2012

The Problem of Measuring Legal Change, with Application to Bell Atlantic v. Twombly

William Hubbard

CHICAGO

The Problem of Measuring Legal Change, with Application to *Bell Atlantic v. Twombly*

William H. J. Hubbard

**THE LAW SCHOOL
THE UNIVERSITY OF CHICAGO**

September 2011

The Problem of Measuring Legal Change, with Application to *Bell Atlantic v. Twombly*

William H.J. Hubbard[*]

whubbard@uchicago.edu

July 11, 2011

Abstract

Measuring legal change—i.e., change in the way that judges decide cases—presents a vexing problem. In response to a change in the behavior of courts, plaintiffs and defendants will change their patterns of filing and settling cases. Priest and Klein's (1984) selection model predicts that no matter how favorable or unfavorable the legal standard is to plaintiffs, the rate at which plaintiffs prevail in litigation will not predictably change; thus, legal change cannot be measured with data on court outcomes. In this paper, I extend the selection model to develop a methodology for measuring legal change, even in the presence of selection effects. I apply this methodology to a recent, high profile Supreme Court case, *Bell Atlantic Corp. v. Twombly.* My model generates novel predictions, which are confirmed in the data, and I find that *Twombly* caused no legal change, even after accounting for possible selection effects.

1 Introduction

How does one measure legal change? There is of course a formal sense in which the law changes every time Congress passes a new statute, or the Supreme Court decides a new case. This paper, however, concerns itself with legal change in a different, more functional sense: a change in the way judges decide cases. It is this latter sense in which legal change affects the expectations that individuals form about the risks of liability for different activities, and is thus the primary means by which the law influences human behavior.[1] It is legal change in this sense, therefore, that is of primary interest to the policymaker.

This poses the question whether a given Supreme Court opinion or new statute actually changes how courts decide cases. To answer this question, we would like to see whether cases that in the past had been decided one way are now being decided differently. The problem is that this comparison requires us to hold fixed the set of litigated cases before and after the potential law change. But in response to a legal change, plaintiffs and defendants may change their litigation strategy, settling cases that previously would have been litigated, and litigating cases that previously would have settled. The rate at which plaintiffs or defendants prevail in litigation may not change, even after a sharp change in how courts decide cases. This is the problem of measuring legal change.

Priest and Klein (1984) were the first to develop this argument. Their model essentially argues that because trying a case is costly, parties have an incentive to settle before trial. If the plaintiff and the defendant have similar beliefs about the expected outcome of the trial, then they are both better off settling for (approximately) the expected amount of damages and saving their attorney fees. It is only when the plaintiff and the defendant have sharply divergent expectations about the litigation that settlement becomes impossible. And

---

[1]Holmes (1897) put it colorfully: "The reason why [law] is a profession, why people will pay lawyers to argue for them or to advise them, is that in societies like ours the command of public force is intrusted to the judges in certain cases .... People want to know under what circumstances and how far they will run the risk of coming against what is so much stronger than themselves, and hence it becomes a business to find out when this danger is to be feared."

since the parties' expectations are most likely to diverge in "close" cases—cases where the defendant is on the borderline between liability and non-liability—we should expect that the cases that do go to trial will be close cases under whatever legal standard governs.[2] Therefore, Priest and Klein predict that (subject to important qualifications) plaintiffs will prevail approximately 50 percent of the time in litigation and that (*not* subject to qualifications) a change in a legal standard will have no predictable effect on the rate at which plaintiffs or defendants prevail in litigation.

Take the recent Supreme Court case *Bell Atlantic Corp. v. Twombly* (550 U.S. 544 [2007]). In *Twombly*, the Supreme Court addressed the standard under which a judge decides whether to dismiss a case based on the allegations in the plaintiff's complaint. This rule applies in every federal civil case, and thus applies to approximately 250,000 newly filed cases per year. The standard for dismissing a case had been announced in the 1957 Supreme Court case *Conley v. Gibson* (355 U.S. 41 [1957]), and for fifty years, the Supreme Court had never altered the standard. But in *Twombly*, the Supreme Court unexpectedly "retired" language from *Conley* and announced a new standard that made it easier to dismiss a case.

There was an immediate uproar from the bar and legal academe. "No decision in recent memory has generated as much interest and is of such potentially sweeping scope as the Supreme Court's decision in *Bell Atlantic Corp. v. Twombly*." Smith (2009). Countless commentators criticized its (presumably) harmful effect on plaintiffs seeking relief in court.[3] Bills were introduced in Congress to overturn the decision. Notice Pleading Restoration Act of 2009, S. 1504, 111th Cong. (2009); Open Access to Courts Act of 2009, H.R. 4115, 111th Cong. (2009). Lawyers and academics testified that *Twombly* has had a "devastating impact" on plaintiffs, particularly in civil rights cases. Schneider (2010).

But was *Twombly* in fact such a momentous legal change? For that matter, did it effect

---

[2]Hylton and Lin (p. 1) credit Holmes (1881, p. 127) as the first to argue this point: "Legal, like natural divisions, however clear their general outline, will be found on exact scrutiny to end in a penumbra or debatable land. This is the region of the jury, and only cases falling on this doubtful border are likely to be carried far in court."

[3]See Part 2.2 for citations.

3

any legal change at all? How can we know? If the Priest and Klein model is right, we might expect to see plaintiffs prevailing on a motion to dismiss about 50 percent of the time before *Twombly*, and about 50 percent of the time after *Twombly*—even if *Twombly* radically changed the pleading standard. In a sample of nearly 13,000 published decisions on motions to dismiss, this is exactly what I find. See Table 1. And if the Priest and Klein model is right, then our enterprise, the empirical measurement of legal change, is doomed.

Well, not quite. In this paper, I will argue that a refinement of the Priest and Klein (or "selection") model may allow us to distinguish empirically a change in a legal standard from no change.[4] This paper takes two initial steps toward that goal.

First, although styled "The Selection of Disputes for Litigation," Priest and Klein (1984) did *not* address the selection of *disputes* (e.g., car crashes or broken promises) for *litigation* (i.e., filed lawsuits). Instead, Priest and Klein present a model of the selection of *filed lawsuits* for *trial*. The literature to date has largely followed suit, either treating the set of filed lawsuits as fixed and exogenous (e.g., Priest and Klein 1984; Waldfogel 1995; Siegelman and Waldfogel 1999) or assuming that all suits with positive expected value are filed (e.g., Wittman 1985; Priest 1985; Siegelman and Donahue 1995).[5] But the rate at which cases are filed is an outcome of interest, and it is an empirical reality that many disputes are settled before a lawsuit is ever filed. Hence, the first contribution of this paper is an application of the selection model to filings.

Second, the Priest and Klein model distinguishes only between "trial" and "settlement," making no allowance for the stages of litigation that precede trial. There are at least two points in the litigation process, prior to trial, at which the judge may decide the case in favor

---

[4]The original Priest and Klein model has survived virtually unchanged since it was first published twenty-seven years ago. Hylton and Lin (2010) note that "[a]lthough the Priest and Klein hypothesis is widely cited in the law and economics literature and has been tested empirically, it still lacks formal treatment.[ ] The original Priest-Klein article provides an informal argument, as do later articles testing it." Waldfogel (1995, 1998b) formalizes some aspects of the model; Hylton and Lin (2010) themselves formalize Wittman's (1985) version of the selection model. This is a testament to the model, as is the large empirical literature it has spawned. (See Waldfogel [1998b] and Hylton and Lin [2010] for citations to this literature.)

[5]By positive expected value, I mean the expected judgment award to the plaintiff exceeds the plaintiff's expected litigation costs.

of the plaintiff or the defendant. Early in the case, a defendant may file a motion to dismiss (as in *Twombly*), and after discovery either party may file a motion for summary judgment. Presumably, the selection process in the Priest and Klein model applies to such dispositive motions as well.[6] Surprisingly, however, virtually all of the empirical work on the Priest and Klein model has focused on trials rather than other types of adjudication.[7] No prior work has directly addressed the interaction between the selection of disputes for filing and the selection of filed lawsuits for adjudication. Thus, the second contribution of this paper is to generalize the selection model to account for multi-stage litigation by nesting the process of selecting lawsuits for adjudication within the process of selecting disputes for filing.

I will refer to this generalization of the selection model as the "nested selection model." In this model, the process of selection of filed lawsuits for litigation of a motion to dismiss is nested within the larger process of the selection of disputes to become filed lawsuits. At each stage of the dispute-resolution process, some disputes are resolved by a settlement between the parties, while other advance to further litigation. The nested selection model attempts to capture this dynamic. It not only replicates many of the predictions of the original selection model, but generates addition predictions, which I validate empirically.

With this nested selection model, I return to the empirical problem of measuring legal change. To hold the selection of disputes for filing fixed while allowing adjudication to respond endogenously to a legal change, I exploit the fact that the stages of litigation (e.g., filing the suit, litigating a motion to dismiss) are separated in time by weeks or months. Because of this fact, the nested selection model, unlike the original Priest and Klein model, makes unambiguous predictions of the effects of a legal change on observable outcomes.

In particular, the nested selection model generates predictions that allow one to test

---

[6]Indeed, Priest and Klein (1984, p. 30) conjecture that their model "will explain the outcomes of preliminary, but contested, issues of procedure or substance."

[7]I will follow Waldfogel (1995) and call the resolution of a case through either dispositive motion or trial "adjudication." Note that studies such as Eisenberg (1991) and Waldfogel (1998a) compare plaintiff success rates before trial and at trial, although they do not apply the selection model to the pre-trial stages of litigation.

for legal change, even in the presence of the selection effects described by the Priest and Klein model. Importantly, the nested selection model is itself a testable model. Further, one can test the predictions of the nested selection model using familiar empirical methods, such as linear regression, and only a small innovation is required in the collection of court data—collection of filing date information for all cases.

I then apply this theoretical framework to *Twombly*. As noted above, most commentators have assumed that *Twombly* was a major legal change and therefore precipitated a rise in both the rate at which motions to dismiss are granted and the total number of dismissals. I will call this view the "conventional hypothesis." A minority of commentators have taken the position that *Twombly* caused at most a modest change in how judges decide cases, and therefore its effects on dismissals should be modest. I will call this the "contrarian hypothesis." Neither of these views, nor the empirical literature attempting to measure the effect of *Twombly*, accounts for selection effects in litigation.

In this paper, I examine the effects of *Twombly*, while accounting for selection effects with the nested selection model. If *Twombly* did change how district courts decide cases, then the nested selection model makes a very specific prediction: for cases filed before *Twombly* but adjudicated after *Twombly*—those cases for which the selection of disputes for filing is held fixed—the rate at which motions to dismiss are filed will rise, but the rate at which they are granted will not change (much). This, in turn, implies that dismissals as a percent of all cases will rise. I will call this the "selection/change hypothesis." The nested selection model in the *absence* of legal change, however, is indistinguishable from the contrarian view—nothing changes.

To test these predictions, I bring two new datasets to bear. These datasets have two important advantages over previously used datasets. First, they are very large. One is a set containing information on outcomes for nearly 13,000 published district court opinions ruling on motions to dismiss between May 21, 2006 and May 21, 2008 (a year before and

after *Twombly*). The other is a set of administrative data on every federal civil lawsuit filed since 1978. My analysis will focus on a subset of cases filed in the 2005-2008 period (over 250,000 observations). Second, I have the filing date for every case in these datasets, allowing me to estimate effects separately for cases filed before *Twombly* but adjudicated after *Twombly*. Using these two datasets, I test the conventional, contrarian, and selection hypotheses, and find no support for the view the *Twombly* has a large effect, under either the conventional hypothesis or the selection/change hypothesis. Instead, the data support the view that *Twombly* did not change the legal standards applied by district courts.

In sum, this paper makes two contributions. First, and most broadly, it provides an approach to measuring legal change that accounts for the ways in which actors in the legal system respond endogenously to changes in the legal environment. Second, and more narrowly, it helps resolve a contentious debate about the most controversial civil procedure decision in a generation: *Bell Atlantic v. Twombly.*

The balance of this paper is ordered as follows: Part 2 describes *Twombly* and briefly addresses the large literature that it has spawned. In Part 3, I list the central empirical predictions made about the effects of *Twombly.* I then develop the nested selection model and define my methodology for testing the competing hypotheses. Part 4 describes the two data sets I use, their advantages and their limitations. Part 5 presents my estimation specifications and results. I find no evidence that *Twombly* marked a change in the law of pleading. The most "sweeping" Supreme Court decision "in recent memory," it appears, has had little or no effect on how judges decide cases. Part 6 concludes with a discussion of the limitations of and the opportunities presented by the nested selection model.

2 The Story So Far

2.1 From *Conley* to *Twombly*

Civil actions in federal court begin with the filing of a complaint by the plaintiff. This pleading "is the key to the courthouse door." (Steinman 2010, p. 1295). Federal Rule of Civil Procedure ("Rule") 8(a) governs the content of a complaint: "A pleading that states a claim for relief must contain . . . a short and plain statement of the claim showing that the pleader is entitled to relief." Once a plaintiff files a complaint, the defendant can file a motion to dismiss for "failure to state a claim upon which relief can be granted" under Rule 12(b)(6).[8] If the court denies the motion to dismiss, the case proceeds to discovery and perhaps settlement or trial. If the court grants the motion, the case is over; the plaintiff's only recourse is appeal to a higher court.

For 50 years, the standard governing a Rule 12(b)(6) motion seemed unassailable. *Conley v. Gibson* (355 U.S. 41, 45-46 [1957]), described "the accepted rule that 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." On its face, this statement seems to imply that a complaint would survive a motion to dismiss so long as the plaintiff did not plead facts that contradicted his legal claim. While perhaps never interpreted quite so literally, *Conley* reflected a liberal approach to the principle of "notice pleading" embodied by the Federal Rules.[9]

So began five decades of apparent tranquility in the law of pleading. In this period, "pleading standards were widely viewed as 'well established' and 'relatively straightforward'"

---

[8]Herein, "motion to dismiss" refers specifically to a Rule 12(b)(6) motion to dismiss; of course, there are other grounds for dismissal, such as lack of jurisdiction or failure to prosecute the claim, which are not relevant here.

