Selection in Motion:
A Formal Model of Rule 12(b)(6) and the *Twombly-Iqbal* Shift in Pleading Policy

Jonah B. Gelbach[*]

Yale Law School Class of 2013

and

Senior Research Fellow
Program in Applied Economics in Policy
Yale University Department of Economics

August 29, 2012

Abstract

I build a sequential model of pre-discovery litigation and analyze it using standard game-theoretic tools. The model is rich enough to allow me to capture key aspects of pre-discovery litigation and tractable enough to allow me to yield several important results concerning the effects of the Supreme Court's *Twombly* and *Iqbal* cases, which many observers believe raised the pleading standard in federal civil litigation. I show that the Rule 12(b)(6) motion filing and grant rates each might take on any value between zero and 100 percent. I also show that the parties' behavioral responses to more demanding pleading—what I call party selection—might cause any combination of increases, no change, or decreases in the Rule 12(b)(6) motion filing rate, the grant rate, and the number of cases filed. This analytical result casts suggests that very little is at stake in the heated debate concerning recent empirical studies of *Twombly* and *Iqbal*'s effects on the Rule 12(b)(6) grant rate. I then use the model to verify the validity of the empirical approach in an earlier paper, and I reprise that paper's empirical finding that switching to the *Twombly*/*Iqbal* pleading regime had negative effects on plaintiffs in a substantial share of cases that had Rule 12(b)(6) motions filed (or filed and granted) in the post-*Iqbal* period.

I then raise three normative issues related to *Twombly* and *Iqbal*. The first issue concerns the impact of switching to more demanding pleading on parties' expected payoffs from litigation. I show analytically that, on the whole, switching to a more demanding pleading standard will redistribute from defendants to plaintiffs. Second, I tee up the question of how one might tell whether *Twombly* and *Iqbal* have affected only low-merit suits or have had effects that the cases' authors did not intend. Third, I explore the multi-faceted ways in which pleading standards might affect primary behavior, illustrating that the choice of pleading standards involves tradeoffs between reducing unlawful behavior and chilling lawful behavior.

I. Introduction

For the fifty years following *Conley v. Gibson*, the Supreme Court's most widely known on language on pleading was 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."[1] But in 2007's *Bell Atlantic v. Twombly*, the Court declared that the *Conley* no-set-of-facts test had "earned its retirement," at least in *Twombly*'s substantive context of parallel conduct allegedly violating the antitrust laws.[2] In its place, the *Twombly* Court introduced the requirement that, to meet Rule 8(a)'s standard for "showing that the pleader is entitled to relief,"[3] a complaint must recite "enough facts to state a claim to relief that is plausible on its face."[4] Two years later, the Court in *Ashcroft v. Iqbal* eliminated any doubt as to *Twombly*'s trans-substantivity, declaring the plausibility standard applied to "all civil actions."[5]

Many commentators have criticized *Twombly* and *Iqbal* on the ground that they will reduce access to justice for plaintiffs whose claims are meritorious but difficult to document without discovery.[6] Because "[p]leading is the gateway to the federal courts,"[7] changing pleading rules will change more than whether a given case, once stopped at the gateway, is allowed passage.[8] So, *Twombly* and *Iqbal* could affect many cases that settle without filing.[9] And they could affect the frequency with which filed cases are attacked via Rule 12(b)(6) motions to dismiss for failure to state a claim;[10] indeed, empirical evidence suggests that the rate at which

---

[1] *Conley v. Gibson*, 355 U.S. 41, 45-46 (1957).

[2] *Bell Atl. Corp. v. Twombly*, 550 U.S. 544, 563 (2007).

[3] Fed. R. Civ. P. Rule 8(a).

[4] *Id.*, at 570.

[5] *Ashcroft v. Iqbal*, 556 U.S. 662, 684 (2009) (quotation marks removed).

[6] See, Joshua Civin & Debo P. Adegbile, *Restoring Access to Justice: The Impact of* Iqbal *and* Twombly *on Federal Civil Rights Litigation*, AM. CONST. SOC'Y FOR L. & POL'Y (2010), http://www.acslaw.org/sites/default/files/Civin_Adegbile_Iqbal_Twombly.pdf Michael C. Dorf, Iqbal *and Bad Apples*, 14 LEWIS & CLARK L. REV. 217, 218 (2010).

[7] Christopher M. Fairman, *The Myth of Notice Pleading*, 45 Ariz. L. Rev. 987, 988 (2003).

[8] Michael Dorf has gone so far as to write that "[u]nless overturned by Congress or the Rules Advisory Committee process, the *Twombly/Iqbal* pleading rule will play a potentially decisive role in every federal civil case." Dorf, *supra* note 6, at 218.

