The potential injurer will comply with the law if and only if

$$(\theta_{nc}r_{nc} - \theta_cr_c) + (\theta_{nc}s_{nc} - \theta_cs_c)\gamma_d > x, \quad (\text{A7})$$

i.e., if the incremental liability of failing to comply exceeds the cost of compliance.

Given that the cost of compliance x is distributed according to the distribution function G , the percentage of actors that comply with the law is simply

$$w = G((\theta_{nc}r_{nc} - \theta_cr_c) + (\theta_{nc}s_{nc} - \theta_cs_c)\gamma_d). \quad (\text{A8})$$

Merit and Incentives to Comply

Here I will introduce the concept of merit in a lawsuit. Low merit claims target the complying actors almost as severely as the non-complying actors. Thus, the incremental liability for failing to comply falls as the merit of lawsuits falls. In terms of the model, this means the ratios r_{nc}/r_c and s_{nc}/s_c decline as the level of merit falls.

Let m represent the level of merit in a lawsuit against a particular actor. From the foregoing, $r_{nc}/r_c = f(m)$ and $s_{nc}/s_c = h(m)$, where $f' > 0$ and $h' > 0$.

The central claim of this model is that the plaintiff's prediction of success serves as the best available proxy for merit. This implies that the plaintiff's prediction of success is equal to the merit measure plus an error term: $P_p = m + \varepsilon$. Using P_p as the only observable measure of merit, the fraction of actors who comply with the law can be expressed as

$$w = G(r_c(\theta_{nc}f(P_p - \varepsilon) - \theta_c) + (\theta_{nc}h(P_p - \varepsilon) - \theta_c)\gamma_d). \quad (\text{A9})$$

Since f and h are both increasing functions, the percentage of law-compliers, w , is also an increasing function of the plaintiff's prediction of success. (A9) and (A6) together define a threshold level of P_p below which lawsuits should be dismissed. By dismissing lawsuits in which P_p is below the cut-off implicitly defined by (A6) and (A9), courts maintain the average level of merit at a level that supports incentives to comply.

It is important to note that P_p is determined by case-specific facts rather than general information about average compliance rates in the population. If P_p were determined entirely by average compliance rate, we would have the equality $P_p = 1 - w$, which would clearly contradict the model in this appendix. However, if the parties have case-specific facts that allow them and courts to distinguish compliers and non-compliers, P_p will no longer be determined by $1 - w$. Under this assumption, dismissal of cases with low P_p is equivalent to dismissing cases in which the plaintiff has failed to provide case-specific facts pointing to liability.