[9]Though remote from my analysis, the story of the regime of "fact pleading" that preceded the adoption of the Federal Rules of Civil Procedure and the development of simplified pleading systems, including the Field Code in the 19th Century, that culminated in the adoption of the Rules in 1938 may be of interest to the reader. For various accounts, see Marcus (1986, pp. 433-444); 550 U.S. at 573-576 (Stevens, J., dissenting); Stancil (2009, pp. 109-114); Campbell (2010, pp. 1196-1218); Dodson (2010, pp. 447-452).

(Steinman 2010, p. 1295). Pleading was so well settled that "commentators lost interest in pleading" (Marcus 1986, p. 434). And as recently as *Swierkiewicz v. Sorema N.A.*, 534 U.S. 506 (2002), and *Leatherman v. Tarrant County Narcotics Intelligence and Coordination Unit*, 507 U.S. 163 (1993), the Supreme Court had unanimously rejected efforts to require pleading of specific facts, even in narrow categories of cases.[10]

And then, on May 21, 2007, the Supreme Court decided an antitrust dispute captioned *Bell Atlantic Corp. v. Twombly* (550 U.S. 544 [2007]). William Twombly and Lawrence Marcus had sued the "Baby Bells"—the four Incumbent Local Exchange Carriers (ILECs) that control nearly all local telephone service in the United States—alleging a class-action claim of conspiracy in violation of the Sherman Act. Their complaint alleged that the ILECs had engaged in "parallel conduct" to thwart competition and raise prices for local phone and high-speed internet service. The factual allegations in support of this claim were essentially the parallel conduct of each ILEC in resisting entry by new, competing carriers into its market area, and in not attempting entry into its neighbors' markets. (544 U.S. at 549-551.)

The Court began by holding that allegations of parallel conduct, even if taken to be true, would not establish a claim under the Sherman Act. It then turned to the question whether the complaint should therefore be dismissed. To the Court, this was the hard question. After all, *Conley* did not require that the factual allegations of a complaint, if true, prove a legal claim; it simply required that the allegations, if true, not *disprove* the plaintiff's legal claim:

> *Conley v. Gibson* [355 U.S. 41 (1957)] spoke ...of 'the accepted rule that 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. This 'no set of facts' language can be read in isolation as saying that any statement revealing the theory of the claim will

---

[10]I should note that unlike the Supreme Court, the lower federal courts periodically made efforts to push pleading standards higher. Marcus (1986) and Fairman (2003) document these efforts in detail. Fairman argues, though, that after *Leatherman* and *Swierkiewicz*, the lower courts retreated from their more aggressive interpretations of the pleading standard. See Fairman (2003, pp. 1008, 1018, 1025, 1030, 1034, 1040).

suffice unless its factual impossibility may be shown from the face of the pleadings. . . . On such a focused and literal reading of Conley's 'no set of facts,' a wholly conclusory statement of claim would survive a motion to dismiss whenever the pleadings left open the possibility that a plaintiff might later establish some 'set of [undisclosed] facts' to support recovery.

(544 U.S. at 561 (citations omitted).) Treating the absurdity of this result as self-evident, the Court announced that "after puzzling the profession for 50 years, this famous observation has earned its retirement." (544 U.S. at 563.) In its place, the Court required that a complaint plead facts "plausibly suggesting (not merely consistent with)" the plaintiff's legal claim. (544 U.S. at 557.) It concluded: "we do not require heightened fact pleading of specifics, but only enough facts to state a claim to relief that is plausible on its face. Because the plaintiffs here have not nudged their claims across the line from conceivable to plausible, their complaint must be dismissed." (544 U.S. at 570.)

The Supreme Court itself would reaffirm *Twombly* almost exactly two years later. On May 18, 2009, it decided *Ashcroft v. Iqbal* (129 S. Ct. 1937 [2009]), which presented a civil rights claim brought by an Arab Muslim detained by the FBI in the aftermath of September 11, 2001. The Court elaborated on *Twombly*, saying, "Two working principles underlie our decision in *Twombly*. First, the tenet that a court must accept as true all of the allegations contained in a complaint is inapplicable to legal conclusions. Threadbare recitals of the elements of a cause of action, supported by mere conclusory statements, do not suffice. . . . Second, only a complaint that states a plausible claim for relief survives a motion to dismiss." (Id. at 1949-50.)

The Court also noted that *Twombly* was not limited to antitrust cases, but rather "expounded the pleading standard for all civil actions." (Id. at 1953 (internal quotation marks omitted).) This confirmed the conclusion that both courts and academics had already reached about the scope of *Twombly*: it applied to all cases. See Bone (2009, p. 881);

Hatamyar (2010, p. 555); Redish and Epstein (2008, p. 26).

2.2 The Aftermath

Twombly was a bombshell. The viability of *Conley* was not part of the question presented to the Supreme Court; it had not even been raised by the parties. *Twombly* struck "[s]eemingly without warning" (Hatamyar 2010, p. 554) and "sent shockwaves through the legal community—for academics, practitioners, and judges alike" (Steinman 2010, p. 1305 (citations omitted)).[11]

The response to *Twombly* was immediate. Almost overnight, it became a fixture in judicial opinions,[12] and after 50 years of near-dormancy, the scholarly literature on pleading exploded.[13] As Reinert (2010, p. 3) notes, a near-consensus emerged "among academic observers that the *Iqbal/Twombly* pleading standard marks a sharp break with the past."[14] Most scholars view *Twombly* "as overturning fifty years of generous notice pleading practice, and critics attack 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.'" (Bone 2009, p. 875.) Spencer (2008, p. 431) sums up the conventional wisdom: "Notice pleading is dead."

---

[11]Other characterizations include "out of the blue" (Epstein 2009, p. 4); "a startling move by the U.S. Supreme Court" (Spencer 2008, p. 431); and "a surprising departure from ingrained federal pleading rules" (Smith 2009, p. 1063).

[12]*Twombly* has been cited in judicial opinions tens of thousands of times and is already one of the most cited decisions in the history of the United States. Steinman (2010, p. 1295 n.9) notes that "[a]s of March 2010, *Twombly* had been cited in nearly 24,000 federal decisions—already number seven of all time." Ward (2008, p. 893) notes that "*Bell Atlantic Corp. v. Twombly* was decided by the U.S. Supreme Court on May 21, 2007, and has already been cited more than 9,400 times as of March 15, 2008."

[13]Stancil (2009, pp. 137-138) writes: "Until recently, the scholarly literature on pleading standards was remarkably thin, with only a few significant pieces written from the 1930s through the early 2000s. Widespread scholarly interest in pleading is a remarkably recent phenomenon, tracing its birth to the Supreme Court's 2007 opinion in *Bell Atlantic Corp. v. Twombly*." For a litany of citations, see Steinman (2010, pp. 1296-1298 and nn.10-14).

[14]Other formulations include "a revolution in pleading " (Schneider 2010, p. 527); a "drastic departure from well-established pleading standards" (Hamburg and Koski 2010, p. 10); "a new era" (Steinman 2010, p. 1310); "a sea change" (Hoffman 2008, p. 1235); "a major departure from the Court's established pleading jurisprudence" (Miller 2010, p. 16); and even a "subversion[ ] of law to achieve the restrictive ends of societal elites" (Spencer 2010, p. 201).

Whether this development is a good thing or a bad thing has been the center of contentious debate. "Without question, *Bell Atlantic v. Twombly*[ ] ranks as one of the most controversial decisions of the United States Supreme Court in recent years." (Epstein 2009, p. 2.) At base, normative arguments about *Twombly* stem from a single, empirical premise: rightly or wrongly, *Twombly* will cause more cases to be dismissed—a lot more cases. Miller (2010, pp. 21) sums up that "the perception among many practicing attorneys and commentators is that the grant rate [for motions to dismiss] has increased, particularly in civil rights cases, employment discrimination, private enforcement matters, class actions, and proceedings brought pro se."[15]

Although *Twombly* was an antitrust case, the greatest concerns raised have been about its effect on civil rights plaintiffs. Reinert (2010) and others note the historical importance of federal courts to civil rights claimants, as well as the particular difficulty that individuals with claims such as employment discrimination will have pleading specific facts about the defendant's motives. Some argue that the need to plead facts establishing a "plausible" claim will often prove an "impossible" (Kilaru 2010, p. 924) or "insurmountable" (Schnapper 2009; Wasserman 2010, p. 178) barrier for civil rights plaintiffs. This has lead to predictions that *Twombly* will disproportionately increase dismissals of civil rights claims (see, e.g., Ward 2008; Wasserman 2010) and that "the shift is undeniably big" (Schneider 2010).

Of course, there have been dissenting voices. Some scholars have argued that, as a doctrinal matter, *Twombly* does not represent a sharp break with precedent (Epstein 2009; Steinman 2010), or is only a modest change in doctrine (Bone 2009). These commentators go on to predict that, consequently, the observed effect of *Twombly* will not be dramatic. This contrarian view has its share of adherents outside the academe. Some practitioners such as Andrew Pincus (2010), an attorney at Mayer Brown, have argued that *Twombly* "is less a

---

[15]In most of the literature, this prediction is perhaps so obvious that it goes unstated; but many make it explicit. See Ward 2008, p. 916 ("Post-*Twombly*, a defendant would predict a higher rate of success on [a Rule 12(b)(6)] motion."); Hannon 2008, p. 1814 ("Generally, any substantive alteration to the pleading standard would have an effect on the dismissal rate under 12(b)(6).").

sea change as it is a recognition of what was already going [on] out there in the trenches."

What is remarkable about this debate is that despite widespread disagreement about the meaning of *Twombly* as a matter of legal doctrine, all sides of the debate (implicitly) agree that whether or not *Twombly* worked a major change in the law will ultimately be resolved as an *empirical* matter: we can measure its effect on the law of civil procedure by measuring how it changes the rate at which cases are dismissed. In other words, this debate implicitly defines legal change as this paper defines it: a change in "the law" is a change in how district courts decide cases. Many more dismissals implies the law changed; no change in dismissals means there was no change in the law. Under this definition of legal change, a new Supreme Court precedent may cause legal change—or it may not. With this notion of legal change in mind, scholars have rushed to compare dismissal rates among published cases decided before and after *Twombly*. (See Hannon 2008; Seiner 2009, 2010; Hatamyar 2010.) Others call for further empirical work. (See Hoffman 2008; Wasserman 2010; Miller 2010.) I now turn to a brief discussion of the existing empirical literature.

2.3   The Evidence to Date

A small literature has emerged that attempts to test whether *Twombly* has led to an increase in motions to dismiss being granted, with particular attention to civil rights claims. These studies are virtually the first attempts to study pleading and motions to dismiss empirically. Chiorazzi et al. (1988, p. 120) compile a massive annotation of 316 empirical studies in civil procedure, but note that "there are no studies in pleading per se, perhaps because of the advent of modern notice pleading."[16] Indeed, For the period before 2008, I know of only one empirical study directly addressing the use of motions to dismiss (Willging 1989).[17]

Hannon (2008) is the first empirical study of *Twombly*. Hannon looks at district court opinions published on Westlaw. He searches Westlaw for every case that cited *Twombly*

---

[16]By way of comparison, they found *five* studies on the use of telephonic conferences in civil litigation.

[17]A second, Connolly and Lombard (1980), contains some data related to motions to dismiss.

in the context of a motion to dismiss from June-December 2007, and every case that cited *Conley* during most of the period June 2006-September 2007.[18] His final sample includes 3,287 cases. Comparing the cases citing *Conley* with the set of cases citing *Twombly*, he finds that among civil rights cases, the grant rate for motions to dismiss jumps from 41.7 percent in the *Conley* group to 52.9 percent in the *Twombly* group. For all other cases, the difference is meager: 36.9 percent before versus 37.4 percent after *Twombly*.

Three subsequent academic studies have followed Hannon's methodology and reached similar results, although they varied the dates covered and used smaller samples of cases. Hatamyar (2010) and Seiner (2009, 2010) conclude that at least in the context of civil rights claims, *Twombly* is leading to more dismissals.[19]

These studies on the effects of *Twombly* are quick to acknowledge that inferences drawn from them must be qualified by the limitations of their data and methodology. First, the estimated effects are relatively small and none are statistically significant.[20] Second, any study relying on opinions ruling on motions to dismiss can at best quantify only the share of motions to dismiss granted, not the overall rate at which filed cases are dismissed. Third, nearly all the studies rely on opinions published in Westlaw, when in fact only a fraction of all district court orders are published. See Hoffman, Izenman and Lidicker (2007); Lizotte (2007). It is not known the extent to which the selection of decisions for publication may affect the representativeness of the samples drawn from Westlaw. Fourth, a related concern is that focusing only on cases citing *Conley* or *Twombly*, rather than all cases ruling on motions to dismiss, introduces an additional selection bias.[21]

---

[18]He does not explain why he includes cases decided after May 21, 2007 in his *Conley* sample, nor does he address the fact that cases after May 21, 2007 may cite both *Twombly* and *Conley*.

[19]A survey by the FJC provides some qualitative evidence of little effect. Willging and Lee (2010) survey both plaintiffs' and defense attorneys across a range of practice areas and report that "[m]ost interviewees indicated that they had not seen any impact of the two cases in their practice."

[20]Hannon (2008) reports one statistically significant result, but this is due to a specification error. A corrected regression on the same data (not reported, on file with author) yields no significant effect.