[9] This point is a major theme of the game-theoretic analysis below, but it is certainly not novel. *See, e.g.,* Richard A. Epstein, *Of Pleading and Discovery: Reflections on* Twombly *and* Iqbal *with Special Reference to Antitrust*, 2011 Illinois Law Review 187, 196 ("[A] bad set of legal rules also leads to bad settlements. As a general matter, these settlements reflect the probable outcomes of cases that go to final judgment. Any errors in the overall procedural rules, therefore, are likely to be embedded in the settlements.").

[10] Fed. R. Civ. P. Rule 12(b)(6). Concerning induced changes in the frequency of Rule 12(b)(6) motion filing, *see, e.g.*, Kevin M. Clermont & Stephen C. Yeazell, *Inventing Tests, Destabilizing Systems*, 95 IOWA L. REV. 821, 840 n.70 (2010), who go so far as to state that a defense attorney "commits legal malpractice if he or she fails to move to dismiss with liberal citations to *Twombly* and *Iqbal*."

1

Rule 12(b)(6) motions are filed among filed cases increased substantially following *Twombly* and *Iqbal*.[11]

For these reasons, and others,[12] debate has raged concerning *Twombly* and *Iqbal*. One particularly controverted issue concerns the extent to which judicial behavior has actually changed. To wit: Have trial court judges changed the way they adjudicate Rule 12(b)(6) motions to dismiss for failure to state a claim?[13]

In this paper, I introduce and analyze a game-theoretic model of pre-discovery litigation. The model is rich enough to take real-life federal procedure seriously, yet tractable enough to provide an *analytical* answer to the hotly debated question posed in the previous paragraph. I analyze the effects of what I call "switching to more demanding pleading."[14] I define this concept precisely below, but for the moment, take "switching to more demanding pleading" to mean a change in pleading policy that (i) causes some parties to believe some Rule 12(b)(6) motions will be more likely to be granted than before the policy change, and (ii) doesn't cause any parties to believe any Rule 12(b)(6) motions will be *less* likely to be granted. My analysis generates a result that might surprise those who have taken up arms in the debate over what the existing empirical evidence tells us about changes in pleading policy. Quite simply,

---

[11] *See* Jonah B. Gelbach, *Locking the Doors to Discovery? Assessing the Effects of* Twombly *and* Iqbal *on Access to Discovery*, 121 Yale L.J. 2270, at 2290-2291 (2012) (discussing empirical evidence reported in Joe S. Cecil, George W. Cort, Margaret S. Williams & Jared J. Bataillon, FED. JUDICIAL CTR., *MOTIONS TO DISMISS FOR FAILURE TO STATE A CLAIM AFTER IQBAL: REPORT TO THE JUDICIAL CONFERENCE ADVISORY COMMITTEE ON CIVIL RULES* (2011) (hereinafter "the original FJC report"), available at http://www.uscourts.gov/uscourts/RulesAndPolicies/rules/Publications/motioniqbal.pdf).

[12] For example, some fear that *Twombly* and *Iqbal* will shield unlawful primary behavior from the law's reach; *see, e.g.,* Joshua Civin & Debo P. Adegbile, *supra* note 6, at 2 (expressing fear that *Twombly* and *Iqbal* might "create an undesirable safe harbor that effectively placessome defendants beyond the reach of civil rights laws"). Others have criticized the process by which the Supreme Court set about changing pleading doctrine, since the Court did not follow the procedure for reforming rules of procedure set out by the Rules Enabling Act, codified at 28 U.S.C. §§ 2071-2077. *See, e.g.,* Arthur R. Miller, *From* Conley *to* Twombly *To* Iqbal*: A Double Play On The Federal Rules of Civil Procedure*, 60 Duke Law Journal 1, 85; and Stephen B. Burbank's rebuttal in Mark Herrmann, James M. Beck, & Stephen B. Burbank, Debate, *Plausible Denial: Should Congress Overrule Twombly and Iqbal*?, 158 U. PA. L. REV. PENNUMBRA 141, 148 (2009), http://www.pennumbra.com/debates/pdfs/PlausibleDenial.pdf ("The Supreme Court did not 'clarify the standards for courts to assess complaints upon motions to dismiss' in its recent pleading decisions. It changed them. It did so, moreover, through a process that was illegitimate and inadequate given the statutory requirements of the Rules Enabling Act, 28 U.S.C. § 2072 (2006).") (Burbank). On the other hand, courts *interpret* the Federal Rules of Civil Procedure all the time. This debate thus hinges on what constitutes a change in the rules as opposed to a change in their interpretation.

[13] To be more precise, the question is probably best posed as whether a given judge would decide a given Rule 12(b)(6) motion in a given case differently after *Twombly/Iqbal*. For empirical evidence concerning changes in how both district and appellate judges handled pleading standards following *Twombly*, *see* Part IV of Martin H. Redish and Lee Epstein, Bell Atlantic v. Twombly *and The Future of Pleading in the Federal Courts: A Normative and Empirical Analysis* (Northwestern University School of Law Public Law Research Paper No. 10–16, revised draft of November 18, 2008), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1581481.