[21]District courts ruling on a motion to dismiss may cite other Supreme Court cases, precedent from the circuit courts, or even other district court opinions. Given the age of *Conley* during the sample periods in the studies above, district court opinions choosing to cite *Conley* directly, rather than any of the hundreds or thousands of more recent (and equally controlling) precedents, are unlikely to be a representative sample

The most authoritative study to date is Cecil et al. (2011), a large study of docket records from a sample of federal district courts by researchers at the Federal Judicial Center (FJC). It compares cases from 2005-2006 with cases from 2009-2010, bracketing an interval during which both *Twombly* and *Iqbal* were decided. Using a dataset of motion to dismiss filings, Cecil et al. (2011) find that motions to dismiss were filed in a larger share of cases in 2009-2010 (6.2 percent) than in 2005-2006 (4.0 percent). Using a second database of rulings on motions to dismiss, they find no change in the rate at which motions to dismiss are granted. Because of differences between the two samples, Cecil et al. (2011) could not estimate whether the fraction of all filed cases terminated on a motion to dismiss had risen.

None of these studies finds a statistically significant increase in the rate at which motions to dismiss are granted after *Twombly*. This is consistent with no change in the legal standards applied by district courts after *Twombly*. It is also consistent, however, with *Twombly* inaugurating a higher standard for motions to dismiss, and the selection of cases responding endogenously in ways that leaves unchanged the rate at which motions to dismiss are granted. These studies do not address the question of how to measure the effect of *Twombly* if there are significant effects of selection of disputes into litigation.

3 Legal Change and the Nested Selection Model

Both the praise and the scorn directed at *Twombly* are a consequence of the belief that *Twombly* has, in fact, raised the bar for pleading. But has it? The competing accounts in Part 2 offer contrasting predictions. The conventional view proceeds from the premise that *Twombly* has raised the pleading standard and reaches these conclusions:

Prediction 1. Twombly will (significantly) increase the fraction of motions to dismiss that are granted under Rule 12(b)(6).

---

of rulings prior to *Twombly*. For cases citing *Twombly*, this bias may be less of a problem, given its recent vintage.

Prediction 2. Twombly will (significantly) increase the fraction of filed cases that are dismissed under Rule 12(b)(6).

As discussed above, the conventional account suggests that the effects of *Twombly* will be largest among civil rights cases, where evidence of elements such as discriminatory intent may not be available to plaintiffs absent discovery. Because discovery is generally not available until after a ruling on a motion to dismiss, the argument goes, many civil rights plaintiffs will be left without the ability to show that their claim is "plausible" as required by *Twombly*. Thus, for these hypotheses the conventional account would add the following corollary:

Corollary. The effects above will be (significantly) greater for civil rights cases.

The contrarian view, of course, rests on the belief that *Twombly* did not, in fact, change the pleading standard much. Hence, advocates of this view argue that any effects will be quite modest.

As noted above, these views do not account for selection effects predicted by Priest and Klein. In this Part, I will develop what I call a nested selection model, which formalizes and extends the original model described in Priest and Klein (1984). With this model, I will then generate testable predictions distinct from both the conventional view and the contrarian view.

3.1   The Nested Selection Model

I present a more formal exposition of the nested selection model in the Model Appendix. Here I present the model in an informal fashion. Consider the set of all disputes of a given type (e.g., car crashes); these disputes may or may not become filed lawsuits. Let's say we have a measure of "quality" for each dispute, such that disputes with quality above some threshold level will result in a plaintiff victory at trial (if there is a trial), and in disputes with quality below that threshold, the defendant will prevail. If we apply the logic of Priest and Klein to the selection of disputes for filing, we would predict that for cases far above or

below the victory standard, the parties will settle rather than litigate, because they agree that the dispute has a foregone conclusion. The quality of filed cases will tend to be close to the victory standard, because those are the disputes for which the parties are least likely to agree on the result. Filed cases with quality that is high or low relative to that standard will be few.

Now consider a different legal standard, the pleading standard. In the terms of the selection model, the pleading standard will be a fairly low standard relative to the victory standard. Thus, there will be relatively few filed cases that are close to the pleading standard. But if we apply the logic of the selection model to pleadings, the model predicts that only cases close to the pleading standard will have a motion to dismiss (in other cases, the parties will agree that the outcome of a motion to dismiss is a foregone conclusion, and act accordingly). This in turn means that relatively few cases will have litigated motions to dismiss (and therefore few cases will be dismissed). Figure 1 illustrates this. The dashed line indicates the distribution of the quality of filed cases, relative to the victory standard V. The pleading standard is S. The distribution of cases with a motion to dismiss is indicated by the solid line. Of this small group, those below the standard are dismissed. (This is the shaded area in Figure 1.)

I will call this model of (1) the selection of disputes for filing and (2) the selection of filed lawsuits for litigation of a motion to dismiss the "nested selection model." In Figure 1, the distribution of litigated motions to dismiss is literally nested within (and almost a miniature version of) the distribution of filed cases—itself a selected subset of the distribution of all disputes. Within the set of filed cases, a small subset, with case quality centered slightly above the pleading standard S, will litigate a motion to dismiss.[22] Within this subset, defendants will prevail somewhat less than half the time. (The shaded area in Figure 1 represents the set of dismissed cases.) Already this model generates two predictions about

---

[22]Because the distribution of filed cases has most of its mass near the victory standard V, the distribution of cases with motions to dismiss will not be symmetrically distributed around S, but will be shifted slightly toward V. Hence, more than 50 percent of cases with motions to dismiss will have quality above S.

pleadings practice: first, motions to dismiss will be filed in a relatively small fraction of cases; and second, motions to dismiss will be granted a large fraction (but less than half) the time.

Without the context of the nested selection model, these two predictions are surprising, or even contradictory: if a defendant, at relatively low cost, can get maybe a one-in-three chance of dismissing a lawsuit (that otherwise could be very costly), why would so few even bother to file a motion to dismiss? Yet the available data tend to confirm these predictions.

First, very few motions to dismiss are filed. Willging (1989) reviews three sets of district court docket records, and finds rulings on Rule 12(b)(6) motions to dismiss in only 6 to 12 percent of cases. And in a massive study of federal district court docket records, Cecil et al. (2011) find motions to dismiss filed in only 4 to 6.2 percent of cases.[23]

Second, courts frequently dismiss cases in response to motions to dismiss. In a sample of district courts, Willging (1989) finds that in rulings on motions to dismiss, the courts would grant the motion and dismiss the case 18 to 51 percent of the time. Cecil et al. (2011) find that between 23 and 28 percent of all rulings on motions to dismiss result in the termination of the lawsuit within 90 days. It is important to note here that the grant rate for motions to dismiss is usually measured in terms of whether the defendant obtained her requested relief, even if only part of the plaintiff's case was dismissed. These grant rates tend to be higher—in the order of 50 percent. See Willging (1989); Cecil et al. (2011). The nested selection model, however, abstracts away from partial dismissals and models only motions to dismiss the entirety of plaintiff's case. For this reason, I report numbers for motions to dismiss that result in the case being terminated.

My data is consistent with these findings as well. As noted above, the fraction of opinions granting a motion to dismiss in my data is approximately 50 percent. See Table 1. Of these, about 70 percent resulted in the case terminating within 90 days; all told, 34 percent of cases

---

[23]The numbers presented by Cecil et al. (2007) should be taken as a conservative estimate; their sample was limited to motions to dismiss filed within 90 days of the suit being filed. This scope should capture most, but not all, motions to dismiss.

in the sample were terminated within 90 days of a motion to dismiss being granted.[24] And yet, my administrative data indicates that perhaps only 2 percent of all cases are terminated by a motion to dismiss. See Table 3. This is consistent with Willging's finding that the fraction of all sampled cases dismissed varied from 1 to 6 percent.[25]

3.2 A Methodology for Measuring Legal Change

We now have the nested selection model. It generates some novel predictions that are consistent with observed facts (such as the rates at which motions to dismiss are granted) that otherwise would be difficult to explain. It also gives us some leverage on the problem of measuring legal change.

If one can hold the distribution of filed cases (the dashed line) fixed, but at the same time raise the pleading standard (to S'), then the nested selection model predicts that more cases will have motions to dismiss, and thus more cases will be dismissed. See Figure 2. The original selection model could not reach this prediction for two reasons. First, because it did not apply the logic of selection to the filing of lawsuits, it did not recognize that the distribution of quality among filed cases would be concentrated around the victory standard. However, this is the essential logic of the Priest and Klein model—the choice to file is dependent on the parties not settling first. Second, in the original selection model, it is impossible to hold the distribution of filed cases fixed; when the pleading standard changes, the set of filed cases (the dashed line in Figure 2) may change, and in an unpredictable way. This is because in the original selection model, all the stages of dispute resolution—pre-filing settlement, filing, motions to dismiss, discovery, summary judgment, and trial—occur simultaneously in a single period. Thus, changes in legal rules occur simultaneously with changes in the composition of filed cases.

---

[24]Focusing only on the categories of cases included in the AO Data analysis (see Part 4.1) lowers the rate of termination within 90 days to 23 percent. See the Data Appendix for details on data processing.

[25]Because of their methodology, Cecil et al. (2011) were not able to directly estimate the fraction of all sampled cases that were dismissed.

The empirical reality, however, is that litigation is time consuming and the stages of litigation are often spread over months or years. Thus, for any given change in a legal rule, there will be cases for which choices in the first stage (settlement versus filing) occur under the old legal rule, but choices in the second stage (litigating a motion to dismiss) occur under the new legal rule. Thus, in applying the nested selection model, I can hold the distribution of filed cases fixed, even as the distribution of motions to dismiss changes in response to a new pleading standard. I do this by looking at cases in which the complaint was filed before *Twombly* (and thus the filing decision was not affected by *Twombly*) but a motion to dismiss, if any, was litigated after *Twombly* (and thus the motion to dismiss decision was affected by *Twombly*).

Figure 2 illustrates this comparison. The distribution of filed lawsuits is held constant, but the dismissal standard increases from S to S'. The new set of cases with a motion to dismiss will contain many more cases because its mean, S', is closer to the mean of Y. From this, the nested selection model generates two unambiguous empirical predictions. In for the "short run," i.e., holding the distribution of filed cases fixed after the (ostensibly) new legal rule is announced, we predict the following: First, the share of all motions to dismiss that are granted will not change (much). Second, the share of all filed cases that are dismissed will rise. Together, these predictions constitute what I will call the "selection/change hypothesis," which is the joint hypothesis that the nested selection model is operating and that *Twombly* effected a (significant) change in the legal standard.

There is also the "selection/no change hypothesis," which is the joint hypothesis that the nested selection model is operating and that *Twombly* was not a (significant) legal change. This hypothesis yields predictions identical to the "contrarian" hypothesis, i.e., no significant changes in either grant rates for motions to dismiss or total numbers of dismissals.

In Table 2, I array these predictions against the predictions of the conventional hypothesis and the contrarian hypothesis. The conventional and contrarian views implicitly constitute

joint hypotheses as well; both assume the absence of selection effects. Effects on the motion to dismiss grant rate distinguish the conventional hypothesis from the other two; effects (or lack thereof) on the rate at which cases are dismissed distinguishes the contrarian hypothesis from the other two. These predicted effects are for the "short run"—i.e., for cases filed before *Twombly* but in which any motion to dismiss would have been litigated after *Twombly*.

In short, the original Priest and Klein model yielded the result that it is impossible to detect the effects of legal change in litigation outcomes. This paper shows that it is possible to identify observable effects of a change in a legal standard, even in a world in which the selection model operates.[26] I now turn to the data with which I will test these predictions.

4 Data

I have compiled two large datasets of federal district court cases. They are notable for two reasons. First, together they form the largest collection of case information yet brought to bear on the question whether *Twombly* has had an effect. Second, although other empirical studies of legal change have used either administrative data or opinions data, as I do here, prior studies in general were not able to configure their data in a way that permitted drawing inferences about changes in a legal standard under the conditions of the selection model. As I show above, the selection model generates unambiguous predictions only for a "short run" sample of cases, cases that are filed under the old standard but decided under the new standard. This requires sorting cases by filing date, and distinguishing those cases filed before the (apparent) legal change but decided in relevant part after the legal change.

Datasets of published opinions, however, inevitably are organized by the date of the opinion, not the date of case filing. (It is not even clear that a set of published opinions,

---

[26]Under some circumstances, this methodology also allow us to test the nested selection model. Support for the conventional hypothesis would undermine the nested selection model; support for the selection/change hypothesis would tend to refute the conventional hypothesis. But as all models make the same prediction (no effects) in the absence of a legal change, I cannot distinguish the contrarian view from the selection/no change hypothesis in the data.

standing alone, *could* be organized by filing date.) The administrative data, which is collected and published by the Administrative Office of the United States Courts (AO), likewise is organized by the (fiscal) year in which each case was terminated. The AO data, unlike published opinions, does provide filing dates for all cases. Thus, by combining all available annual datasets of terminations with the most recent data on pending cases, I create a database constituting a census of all cases filed on any given date (since 1978). Prior studies have relied on limited sets of annual datasets of terminated cases; this prevents reconfiguring the dataset into even a representative sample (let alone census) of cases filed on any given date.

To organize my dataset of published opinions by filing date, I merge it with my administrative data. To do this requires matching each opinion to a case in the AO data, keeping in mind that a case may have more than one published opinion (or, more likely, none at all), and that, conversely, a single opinion may decide motions pending in several (or dozens or hundreds of) related cases. To do this, I note that each case filed in the federal system is assigned, at filing, a "docket number," which is essentially a seven-digit serial number. Each federal district court is subdivided into one or more "offices," and within each office, the docket number is a unique identifier for every case ever filed. For every federal civil case, its district, office, and docket numbers together form a unique ID. Equally important, every published district court opinion contains in its case caption the name of its district, the office, and the docket number(s) of the case. By collecting and processing this information, I match filing dates to my opinion data.[27]

4.1   Administrative Data

The AO collects information on every civil case terminated in the federal court system and releases public-use versions of this data annually (on a fiscal year basis). The AO collects,

---

[27]These and other data processing steps are described in greater detail in the Data Appendix.

among other things, information on the following aspects of each civil case: the circuit, district, and office in which it was filed; the docket number under which it was filed; the date of filing and the date (if any) of termination; the "nature of suit," which is the category of legal claim (e.g., personal injury, breach of contract) chosen by the plaintiff when filing the suit; and the disposition of the suit (e.g., judgment on a jury verdict).[28]

The first year for which complete data is available is fiscal year 1979 (beginning July 1, 1978); the most recent year available is fiscal year 2009 (ending September 30, 2009). The AO has also released data on all civil cases pending as of December 31, 2008. Together, these datasets constitute a census of all federal civil cases filed or terminated from July 1, 1978 to December 31, 2008 (and all cases terminated through September 30, 2009), amounting to about 8 million observations. I will refer to this data (and subsets thereof) as the "administrative data" or "AO data."

The most relevant subset period for purposes of this paper are the cases filed from approximately 2005 through 2008. In this subset, I exclude several categories of data in order to avoid the excessive effects of outliers. These categories, and all processing of the AO Data, are described in fuller detail in the Data Appendix. Although I believe this subset of the data is the best configuration of the data for purposes of the questions posed, it is worth noting that none of these results reported herein are sensitive to any of the exclusions listed above.

Summary statistics for this slice of the AO Data appear in Table 3. One important fact jumps from this table. In the aggregate, the rates of dismissal before and after *Twombly* appear to be identical. If *Twombly* has changed courts' willingness to dismiss cases, its effect is being masked by other factors.

---

[28]The AO uses this data to generate aggregate statistics about total case filings, terminations, and other information, which it publishes annually. See http://www.uscourts.gov/Statistics.aspx. Also, the AO Data have been used in a large number of papers over the years, although most studies have relied on the published aggregate statistics rather than the underlying microdata. Eisenberg and Schlanger (2003, nn.2-17) and Hadfield (2005, n. 5) collect dozens such citations.

4.2 Opinions Data

Lee Epstein, William M. Landes, and Richard A. Posner have created a dataset of information on all federal district court opinions ruling on a Rule 12(b)(6) motion that were published on Westlaw during the two-year period around *Twombly* (May 21, 2006 to May 21, 2008). This dataset includes, among other things, information on the circuit and district in which the case was litigated, the month and year of the ruling, and indicators for whether the motion to dismiss was granted in full and for whether the case involved a civil rights claim. To these fields I added information on the office in which the case was filed, the docket number of the case, and the date of the ruling. After processing this data (which I describe in detail in the Data Appendix), this dataset contains 12,717 unique observations. I will refer to this dataset as the "opinions data." Summary statistics for the Opinions Data appear in Table 1. Again, the summary statistics reveal little change in dismissals after *Twombly*.

4.3 A Note about Limitations of the Data

These two datasets differ in the questions they can answer. The AO Data contains no information on whether the defendant filed a Rule 12(b)(6) motion. Thus, the AO Data cannot address what fraction of cases involve motions to dismiss, or what fraction of such motions are granted. It can address what fraction of all cases are dismissed under Rule 12(b)(6) (i.e., Hypothesis 2). Conversely, published opinions ruling on a motion to dismiss only exist for cases in which a party made a motion to dismiss. Thus, a sample of published opinions ruling on motions to dismiss can indicate the fraction of motions to dismiss granted (i.e., Hypothesis 1), but not the share of cases in which such a motion is filed.

These data sets address many of the limitations of existing studies. First, their huge sample sizes allow me to draw inferences with considerable precision. Second, while the Opinions Data, like earlier studies, can use rulings on motions to dismiss to test Prediction

1, the Administrative Data can test Prediction 2, which has not yet been tested.[29] Third, while the Opinions Data relies on published opinions, and thus may be subject to publication bias, the AO Data is a census of all cases.[30] Fourth, unlike all samples of published opinions in earlier studies, the Opinions Data includes all rulings on motions to dismiss, regardless of whether *Conley* or *Twombly* was cited by the judge.

Nonetheless, imperfections remain. First, neither the Opinions Data nor the AO Data provides information on how many motions to dismiss were filed. A rise in the number of motions to dismiss filed would be another likely effect of *Twombly.* Cecil et al. (2011) report a higher rate of filings of motions to dismiss in 2010 than 2006, which suggests that *Twombly* or *Iqbal* may have had an this kind of effect. Second, as Burbank (2004, p. 579) notes, "the AO does not exist to serve the interests of the research community, let alone of individual researchers, but rather those of the federal judiciary." Its purpose is to assist the AO in allocating resources to the federal districts, and the organization and coding of the case records reflects this purpose. Thus, while much of the data is organized based on legally significant categories (such as the basis for federal jurisdiction, or the nature of suit), some is not. Most importantly for my purposes, the *disposition* field does not distinguish between the two types of motions by which a case can be terminated on the merits before trial: the Rule 12(b)(6) motion to dismiss and the Rule 56 motion for summary judgment. Both fall into the disposition category "Judgment on motion before trial."[31]

To address this concern, I divide all cases coded as "Judgment on motion before trial" into two groups, based on the duration of the case at the time of termination and the prevailing party. I deem cases with durations of at least 45 and less than 225 days, and for which the prevailing party was the defendant (or prevailing party information was missing),

---

[29]Recall that Prediction 1 states that *Twombly* will increase the fraction of motions to dismiss that are granted under Rule 12(b)(6); and Prediction 2 states that *Twombly* will increase the fraction of filed cases that are dismissed under Rule 12(b)(6).

[30]The FJC has identified only one case (in the District of Oregon in 1987) that existed but which was not documented in AO Data.

[31]This category may also capture Rule 12(c) motions, which are essentially equivalent to Rule 12(b)(6) motions, but appear to be much less frequent.

to be terminated on a Rule 12(b)(6) motion; I create the dummy variable *dismissal* and set it equal to one for these cases. Cases of the same duration, but in which the plaintiff prevailed, may be Rule 12(c) dispositions or types of judgments (such as defaults). I deem cases with durations of 225 days or more to be terminated on a Rule 56 motion. I chose these boundaries to correspond approximately to the minimum and maximum amounts of time from filing of a complaint to disposition of a motion to dismiss in what might be a typical case.[32] This time range is consistent with findings of a study of docket records in eight district courts, which found that average times to ruling varied widely across districts, but the district averages all fell in the range of 63 to 176 days. See IAALS (2009); Kourlis and Singer (2009). I also test this duration range by comparing the dismissal and summary judgment rates it generates against FJC studies that have calculated Rule 12(b)(6) dismissal and Rule 56 summary judgment rates for specific years and districts. See Willging (1989); Cecil et al. (2007). The Data Appendix details the methodology and results; in short, the rates generated are comparable for both categories. These precise boundaries are, of course, arbitrary, but the results below are not sensitive to adjustments to the bounds.

4.4 A Note about Attenuation Bias

Finally, I note an issue of measurement that is particularly implicated when estimating treatment effects over short time frames. If actors anticipate a future legal change, their behavior may begin to change in response to the new legal rule *before* the rule is even implemented. (For a treatment of this issue, see Malani and Reif [2010].) In the case at hand, one could imagine that district courts, anticipating *Twombly*, began to increase dismissals shortly before the decision, recognizing that their rulings would not be reviewed on appeal until after *Twombly*. This will have the effect of attenuating any measured effect of *Twombly*, by shrinking the difference in outcomes before and after that decision.

---

[32]As a default, the defendant has 20 days to file a motion to dismiss (from the date the complaint is served; see Rule 12(a)(1)), and a court may take between one and six months to rule.

Fortunately, anticipation effects appear not to be a concern here. First, as noted above, *Twombly* was a surprise; the standard for pleading was not even raised by the parties to the case. Second, the data reveal no noticeable short-run shifts in filings or dismissals in advance of *Twombly*.

5 Results

I now present my estimation specifications and regression results for each of the predictions of the hypotheses summarized in Table 2.

5.1 Prediction 1: Motion to Dismiss Grant Rate

The first prediction of the conventional view is that the percentage of motions to dismiss that are granted will rise after *Twombly*. The contrarian hypothesis predicts no (significant) change. The selection/change hypothesis predicts no change. Thus, prediction 1 allows one to distinguish between the conventional hypothesis, on the one hand, and the contrarian and selection/change hypotheses, on the other. By itself, it cannot distinguish the contrarian hypothesis from the selection/change hypothesis.

To test this predictions, I use the opinions data. Because my focus is the "short run," I include only cases filed before May 22, 2007 (the date *Twombly* was decided). Motions to dismiss decided before May 22, 2007 represent the control group; those decided after May 22, 2007 represent the treatment group. I employ a linear probability model with the following specification:

$$Granted_i = \alpha + \beta \cdot Twombly_i + \gamma \cdot Civil\_Rights_i + \delta_i + \varepsilon_i \qquad (1)$$

Granted is an indicator for whether the motion to dismiss was granted in whole. *Twombly* is

an indicator for the opinion being decided after *Twombly*; $\hat{\beta}$ therefore estimates the effect of *Twombly*. *Civil_Rights* is an indicator for a civil rights case, and δ is a vector of circuit fixed effects. I also run regressions restricted only to civil rights cases, as well as logit regressions (reporting marginal effects). All standard errors are clustered at the district level.

Table 4 presents my regression results. Nowhere is the coefficient on *Twombly* statistically significant, and the point estimates are all very small in magnitude relative to the overall rate of granting motions to dismiss, which is essentially 50 percent. [33] Notably, if I do not restrict the sample to the "short run," but include all cases in the Opinions Data, including cases filed after *Twombly*, the effect of *Twombly* remains insignificant. And Cecil et al. (2011) reach nearly identical results using data for an even longer-run window: 2006-2010. They find that the rate at which motions to dismiss were granted (without leave to amend) did not change significantly between 2006 and 2010.

Of course, to say that the effect is not statistically different from zero is not to say that the effect is zero. One might be concerned that *Twombly* had a positive effect on the rate at which motions to dismiss are granted, but I cannot detect this effect due to an imprecise estimate. To address this concern, I also report the results of a "power analysis" that asks, what is the smallest null hypothesis of the effect of *Twombly* that the data reject? In other words, can we reject a null hypothesis that the effect of *Twombly* was an increase of (at least) 5 percent in the rate at which motions to dismiss are granted? Can we reject the null that *Twombly* had an effect of at least a 3 percent? How about a 1 percent increase?

As it turns out, my results are fairly precise. In Table 4, I report the lowest positive effect of *Twombly* that the data reject at the 5 percent level.[34] For all cases, the hypothesis that *Twombly* increased the rate at which motions to dismiss are granted by 1.27 percent is rejected. For civil rights cases, the hypothesis that *Twombly* increased the rate at which

---

[33]A difference-in-difference specification (not reported, available from author) in which *Civil_Rights* is interacted with *Twombly* to measure the differential effect of *Twombly* on civil rights cases also finds no effect.

[34]Here I employ a one-tailed test. The question posed is whether the effect of *Twombly* is at least as great as the null supposes.

motions to dismiss are granted by 1.72 percent is rejected. This result casts doubt on the conventional view.

5.2 Prediction 2: Dismissals as a Fraction of All Filings

The second prediction of the conventional view is that dismissals (cases terminated on a motion to dismiss) will rise as a share of all cases after *Twombly*. The contrarian hypothesis predicts no (significant) change. The selection/change hypothesis predicts a rise in dismissals. Prediction 2 allows one to distinguish between the contrarian hypothesis, on the one hand, and the conventional and selection/change hypotheses, on the other. Predictions 1 and 2 together allow one to separate these three hypotheses.

To test this prediction, I use the administrative data. I employ a linear probability model with the following specification:

$$Dismissed_i = \alpha + \beta \cdot Twombly_i + \gamma_i + \delta_i + \varepsilon_i \qquad (2)$$

Dismissed is an indicator for whether the case was dismissed as defined in Part 4 above. *Twombly* is an indicator for the case being in the later group (i.e., litigation of a motion to dismiss occurring after *Twombly*); $\hat{\beta}$ therefore estimates the effect of *Twombly*. There are two sets of fixed effects: γ_i is a vector of nature-of-suit fixed effects, and δ_i is a vector of circuit fixed effects.[35] I also run regressions restricted only to civil rights cases, as well as logit regressions (reporting marginal effects). All standard errors are clustered at the district level.

I include cases filed April 6, 2006 to May 21, 2006 and cases filed April 6, 2007 to May 21, 2007. All cases in this sample were filed before *Twombly*; thus this sample holds the selection of disputes into filed cases fixed. The former group litigated their motions to dismiss before

---

[35]Equation (2) has a full vector of nature-of-suit fixed effects, which Equation (1) has only a civil rights nature-of-suit dummy. This is because the Opinions Data contains only the single nature-of-suit category.

Twombly, while the latter group litigated their cases after *Twombly* (May 21 is 45 days after April 6). Thus, in the latter group of cases, the litigation of motions to dismiss may have endogenously responded to *Twombly*. This "short run" sample allows me to distinguish between the selection/change hypothesis and the contrarian (and selection/no change) view.

Table 5 presents my results of Prediction 2, including all cases. I present separate results for civil rights cases in Table 6. None of the estimated effects of *Twombly* are statistically significant, and the point estimates are extremely small in magnitude.

I then test the power of this empirical specification by determining the smallest positive effect that the data can reject. As Tables 5 and 6 indicate, my results are quite precise. In all specifications, I can reject a null hypothesis of a 0.4 percent effect—i.e., the hypothesis that at least 1 in 250 cases was affected by *Twombly*. Also notable, a "longer run" sample of cases filed up to two years before *Twombly* and up to a year after *Twombly* (not reported, available from author), yields virtually identical results.

One potential concern with this estimation approach is that it is possible that in the wake of *Twombly*, courts began to take longer to decide motions to dismiss, and by fixing the time window for a dismissal (at 45-224 days), I have biased estimates of the effect of *Twombly* downwards. If this is happening, we should expect to see post-*Twombly* dismissal rates rise relative to pre-*Twombly* rates if we lengthen the time window for dismissals, and we should also expect to see the average time from filing to dismissal rise after *Twombly*. Neither of these patterns appears in the data. See Table 7. In sum, the empirical test of Prediction 2 rejects both the conventional view and the selection/change hypothesis.

6   Conclusion

The empirical study of legal change within the courts presents a vexing problem. There is no doubt that the cases adjudicated by courts comprise a highly selected sample of all disputes;

because of this, Priest and Klein (1984) argue, a change in a legal standard as employed by the courts will have no predictable effect on observable outcomes in litigation.

The *Twombly* case presents a perfect example of this problem. While traditional doctrinal analysis of the case has led scholars to divergent views on whether it marks a significant change in the law governing pleading, both sides of the debate agree that its effect on the law should be inferable from its effect on practice in the district courts. Yet the original selection model of Priest and Klein would predict no change in the rate at which cases are dismissed, even if *Twombly* raised the bar for pleadings. If so, empirical study of the courts cannot shed light on the significance of *Twombly*.

In this paper, I show that it is possible both to account for selection effects in litigation and empirically measure the effects of legal change on the courts. I begin the project of generalizing and extending the selection model to account for multiple stages of litigation and to make the problem of measuring legal change, in the presence of selection effects, tractable. I show that fairly weak assumptions lead to a refinement of the selection model, which I call the nested selection model. This, in turn, generates new predictions about the prevalence and success rates of motions to dismiss; these predictions are borne out empirically.

This nested selection model further predicts that legal changes will have observable effects in the "short run," even if no changes can be observed in the long run, as the original model predicts. Applying this methodology to *Twombly*, I find fairly precise zeros for the effects of *Twombly* on both the grant rate of motions to dismiss and the overall rate of dismissals among filed cases. These results support the contrarian view and the selection/no change hypothesis, but tend to reject the conventional view and the selection/change hypothesis. Thus, while these results do not constitute a test of the nested selection model itself—I cannot reject both non-selection hypotheses—I can conclude that *Twombly* has had no effect even accounting for the possibility that the selection of disputes changed in response to the case.

To this result I must attach two qualifications. First, due to the limitations of my data, the longest "long run" I can examine lasts through December 31, 2008. While I am unable to observe changes in dismissals even when looking at the "long run," it is possible that the effects of *Twombly* will only show up beyond this window; rather than directly changing the behavior of district courts, *Twombly* may have touched off a period of evolution in the treatment of pleadings by district courts, which eventually led to plaintiffs facing a significantly higher bar to pleadings. The methodology presented in this paper cannot address this hypothesis.[36] Nonetheless, I can say that the effect of *Twombly*, if it was significant, was not immediate. This suggests something about the Supreme Court's control over the "lower" courts' treatment of pleadings: at best, the lower courts' response (if any) to *Twombly* was not so much a quick shift to a higher standard, as it was a period in which the lower courts negotiated a new standard for pleadings among themselves. Indeed, the *Iqbal* decision might be seen as part of this process as well.

Second, it is quite possible that while the effect of *Twombly* on broad categories of civil litigation is negligible, *Twombly* may have had an impact in a small subset of cases— indiscernable to the empirical analysis used in this paper but perhaps important nonetheless. The opinion in *Twombly* expressed grave concerns about the cost of litigating low-merit, high-stakes class actions. Perhaps if *Twombly* tipped the balance in favor of granting a motion to dismiss in only a handful of cases that, by some criteria, fell into this category, this was exactly the effect the Supreme Court intended. Of course, it is also possible that *Twombly* has had dramatic effects on very different types of cases; in either event, such an effect, if it exists, would have to be concentrated among a very small number of cases.

With these caveats, this paper strongly rejects the view the *Twombly* constitutes a major change in how district courts have applied the law of pleading. (*Why* it has had no effect is a question for another paper. See Hubbard 2011.)

The nested selection model in this paper is a first step toward a new, more empirically

---

[36]Further, the results in this paper do not address the impact, if any, of the *Iqbal* decision.

tractable theory of the litigation process. I will close by noting some potential next steps. First, it may be fruitful to generalize this model to other stages of litigation, including summary judgment and trial. Second, in this paper, I abstract away from the process by which parties may gain information in each stage of litigation. This is an essential feature of litigation—indeed, it is the purpose of discovery—and ultimately any selection model should help explain this process. Watanabe (2009) takes a notable step in this direction. Finally, an unstated premise of Priest and Klein (1984) and all subsequent selection models (including in this paper) is that, in essence, there is *one* court in which the plaintiff can file if the dispute does not settle. The reality, at least in the United States, is that many disputes can be filed as lawsuits in either state court or federal court, or in one of several state courts. This issue of *forum choice* has been carefully studied through the lens of personal jurisdiction, federal jurisdiction, and conflict of laws, but ignored by selection models. I leave to future work the possibility that accounting for forum choice may change the predictions of the nested selection model or open up new possibilities for measuring legal change in a multi-jurisdiction environment.

Model Appendix

6.1   The Selection of Disputes for Filing

Take a given category of dispute, defined in terms of subject matter, amount at stake (which I will call J), etc. For this category, the factual characteristics of any given disputes can be summarized in a single index of case quality, i.e., how strong the plaintiff's case would be, relative to the governing legal standard(s), were the case litigated in court. Call the quality of a given dispute Y. The distribution of all disputes (within the given category) is $F(Y)$. Call the corresponding density function of case quality $f(Y)$.[37]

---

[37]Treating $F(Y)$ as a distribution implies a normalization of the total number of disputes to 1.

There is a set of legal standards governing how the court will decide contested motions and how the court (or jury) would ultimately decide who prevails at trial. For this paper, I concentrate on a single stage of litigation, the motion to dismiss, and deal only abstractly with subsequent stages of litigation. Call the legal standard for granting a motion to dismiss S. If a lawsuit is filed, and the defendant files a motion to dismiss, the court will dismiss the case if the case quality Y is lower than the dismissal standard S and will deny the motion otherwise.

Before a dispute becomes a lawsuit, the (potential) plaintiff and (potential) defendant may settle the dispute. The parties to the dispute will look ahead to litigation and estimate their chances of success, based on the dismissal standard S and all of the other applicable legal standards (call this set Q). The legal standard for the plaintiff to prevail in litigation is then summarized by $V = v(S, Q)$. V is the minimum case quality Y such that the plaintiff will prevail if the case is litigated. Of course, the application of legal rules to the facts of any given case (or vice versa) is uncertain ex ante, and so the parties to the dispute do not know the standard V with precision (nor do they know S precisely).

Each party forms an estimate $\hat{V}_i$ of V with some error:

$$\hat{V}_i = V - \epsilon_i \tag{3}$$

where $i = p, d$ indicates the (potential) plaintiff or (potential) defendant, and the error terms ϵ_p and ϵ_d are independent, normally distributed random variables with zero mean and identical standard deviations σ_ϵ.[38] The parameters of the error terms are known to the parties. Each party, at the time of the settlement decision, has an unbiased (but different) estimate of the true case quality relative to the liability rule. With $\hat{V}_i$, each party forms an

---

[38]Identical standard deviations is a convenience; to the extent that relaxing this assumption affects the analysis, I note it below.

estimate P_i of the probability that the plaintiff will prevail in litigation:

$$P_i = Pr(Y > V | \hat{V}_i) = \Phi \left(\frac{Y - \hat{V}_i}{\sigma_\epsilon} \right) \tag{4}$$

The parties have the opportunity to settle the case rather than litigate. The judgment the plaintiff would receive if successful in court is J, and the cost of litigation (net of the cost of settlement) to each of the parties is C_i.[39] These values are common knowledge to the parties. The plaintiff is willing to settle for any amount at least as great as his expected net judgment in litigation. Hence, his "asking price" A is

$$A = P_p J - C_p \tag{5}$$

The defendant is willing to settle for any amount up to her expected net costs in litigation. Her "bidding price" B is

$$B = P_d J + C_d \tag{6}$$

The model assumes that so long as $B \geq A$, i.e., there exists a settlement amount acceptable to both parties, then the parties will settle. If we define $C = C_p + C_d$, then the condition for litigation is the familiar "Landes-Posner-Gould" condition:

$$P_p - P_d > \frac{C}{J} \tag{7}$$

Cases that satisfy this condition will make up the population of filed lawsuits. Note that in addition to the requirement that the parties no mutually agreeable settlement amount exist, the model assumes that a plaintiff will only file suit if it has positive expected value (PEV). In the context of the selection of disputes for filing, the PEV condition is $P_p J - C_p > 0$,

___

[39]Note that my description of the model does not separately address the possibility of settlement *after* filing. Of course, settlement after filing is in practice a very common outcome. This model allows parties to account for the possibility of settling after filing; expectations about post-filing settlement simply adjust C_p, C_d, and J.

which is always satisfied if the Landes-Posner-Gould condition is met. Hence, we need not discuss the PEV condition further here.

To derive the distribution of case quality among filed lawsuits, first note that conditional on Y and ϵ_p, Equations (4) and (7) imply that the probability that a given dispute becomes a filed lawsuit is given by

$$
\begin{aligned}
Pr(file|Y,\epsilon_p) &= Pr\left(P_p - P_d > \frac{C}{J}\right) \\
&= Pr\left(\Phi\left(\frac{Y-V+\epsilon_p}{\sigma_\epsilon}\right) - \Phi\left(\frac{Y-V+\epsilon_d}{\sigma_\epsilon}\right) > \frac{C}{J}\right)
\end{aligned}
\tag{8}
$$

Solving for the distribution of case quality among filed cases (call this $G(Y)$) yields

$$
G(Y) = \int_{-\infty}^{Y}\int_{-\infty}^{\infty} \Phi\left(\frac{V-y}{\sigma_\epsilon} + \Phi^{-1}\left[\Phi\left(\frac{y-V+x}{\sigma_\epsilon}\right) - \frac{C}{J}\right]\right) d\Phi\left(\frac{x}{\sigma_\epsilon}\right) dF(y)
\tag{9}
$$

Note that $G(Y)$ is *not* a probability distribution, in the sense that it has a density function $g(Y)$ that integrates to 1. Instead, $g(Y)$ would integrate to the share of all disputes that are filed lawsuits. (Recall that $F(Y)$, the distribution of disputes, is a proper distribution and reflects a normalizing of the number of disputes to 1.) For convenience, I will abuse terminology somewhat and refer to $g(Y)$ as the "density function" of quality among filed disputes.

Because the distribution of case quality among all disputes $F(Y)$ is unobserved, the precise shape of the distribution of case quality among filed cases is also unknown. Nonetheless, for a broad range of possible distributions $F(Y)$, the selection effects are sufficiently strong that the general characteristics of $G(Y)$ can be described with some confidence. Even for underlying distributions that are highly skewed and have very little mass near the legal standard V, the distribution of Y is approximately normal with a mean near V.[40]

---

[40]If parties' uncertainty about V is not symmetrical, i.e., if σ_ϵ is not the same for both parties, the distribution of filed cases will be shifted to the left if plaintiff's uncertainty is greater, and shifted to the right if defendant's uncertainty is greater.

If $F(Y)$ is uniform, then $G(Y)$ looks like a normal distribution with mean V. This yields the strong form of the Priest and Klein hypothesis, at least as applied to all filed lawsuits, rather than trials: of the cases that are filed, plaintiffs will win half. (A skewed distribution $F(Y)$ would shift the plaintiff win rate away from zero.) Waldfogel (1995) and Siegelman and Waldfogel (1999) implicitly test the result that quality among filed cases is normally distributed (though not necessarily centered on V), because the versions of the model they test assume a normal distribution of case quality. Their results support the model.

Before turning to the selection of filed lawsuits for litigation on a motion to dismiss (the nested selection process), I will take stock of my model so far. I have applied the logic of Priest and Klein (1984) to the decision to settle a dispute or file a lawsuit. This accounts for the empirical reality that many (I would say the vast majority) of disputes are settled without a lawsuit ever being filed.[41] Leaving unspecified the (unobserved) distribution of case quality among all disputes, I have been able to characterize, at least in general terms, the distribution of the subset of disputes that become filed cases. The prediction that the quality of filed cases will be concentrated around the legal standard V goes beyond Priest and Klein (1984), who assume that the distribution of filed cases may take any shape; and it diverges from Priest (1985), Wittman (1985), and Siegelman and Donahue (1995), who assume that all or nearly all present expected value (PEV) disputes are filed. Indeed, contrary to Priest (1985) and Wittman (1985), whose reasoning implies that the highest-expected-value disputes are the disputes most likely to become lawsuits, my model predicts that they are among the *least* likely disputes to become lawsuits.

---

[41]Evidence on the number of *all* disputes relative to litigated disputes is hard to come by. In some contexts, one way to characterize the set of disputes is to compare the number of all injury claims made to an insurer against the number of such injuries that become the subject of a lawsuit. Conard et al. (1964), in a seminal study of automobile accident injury claims, find that about 5 percent of all injuries become the subject of litigation. Also, the Landes-Posner-Gould condition (equation (7)) predicts that likelihood of a filed lawsuit rises with the amount at stake, i.e., the injury, and this is precisely what the data indicate. Conard et al. (1964) report that 27 percent of injuries over $25,000 (in 1958 dollars) lead to lawsuits, but only 2 percent of smaller injuries do. Black et al. (2008), using data on commercially insured personal injury claims, do not measure lawsuit filing rates, but compare total claims to claims ending in fully litigated trials, and find that claims involving injuries over $25,000 (in 1988 dollars) are about 10 times more likely to end in a lawsuit and trial than smaller claims. Note that only 0.26 percent of all claims in the Black et al. (2008) data ended in a trial.

One thing that this new model cannot do (so far) is budge the negative prediction of the original selection model: a change in a legal standard that changes V will lead to no unambiguous prediction about the rate at which cases are filed, or the share of cases won by plaintiffs. Say that there is an upward shift in V. A new distribution of filed cases will emerge, but because we cannot observe $F(Y)$, the number of filed cases and the share that plaintiffs will win could in principle rise, fall, or stay the same.

6.2   The Selection of Disputes for Dismissal

Among filed cases, the first major decision the parties generally face is whether to litigate a motion to dismiss or not. If the defendant files a motion to dismiss, the parties litigate the issue and the court dismisses the case if case quality Y is below the dismissal standard S. How does the motion to dismiss standard S relate to the ultimate success threshold V? Litigation of a motion to dismiss is the first stage of litigation and one which only the defendant can win. Plaintiffs who prevail at the motion to dismiss stage may go on to win or lose at a later stage of litigation. Thus, since plaintiffs who meet the threshold S may fail to meet the threshold V, it must be the case that S is a lower threshold that V.[42] Indeed, given the many stages of litigation that occur *after* a motion to dismiss—such as discovery, summary judgment, and trial—we would expect that S is much lower standard than V. And this is generally how lawyers perceived the standard for a motion to dismiss—at least until *Twombly*.

The parties are uncertain about the application of the pleading standard to the facts of their case, so they form estimates:

$$\hat{S}_i = S - e_i \tag{10}$$

where the plaintiff's and defendant's errors are e_p and e_d, respectively. Among all disputes, I assume that e_p and e_d are each independent and normal with mean zero and standard

---

[42]Of course, as noted above, V may depend on S.

deviation σ_e. It is important to note that σ_e will be smaller than σ_ϵ. This is because the outcome of a motion to dismiss depends only on the facts alleged in the plaintiff's complaint and the judge's application of the law to those facts; the outcome of a fully litigated lawsuit, however, depends not only on the plaintiff's allegations, but the defendant's allegations (which come after the motion to dismiss), the evidence produced by the parties, the testimony of witnesses, and the judge's (and possibly a jury's) perceptions of all of these, not to mention the judge's application of the law not only to the plaintiff's allegations, but questions of evidence and procedure. Hence, parties' uncertainty about the overall litigation outcome should be much greater than uncertainty about the outcome of a motion to dismiss.[43]

With $\hat{S}_i$, each party forms an estimate P_i^{MTD} of the probability that the plaintiff will prevail on the motion to dismiss:

$$P_i^{MTD} = Pr(Y > S | \hat{S}_i) \tag{11}$$

As the defendant has the option to file, the roles of the parties are reversed, but otherwise the analysis of the selection process proceeds as above. There will be a subset of filed cases in which the parties litigate a motion to dismiss, and the quality Y of these cases will tend to be concentrated around the dismissal standard S. The rest of the cases will "settle."

Here it is important to note what "settlement" entails in this context. In the context of this (or any) selection model, settlement is any agreement between the parties to save the costs of litigation by avoiding litigation. Here "litigation" specifically means a motion to dismiss and the "costs of litigation" are specifically the costs of litigating a motion to dismiss. Thus, settlement is not limited to settling the case before a motion to dismiss is filed (although this in fact happens); it can include the decision to simply forgo the issue (which implies a settlement of zero) and move forward with the litigation.

---

[43]However, because $V = v(S, Q)$, we might expect e_p to (positively) covary with ϵ_p and to allow e_d to (positively) covary with ϵ_d. Whether this is so does not affect the discussion below.

The defendant will settle for any amount less than she expects to pay if she does file a motion to dismiss. Her "bidding price" B^{MTD} is

$$B^{MTD} = P_d^{MTD} J_d^{MTD} + C_d^{MTD} \qquad (12)$$

where C_d^{MTD} is the defendant's cost of litigating the motion to dismiss, and J_d^{MTD} is the defendant's belief about the plaintiff's expected judgment, conditional on the motion to dismiss being denied.[44] Analogously, the plaintiff will settle for any amount more than he expects to receive if the defendant files a motion to dismiss. His "asking price" is

$$A^{MTD} = P_p^{MTD} J_p^{MTD} + C_p^{MTD} \qquad (13)$$

The condition for settling, rather than litigating a motion to dismiss, is $B^{MTD} \geq A^{MTD}$. Thus, just as was the case for the selection of disputes for filing, there are two conditions for filing a motion to dismiss. First, there must not be a mutually agreeable settlement amount that would forestall a motion to dismiss. No such settlement exists if $B^{MTD} < A^{MTD}$. Second, a motion to dismiss must have positive expected value (PEV) for the defendant. Unlike in the selection of disputes for filing, in the case of the selection of disputes for dismissal, the first condition does not imply the second. (When the plaintiff is deciding whether to file suit, his outside option is receiving nothing. But when the defendant is deciding whether to file a motion to dismiss, her outside option is the expected cost of moving on to subsequent stages of litigation.)

Both of these conditions will limit the number of motions to dismiss that get filed. However, unlike in the selection of disputes for filing, where the PEV condition was redundant,

---

[44]In the first (filing) stage, I assume a known amount J that is in dispute. The amount the plaintiff "wins" if a motion to dismiss is denied, however, is the expected value of his judgment (conditional on surviving a motion to dismiss, but with other stages of litigation remaining). This depends both on J and on each party's expectation of plaintiff's likelihood of ultimately prevailing (again, conditional on surviving the motion to dismiss). Because of the selection of disputes into filed cases, the parties' beliefs about the plaintiff's expected gain from surviving a motion to dismiss will diverge. Hence, I call these beliefs J_i^{MTD} for party i.

here the PEV condition is likely to be more important in limiting the number cases with litigated motions to dismiss. Why? The relevant universe of disputes in which a motion to dismiss may be filed is the set of all disputes in which a lawsuit is filed. As discussed above, filed lawsuits are disputes in which each party is relatively optimistic about its chances of prevailing at trial. This implies that the defendant's belief about plaintiff's expected judgment (J_d^{MTD}) will be lower than plaintiff's belief about his expected judgment (J_p^{MTD}). Add to this the fact that the cost of litigating a motion to dismiss may be small relative to the amount at stake, and there may be a large number of cases for which $B^{MTD} < A^{MTD}$; i.e., no settlement range exists.

On the other hand, filed lawsuit also tend to be disputes where true case quality Y is close to the victory standard V. This means that Y is far above the standard for a motion to dismiss S. Consequently, both plaintiff and defendant are likely to recognize that any given case—conditional on being a filed lawsuit—is very unlikely to have case quality below S. In other words, P_d^{MTD} is very close to one, and even if a motion to dismiss is not very expensive, it is unlikely to have positive expected value. Only in the handful of filed lawsuits with true case quality near S will a defendant perceive the likelihood of a motion to dismiss succeeding to be much above zero. In plain English, if the parties know that the court is going to deny the motion, the defendant will save her money and move on.

In short, the distribution $G(Y)$ of case quality among filed cases implied by my model suggests that motions to dismiss will be rarely filed. S is necessarily below V, and the selection of disputes for filing predicts that the case quality of filed cases will be concentrated around V. Thus, a relatively small fraction of cases—only those with case quality near S— will see motions to dismiss. Figure 1 illustrates this. The broken line is the density $g(Y)$ of case quality among filed cases, and the solid line indicates the distribution of case quality in the subset of cases in which a motion to dismiss is litigated (call this distribution is $H(Y)$; again, this is not a probability distribution).

Data Appendix

6.3   AO Data Processing

The AO data are compiled and published as public-use database files by the Federal Judicial Center (FJC) and made available through the Inter-University Consortium for Political and Social Research (ICPSR)[45] entitled Federal Court Cases: Integrated Data Base. See FJC 2005a-d, 2006a-b, 2007, 2009a-b, 2011.

My processing of the AO Data includes the following steps: I begin with the civil terminations files for SY 1970-2009, and the 2008 pending cases file.[46] I drop all observations litigated in non-Article III district courts (these are district courts for U.S. territories; all district courts in U.S. states are included in the sample). The excluded courts correspond to $district > 90$ in the AO Data. I drop all observations in which the jurisdictional basis is "local question" ($jurisdiction = 5$), a code that should only apply in territorial courts and to the District of D.C. before the creation of the D.C. Superior Court in 1970. For the same reason I drop observations coded with a "local question" nature of suit ($natureofsuit = 910$, 920, 930, 940, or 992). The code $natureofsuit = 990$ ("other") is ambiguous as to whether it applies only to local question cases; I retain these observations. I calculate a new variable *duration*, case duration in days, by subtracting *filedate* from *termdate*. I drop all observations with negative duration. (Note that observations with zero duration are not, as a rule, errors. Cases can be filed and terminated in the same day, often as a result of previously negotiated settlement.) I drop observations coded as "Land Condemnation" cases (*progress* $= 0$).

For SY 1972-1978, disposition data is missing, so I drop all observations either filed or

---

[45]See http://www.icpsr.umich.edu/icpsrweb/ICPSR/.

[46]For SY 1970-1991, the AO used a July-to-June reporting period; e.g., SY 1986 covered July 1, 1985 through June 30, 1986. During SY1992, the AO changed from a July-to-June reporting period to an October-to-September period that conformed to the federal government's fiscal year. SY 1992 covers a 15-month period. All years beginning with SY 1993 run October-to-September.

terminated on or before June 30, 1978. I drop all observations (other than pending cases) with missing values for *disposition.*

I drop a number of nature of suit categories that are either inconsequential or that reflect peculiar (both in terms of content and timing) types of litigation. These include "State reapportionment," i.e., congressional redistricting litigation, which spikes every 10 years (*natureofsuit* = 400); "Parole Board Review," a quasi-criminal nature of suit code discontinued in 1985 (*natureofsuit* = 520); "Welfare Pension Plans," of which there is only 1 after 1980 (*natureofsuit* = 750); "Selective Service," which was bunched in the early 1970s (*natureofsuit* = 810); "Black Lung," which is clustered around 1976-1981 (*natureofsuit* = 862); and "Appeal of Fee," which is trivial in number and different in kind from actions originating in the district court (*natureofsuit* = 900).

I drop actions (by the government) to recover student loans, overpayments of benefits, and the like, which as Hadfield (2005, 2004) notes usually involves uncontested actions that are not in fact litigated but end in default judgments (*natureofsuit* = 150, 151, 152, 153). I drop observations (*n* = 55) with missing nature of suit codes or with values for *natureofsuit* that do not correspond to any nature of suit code values in the codebooks for the AO Data. I recode the discontinued value "Antitrust-Electrical Equipment" as "Antitrust" (*natureofsuit* = 410). I combine various discontinued codes labeled "Deportation" into the current "Deportation" code (*natureofsuit* = 461), and various discontinued codes labeled "Bankruptcy" into the current "Bankruptcy" code (*natureofsuit* = 422).

I ensure all codes for district, office number, and docket number are numeric values, replacing alphanumeric values with unused, unique, numerical values when necessary. These codes are used to generate identifiers for each unique federal civil case. I generate dummy variables identifying various categories of suits. See Table A1. As reported in Table A1, a number of categories are excluded from the reported specifications. These include categories outside of (what might be called) the "typical" forms of civil litigation described in the main

text. For example, I exclude habeas corpus petitions, which are essentially criminal matters, and social security cases, which are essentially appeals from administrative proceedings.

Further, reported results exclude cases with a pro se or in forma pauperis party. Shortly after *Twombly*, the Supreme Court issued a per curiam opinion that appeared to limit the application of *Twombly* to pro se cases. See *Erickson v. Pardus*, 551 U.S. 89 (June 4, 2007) (per curiam) ("A document filed pro se is to be liberally construed ... and a pro se complaint, however inartfully pleaded, must be held to less stringent standards than formal pleadings drafted by lawyers."). Reported results also exclude categories of cases involving unusual spikes in filings, such as cases in the Eastern District of Louisiana (due to Hurricane Katrina) and nature of suit categories involving "mass filings" (i.e., more than 1000 cases filed in a single day in a single district). I exclude cases in multi-district litigation (MDL) proceedings; these cases are often terminated (including dismissed) in batches of hundreds or thousands. I exclude fraud (including securities fraud) cases, as these cases involved explicitly heightened pleading standards even before *Twombly*. See Rule 9(b); 15 USC §78u-4(b). Finally, I exclude cases involving the review of arbitral awards($disposition = 15$, 16 or $progress = 13$), appeals to a district court ($diposition = 19$, 20), and bankruptcy stays (variously coded, depending on year), all of which are procedurally unique.

For computing dismissal rates, I exclude all observations in which the disposition is a transfer to a new district (including multi-district litigation (MDL) transfers). I exclude cases for which no opportunity arose for a motion to dismiss; given my definition of *dismissal*, these are cases with $duration \leq 45$. For the same reason, I drop cases that were disposed of on grounds that are procedurally antecedent to a Rule 12(b)(6) motion, such as dismissals for want of jurisdiction, dismissals for want of prosecution, default judgments, remands of removed cases to state court, and remands to U.S. agencies.

I also account for double counting of cases in the raw data. If a single case is closed and reopened (this can happen multiple times) or transferred to another district (including for

an MDL proceeding), that case will appear as multiple observations in the AO data, even though there has been only a single opportunity for the case to be terminated because of a Rule 12(b)(6) motion to dismiss (setting aside exceptions which can be ignored for my purposes). For example, many cases coded as being "dismissed," are in fact being dismissed "without prejudice," i.e., the plaintiff can refile the lawsuit at any time. (This is often done when the parties are negotiating a settlement. Upon reaching an settlement, the parties reopen the case in order to obtain a final judgment dismissing the case "with prejudice.")

To address double counting, I de-duplicate case records in the data. For each unique case, as identified by each unique combination of district, office, and docket number, I collapse all observations for that case into a single observation. (About 6 percent of cases had duplicate observations.) For these cases, I retain the first-filed observation (as determined by filing date), but create new fields to capture the termination date, procedural progress, prevailing party, and disposition codes from the last-terminated observation (as determined by termination date). The variable *duration* is calculated as the difference between the first filing date and the last termination date.

Notably, in more than 99 percent of cases with multiple observations, time-invariant variables, such as jurisdictional basis and nature of suit code, had no discrepancies across observations for the same case. This suggests a high degree of reliability in the data. Also, although some studies have found high error rates in the coding of some fields in the AO Data (see Eisenberg and Schlanger 2003; Hadfield 2004), the accuracy of the fields describing case characteristics essential to this paper (such as the jurisdiction, dates of filing and disposition, etc.) has never been questioned, and Hadfield (2005) finds the key outcome field, emphdisposition, to be fairly reliable.[47]

As discussed in the main text, the *disposition* field does not distinguish between Rule

---

[47]Based on audits using docket records of a sample of cases, Hadfield (2005) finds that the "Judgment on motion before trial" code corresponds to final judgments on motions before trial (presumably either Rule 12(b)(6) or Rule 56 motions) in 85-90 percent of cases and the "Judgement: Settled" code corresponds to settlements in approximately 95 percent of cases.

12(b)(6) motions to dismiss and Rule 56 motions for summary judgment. To identify Rule 12(b)(6) dismissals, I first exclude cases coded (in the *judgmentfor* field) as terminated in favor of plaintiff. I then consider the duration of the case at termination. When a plaintiff files a complaint with the court, the plaintiff must also serve the complaint on the defendant. A defendant has 20 days from the date of service or, if the defendant waives the requirement of service, 60 days from the mailing of the request for waiver of service to file a Rule 12(b)(6) motion. See Rule 12(a)(1). Upon filing of a motion to dismiss, the court typically sets a briefing schedule, with perhaps 30 days to file a response in opposition to the motion, and 30 days beyond that for filing a reply in support of the motion. After briefing, the court may schedule a hearing (again, perhaps within 30 days) at which the parties can argue the motion. Thereafter, unless the case has settled, the court will grant or deny the motion. Given this timeline, I would expect that a contested Rule 12(b)(6) motion is unlikely to be granted immediately after the case is filed. Conversely, unless a large number of deadline extensions are granted, or the court is slow to decide the motion, a motion to dismiss is relatively unlikely to terminate a case 9 or 12 months after filing. As noted in the main text, the duration window of 45-224 days accommodates this rough timetable and is consistent with published studies of case duration at time of ruling on a motion to dismiss.

On the other hand, a Rule 56 motion requires that parties have completed sufficient discovery to present evidence demonstrating the presence or absence of factual disputes in the case. Thus, a Rule 56 motion necessarily comes after a Rule 12(b)(6) motion (if one is filed), and usually many months (or years) into a case, depending on the breadth of discovery. For this reason, I assume that Rule 56 motions are decided 225 or more days into a case.

This recoding of the AO Data yields summary statistics (not reported, available from author) consistent with earlier studies of the rates at which cases were terminated due to Rule 12(b)(6) motions to dismiss or Rule 56 motions for summary judgment. See Williging (1989); Cecil et al. (2007). In any event, as noted in the main text, my reported results are robust to variations in the time frame used to define a termination based on a motion to

dismiss in the AO Data.

6.4   Opinions Data Processing

The Opinions Data is derived from a data set collected by Lee Epstein, William M. Landes, and Richard A. Posner. Each observation in this data set represented a published judicial opinion ruling on a motion to dismiss. For each observation, the data set provided the Westlaw citation of the opinion, the month and year of the opinion, the district in which the case was decided, an indicator for whether the case was a civil rights case (based on subjective coding criteria), an indicator for whether the opinion granted the motion to dismiss in its entirety, as well as a number of other fields not used herein.

I took the following steps to process the data: I removed duplicate observations from the data set and then divided the remaining observations among myself and my research assistants by random assignment. Each opinion was then retrieved from Westlaw and the exact date of the opinion, as well the district, office, and docket number information was recorded. The date and district information was then cross-checked against the corresponding fields in the data set and discrepancies were resolved. Additional duplicate observations identified during this process were dropped as well. In less than 1 percent of cases, observations referring to the same opinion (i.e., what I have called "duplicate observations") contained different codes for whether the motion to dismiss was granted in full or for whether the case was a civil rights case. All such observations were dropped from the data set.

This yielded a data set of 12,717 unique opinions. Results using this data set are reported in the main text. This configuration of the data set makes it most comparable to previous studies on the effect of *Twombly* on the rate at which motions to dismiss are granted.

In order to check the reliability of the coding in this dataset and to generate results more comparable to results based on the AO Data, I then reorganized the data set by unique case identifier (given by district, office, and docket number) rather than Westlaw citation.

Some cases had more than one opinion on Westlaw; for these cases, information about the multiple opinions was collapsed into a single observation. Conversely, some opinions dealt with more than one case (this happened in the case of consolidated cases, including multi-district litigation). For these opinions, multiple observations were created, each with the same opinion-specific information, and one observation assigned to each unique case.

Once in this configuration, this data set could be merged with the AO Data by matching on unique case identifier. Of the 13,031 observations generated by this process, more than 98 percent (all but 242) were successfully matched with observation in the AO Data. By matching the Opinions Data to the AO Data, I can check for consistency across those codes that are common to both datasets; despite the fact that the Opinions Data was coded by hand, approximately 98 percent of all observations in the matched Opinions Data have identical values for those fields that also appear in the AO Data.[48] This indicates a high level of reliability, particularly for a hand-collected and hand-coded data set.

I then drop all observations excluded from all specifications using the AO Data (see Table A1) and all observations ($n = 83$) in which there were opinions published both before and after *Twombly*. Matching the Opinions Data with the AO Data also allows me to compare the date on which a motion to dismiss is granted in the Opinions Data to the date on which a case is terminated in the AO Data. From this, I calculate the fraction of cases terminated within 90 days of the granting of a motion to dismiss.

Finally, I re-run the specifications for the Opinions Data using this reconfigured version of the data. All of the results are qualitatively identical to those reported in Table 4 in the main text, and in fact all point estimates for the effect of *Twombly* are closer to zero.

---

[48]Nonetheless, there still remains the possibility of measurement error leading to some attenuation bias in the results. As the outcome variable is a dummy (generally for whether the case was terminated by a Rule 12(b)(6) dismissal), the estimated β will depend on both the covariance of the true disposition with X and the covariance of the measurement errors with X. And because the disposition variable can take only two values, one or zero, measurement errors will systematically vary with X in the opposite direction that the true disposition varies with X.

References

Black, Bernard, David A. Hyman, Charles Silver, and William M. Sage. 2008. Defense Costs and Insurer Reserves in Medical Malpractice and Other Personal Injury Cases: Evidence from Texas, 1988–2004. *Amer. L. & Econ. Rev.* 10(2):185-245.

Bone, Robert G. 2009. *Twombly*, Pleading Rules, and the Regulation of Court Access. *Iowa L. Rev.* 94:873-936.

Burbank, Stephen B. 2004. Keeping Our Ambition Under Control: The Limits of Data and Inference in Searching for the Causes and Consequences of Vanishing Trials in Federal Court. *J. Empirical Legal Stud.* 1:571-590.

Campbell, Ray Worthy. 2010. Getting a Clue: Two Stage Complaint Pleading as a Solution to the *Conley-Iqbal* Dilemma. *Pa. St. L. Rev.* 114:1191-1245.

Cecil, Joe S., George W. Cort, Margaret S. Williams, and Jared J. Bataillon. 2011. Motions to Dismiss for Failure to State a Claim After *Iqbal*. Washington, D.C.: Federal Judicial Center.

Cecil, Joe S., Rebecca N. Eyre, Dean Miletich, and David Rindskopf. 2007. A Quarter-Century of Summary Judgment Practice in Six Federal District Courts. *J. Empirical Legal Stud.* 4:861-907.

Conard, Alfred F., James N. Morgan, Robert W. Pratt, Jr., Charles E. Voltz, and Robert L. Bombaugh. 1964. *Automobile Accident Costs and Payments: Studies in the Economics of Injury Reparation.* Ann Arbor, Michigan: University of Michigan Press.

Connolly, Paul R. J., and Patricia A. Lombard. 1980. Judicial Controls and the Civil Litigation Process: Motions. Washington, D.C.: Federal Judicial Center.

Cooter, Robert. 1987. Why Litigants Disagree: A Comment on George Priest's "Measuring Legal Change." *J Law, Econ, & Org* 3(2):227-234.

Dodson, Scott. 2010. Comparative Convergences in Pleading Standards. *U. Pa. L. Rev.* 158:441-472.

Eisenberg, Theodore. 1991. The Relationship between Plaintiff Success Rates before Trial and at Trial. *J Royal Stat Soc A* 154(1):111-116.

Eisenberg, Theodore, and Margo Schlanger. 2003. The Reliability of the Administrative Office of the U.S. Courts Database: An Initial Empirical Analysis. *Notre Dame L. Rev.* 78:1455-1496.

Epstein, Richard A. 2009. *Twombly*, After Two Years: The Procedural Revolution in Antitrust That Wasn't. *GCP: The Online Magazine for Global Competition Policy* July 2009(2).

Fairman, Christopher M. 2003. The Myth of Notice Pleading. *Ariz. L. Rev.* 45:987-1065.

Federal Judicial Center. 2005a. Federal Court Cases: Integrated Database, 1970-2000 (Computer file ICPSR08429-v7.) Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

Federal Judicial Center. 2005b. Federal Court Cases: Integrated Database, 2001 (Computer file ICPSR03415-v3.) Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

Federal Judicial Center. 2005c. Federal Court Cases: Integrated Database, 2002 (Computer file ICPSR04059-v3.) Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

Federal Judicial Center. 2005d. Federal Court Cases: Integrated Database, 2003 (Computer file ICPSR04026-v5.) Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

Federal Judicial Center. 2006a. Federal Court Cases: Integrated Database, 2004 (Computer file ICPSR04348-v2.) Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

Federal Judicial Center. 2006b. Federal Court Cases: Integrated Database, 2005 (Computer file ICPSR04382-v1.) Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

Federal Judicial Center. 2007. Federal Court Cases: Integrated Database, 2006 (Computer file ICPSR04685-v2.) Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

Federal Judicial Center. 2009a. Federal Court Cases: Integrated Database, 2007 (Computer file ICPSR22300-v2.) Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

Federal Judicial Center. 2009b. Federal Court Cases: Integrated Database, 2008 (Computer file ICPSR25002-v2.) Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

Hadfield, Gillian K. 2005. Exploring Economic and Democratic Theories of Civil Litigation: Differences between Individual and Organizational Litigants in the Disposition of Federal Civil Cases. *Stan. L. Rev.* 57:1275-1327.

Hamburg, Rebecca M., and Matthew C. Koski. 2010. *Summary of Results of Federal Judicial Center Survey of NELA Members, Fall 2009.*

Hannon, Kendall W. 2008. Much Ado About *Twombly*? A Study of the Impact of *Bell Atlantic Corp. v. Twombly* on 12(b)(6) Motions. *Notre Dame L. Rev.* 83:1811-1846.

Hatamyar, Patricia W. 2010. The Tao of Pleading: Do *Twombly* and *Iqbal* Matter Empirically? *Am. U. L. Rev.* 59:553-633.

Hoffman, Lonny S. 2008. Burn Up the Chaff with Unquenchable Fire: What Two Doctrinal Intersections Can Teach Us About Judicial Power over Pleadings. *B.U. L. Rev.* 88:1217-1270.

Hoffman, David A., Alan J. Izenman and Jeffrey R. Lidicker. 2007. Docketology, District Courts, and Doctrine. *Wash. U. L. Rev.* 85:681-751.

Hubbard, William H.J. 2011. The (Other) Paradoxes of *Twombly* and Civil Rights Litigation. Unpublished working paper.

Hylton, Keith N. and Haizhen Lin. 2010. Trial Selection Theory: A Unified Model. B.U. School of Law Working Paper No. 10-12.

Institute for the Advancement of the American Legal System. 2009. *Civil Case Processing in the Federal District Courts.*

Kilaru, Rakesh N. 2010. Comment. The New Rule 12(b)(6): *Twombly, Iqbal,* and the Paradox of Pleading. *Stan. L. Rev.* 62:905-930.

Kourlis, Rebecca Love and Jordan M. Singer. 2009. Managing Towards the Goals of Rule 1. *Fed. Cts. L. Rev.* 2009:1-18.

Lizotte, Brian N. 2007. Publish or Perish: The Electronic Availability of Summary Judgments by Eight District Courts. *Wisc. L. Rev.* 2007:107-149.

Malani, Anup, and Julian Reif. 2010. Accounting for Anticipation Effects: An Application to Medical Malpractice Reform. NBER Working Paper 16593.

Marcus, Richard L. 1986. The Revival of Fact Pleading Under the Federal Rules of Civil Procedure. *Colum. L. Rev.* 86(3):433-494.

Miller, Arthur R. 2010. From *Conley* to *Twombly* to *Iqbal*: A Double Play on the Federal Rules of Civil Procedure. *Duke L.J.* 60(1):1-130.

Priest, George L. 1985. Reexamining the Selection Hypothesis: Learning from Wittman's Mistakes. *JLS* 19:215-243.

Priest, George L. 1987. Measuring Legal Change. *J Law, Econ, & Org* 3(2):193-225.

Priest, George L., and Benjamin Klein. 1984. The Selection of Disputes for Litigation. *J. Legal Stud.* 13:1-55.

Redish, Martin H. and Lee Epstein. 2008. *Bell Atlantic v. Twombly* and the Future of Pleading in the Federal Courts: A Normative and Empirical Analysis. Northwestern Univ. School of Law Public Law and Legal Theory Working Paper No. 10-16.

Reinert, Alexander A. 2011. The Costs of Heightened Pleading. *Indiana L. Rev.* 86:119 (forthcoming).

Schnapper, Eric. 2009. Statement to the House Subcommittee on Courts and Competition Policy on December 16, 2009, Hearings on the Notice Pleading Restoration Act of 2009, S. 1504, 111th Cong. sec. 2 (available at 2009 WL 4829082).

Schneider, Elizabeth M. 2010. The Changing Shape of Federal Civil Pretrial Practice: The Disparate Impact on Civil Rights and Employment Discrimination Cases. *U. Pa. L. Rev.* 158:517-570.

Seiner, Joseph A. 2010. Pleading Disability. *B.C. L. Rev.* 51:95-149.

Seiner, Joseph A. 2009. The Trouble with *Twombly*: A Proposed Pleading Standard for Employment Discrimination Cases. *U. Ill. L. Rev.* 2009:1011-1060.

Siegelman, Peter, and John J. Donahue III. 1995. The Selection of Employment Discrimination Disputes for Litigation: Using Business Cycle Effects to Test the Priest-Klein Hypothesis.

Siegelman, Peter, and Joel Waldfogel. 1999. Toward a Taxonomy of Disputes: New Evidence Through the Prism of the Priest/Klein Model. *JLS* 28:101-130.

Smith, Douglas G. 2009. The *Twombly* Revolution? *Pepp. L. Rev.* 36:1063-1100.

Spencer, A. Benjamin. 2008. Plausibility Pleading. *B.C. L. Rev.* 49:431-494.

Spencer, A. Benjamin. 2010. *Iqbal* and the Slide Toward Restrictive Procedure. *Lewis & Clark L. Rev.* 14:185-201.

Stancil, Paul. 2009. Balancing the Pleading Equation. *Baylor L. Rev.* 61:90-173.

Steinman, Adam N. 2010. The Pleading Problem. *Stan. L. Rev.* 62:1293-1360.

Waldfogel, Joel. 1995. The Selection Hypothesis and the Relationship between Trial and Plaintiff Victory. *JPE* 103(2):229-260.

Waldfogel, Joel. 1998a. Reconciling Asymmetric Information and Divergent Expectations Theories of Litigation. *J L & Econ* 41:451-476.

Waldfogel, Joel. 1998b. Selection of cases for trial. Vol. 3, pp. 419-424 in *The New Palgrave Dictionary of Economics and the Law*, edited by Peter Newman. New York: Stockton Press.

Ward, Ettie. 2008. The After-Shocks of *Twombly*: Will We "Notice" Pleading Changes? *St. John's L. Rev.* 82:893-918.

Wasserman, Howard M. 2010. *Iqbal*, Procedural Mismatches, and Civil Rights Litigation. *Lewis & Clark L. Rev.* 14(1):157-183.

Watanabe, Yasutora. 2009. Learning and Bargaining in Dispute Resolution: Theory and Evidence from Medical Malpractice Litigation. Unpublished working paper.

Willging, Thomas E. 1989. Use of Rule 12(b)(6) in Two Federal District Courts. Washington, D.C.: Federal Judicial Center.

Willging, Thomas E. and Emery G. Lee, III. 2010. *In Their Words: Attorney Views About Costs and Procedures in Federal Civil Litigation.* Washington, DC: Federal Judicial Center.

Wittman, Donald. 1985. Is the Selection of Cases for Trial Biased? *JLS* 14:185-214.

Figures and Tables

Figure 1: Motions to Dismiss and Dismissals
Nested Selection Model
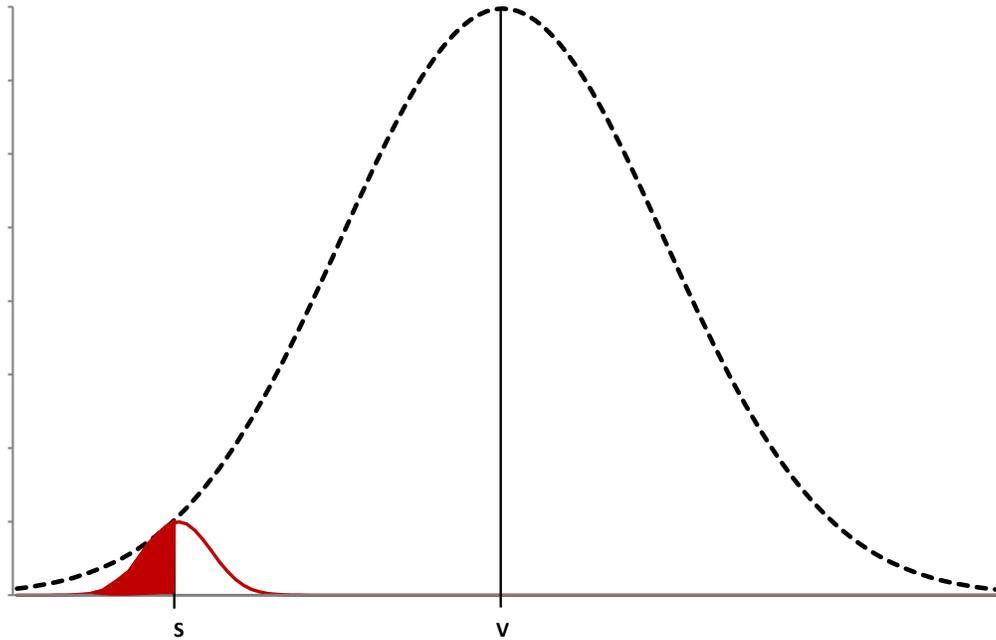
Note: Dashed curve represents the distribution of case quality Y among filed cases. Solid curve represents the distribution of filed cases in which a motion to dismiss is litigated. Shaded region indicates the set of cases dismissed. V is the overall threshold for plaintiff success. S is the pleading standard. (Height of solid curve is exaggerated for clarity. Even near S, only a fraction of filed cases have litigated motions to dismiss.)

Figure 2: Motions to Dismiss and Dismissals
under Higher Pleading Standard S', Short Run
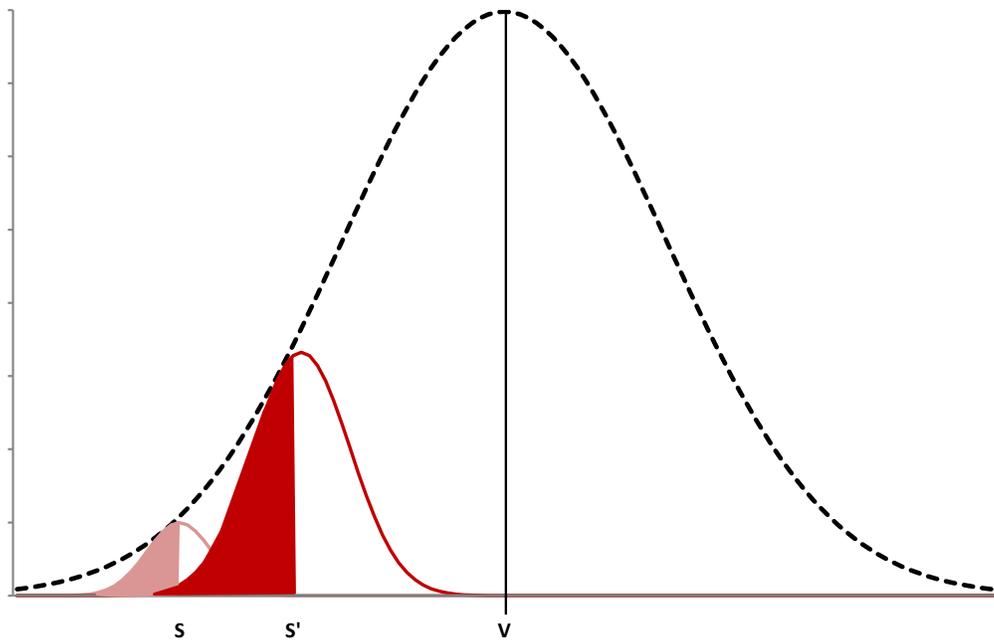
Note: Dashed line represents the distribution of case quality Y among filed cases. Solid line represents the distribution of filed cases in which a motion to dismiss is litigated. Shaded region indicates the set of cases dismissed. V is the overall threshold for plaintiff success. S' indicates a new, higher pleading standard. (Height of solid curve is exaggerated for clarity. Even near S', only a fraction of filed cases have litigated motions to dismiss.)

Table 1: Summary Statistics for Opinions Data

		(1) Pre- *Twombly*	(2) Post- *Twombly*
Civil Rights	Opinions	2,230	2,379
	Dismissal Rate	0.569	0.583
Non-Civil Rights	Opinions	3,851	4,257
	Dismissal Rate	0.466	0.476
Total	Opinions	6,081	6,636
	Dismissal Rate	0.504	0.515

Table 2: Summary of Predictions from Competing Hypotheses

	Conventional	Contrarian; Selection/No Change	Selection/Change
1. MTD Grant Rate	+	0	0
2. Dismissals as Fraction of Filed Cases	+	0	+

Note: "+" and "-" indicate significantly positive or negative changes in response to the treatment. "0" indicates little or no change.

Table 3: Summary Statistics for Administrative Data

		Year Beginning May 22		
		2005	**2006**	**2007**
All Cases	Filings	85,606	83,770	82,682
	Dismissal Rate	0.014	0.015	0.013
Civil Rights	Filings	22,762	21,485	21,193
	Dismissal Rate	0.016	0.016	0.013
Shares	Tort	0.068	0.066	0.066
	Contract	0.260	0.266	0.269
	Civil Rights	0.266	0.256	0.256

Table 4: Opinions Data Regression Results

Model	(1) Linear	(2) Linear	(3) Linear	(4) Logit	(5) Linear	(6) Logit
Twombly	-0.0089 (0.0131)	-0.0112 (0.0132)	-0.0109 (0.0130)	-0.0112 (0.0133)	-0.0170 (0.0207)	-0.0173 (0.0210)
Can Reject at $p = 0.05$ Effect of	0.0127	0.0105	0.0105	0.0107	0.0171	0.0172
Civil Rights			0.1015** (0.0199)	0.1037** (0.0201)	NA	NA
Circuit Dummies		Y	Y	Y	Y	Y
Scope	All	All	All	All	Civil Rights	Civil Rights
Constant	0.5082** (0.0186)	0.6985** (0.0059)	0.6508** (0.0099)		0.7061** (0.0099)	
N	10,288	10,288	10,288	10,288	3,652	3,652

Notes: All errors are clustered at the district level. One and two asterisks denote statistical significance at the 5 and 1 percent level, respectively. Logit coefficients are reported as marginal effects, evaluated at sample mean. Reported values for the constant in regressions with circuit fixed effects reflect the mean in the omitted circuit (DC). The *Twombly* variable is an indicator for cases in which ruling on a motion to dismiss (if any) occurred after the date of *Twombly* (May 21, 2007). "Civil Rights" is an indicator for a civil rights cases. "Scope" denotes whether all cases, or only civil rights cases, were included in a given specification. "NOS" stands for "Nature of Suit."

Table 5: AO Data Regression Results, All Cases

Model	(1) Linear	(2) Linear	(3) Linear	(4) Logistic
Twombly	-0.0023 (0.0022)	-0.0024 (0.0022)	-0.0024 (0.0022)	-0.0021 (0.0020)
Can Reject at p = 0.05 Effect of	0.0013	0.0013	0.0013	0.0011
Constant	0.0187** (0.0020)	0.0187** (0.0014)	0.0208** (0.0029)	
NOS Dummies			Y	Y
Circuit Dummies		Y	Y	Y
N	17,223	17,223	17,223	17,223

Note: See notes following Table 4.

Table 6: AO Data Regression Results, Civil Rights Cases

Model	(1) Linear	(2) Linear	(3) Logit
Twombly	-0.0027 (0.0041)	-0.0027 (0.0040)	-0.0021 (0.0030)
Can Reject at p = 0.05 Effect of	0.0040	0.0039	0.0029
Constant	0.0192** (0.0034)	0.0192** (0.0029)	
Circuit Dummies		Y	Y
N	4,853	4,853	4,853

Note: See notes following Table 4.

Table 7: AO Data: Trends in Dismissals in Longer-Duration Cases

	Filed During Period Ending	
	May 21, 2006	May 21, 2007
Share Dismissed within 365 Days	0.0349 (8,919)	0.0354 (8,304)
Mean Duration, Dismissed Cases	210.9 (311)	219.1 (294)
Share Dismissed within 590 Days	0.0655 (8,919)	0.0625 (8,304)
Mean Duration, Dismissed Cases	330.8 (584)	329.3 (519)

Note: Frequencies appear in parentheses.

Table A1: Nature of Suit Recodes

Recode Category	AO Nature of Suit Codes
Contract	100-196, excluding 191
Real Property	220-230, 290
Torts (excl. Fraud)	310-362, excluding 330, 340, and 360
Fraud (incl. Securities)	371, 850
Antitrust	410
Civil Rights	440-446
Other	191, 210, 330, 720-791, 820-840, 875-890, 895, 950, 990, excluding 750
Mass Filing Categories	340, 360, 365, 368, 370, 380, 710
Other Categories Excluded in Reported Specifications	240-245, 385, 422, 423, 430, 450, 460-490, 610-620, 630-660, 870, 871, 891-894, 625, 690, 970, 860-865, excl. 862, 510-555, excl. 520
Excluded in All Specifications	150-153, 400, 520, 750, 810, 862, 900

Readers with comments should address them to:

Professor William H. J. Hubbard
University of Chicago Law School
1111 East 60th Street
Chicago, IL 60637
        whubbard@uchicago.edu

Chicago Working Papers in Law and Economics
(Second Series)

For a listing of papers 1–550 please go to Working Papers at http://www.law.uchicago.edu/Lawecon/index.html