[14] I use the term "more demanding" rather than "heightened" or "strict" because those terms sometimes are used to refer to specific pleading policies, whereas I mean only to convey that the pleading regime is more demanding than under *Conley*.

If parties adjust their behavior when they perceive that the pleading standard has risen, then no amount of data on Rule 12(b)(6) motion filing rates, grant rates, or the number of cases filed will allow observers to determine whether courts' adjudication of a given set of cases has changed.

The intuition behind this result is simple. In some cases, changing the pleading standard will change the plaintiff's payoff from filing suit, change the defendant's payoff from filing a Rule 12(b)(6) motion in the event that the plaintiff does file suit, or both. Further, in some cases these changes in the parties' payoffs from making various procedural choices will change the parties' ability to settle at all.

Consequently, switching to more demanding pleading will cause what I have previously called party selection[15] to operate in multiple directions. Defendant selection occurs in cases for which a defendant would file a Rule 12(b)(6) motion under *Twombly/Iqbal* but not under *Conley*. Plaintiff selection occurs in cases in which the plaintiff would file suit under *Conley* but not under *Twombly/Iqbal*. And settlement selection occurs in cases that would settle—whether before or after the plaintiff files suit—under one pleading standard but not the other. The model generates several formal results, some of which track the heuristic description I provided in my earlier work on *Twombly* and *Iqbal*.[16] Thus, one contribution of the present paper is to provide a rigorous basis for the analytical claims I made in that work, as well as to provide a rigorous foundation for the empirical approach developed there.[17]

The model generates other important results, too.I show that when the pleading standard is held fixed—meaning essentially that the parties to each controversy have fixed beliefs concerning the probability that a Rule 12(b)(6) motion would be granted if filed—there are equilibria of the litigation game in which the Rule 12(b)(6) motion filing and grant rates each take on any value between zero and 100 percent. And, I show that on the whole, a switch to more demanding pleading redistributes payoffs from defendants to plaintiffs.

Before I continue, I should address an apparent threshold issue, concerning whether *Twombly* and *Iqbal* actually changed anything in the real world. To be sure, lower courts sometimes—maybe even often—strayed from the no-set-of-facts test.[18] But twice in ten years— in *Leatherman v. Tarrant County Narcotics Intelligence & Coordination Unit*, 507 U.S. 163 (1993) (C.J. Rehnquist) and in *Swierkiewicz v. Sorema N. A.*, 534 U.S. 506 (2002) (J. Thomas)— the Supreme Court unanimously swatted down lower courts' imposition of more demanding pleading requirements in civil rights and employment discrimination cases, re-affirming *Conley*

---

[15] *See* Gelbach, *Locking the Doors to Discovery*, *supra* note 11, at 11.

[16] *See* Gelbach, *Locking the Doors to Discovery*, *supra* note 11, Parts III-V.

[17] As I discuss in *supra* note 74, *infra*, there is one minor error in the characterization of party behavior in *Locking the Doors to Discovery*; correcting this error has no impact on any important conclusion in that paper.

[18] For scholarly discussions of trends against the *Conley* no-set-of-facts test, *see, e.g.,* Christopher M. Fairman, *The Myth of Notice Pleading*, 45 Arizona L. Rev. 987, 988 (2003) ("From antitrust to environmental litigation, conspiracy to copyright, substance specific areas of law are riddled with requirements of particularized fact-based pleading."); Richard L. Marcus, *The Revival of Fact Pleading Under the Federal Rules of Civil Procedure*, 86 Colum. L. Rev. 433, 436 (1986) ("federal courts are insisting on detailed factual allegations more and more often, particularly in securities fraud and civil rights cases").

and the no-set-of-facts test.[19] Moreover, as Richard Marcus has pointed out,[20] the fact that the Private Securities Litigation Reform Act of 1995[21] substantially strengthened the letter of the the pleading standards law in securities litigation.[22] If pleading were already so strict in general, then this Congressional action would have had no effect. These developments suggest that as of the moment the Supreme Court granted *certiorari* in *Twombly* in 2007, *Conley* was alive and well in spirit, even if not in every jot and tittle.[23] In any event, what matters for my purposes is simply that some of the parties to controversies *believe* that *Twombly* and *Iqbal* elevated pleading standards, and that point hardly seems even arguable.

The balance of this paper proceeds as follows. In Part II, I briefly discuss the previous empirical literature concerning *Twombly* and *Iqbal*'s effects on Rule 12(b)(6) motion adjudication, as well as the theoretical literature on litigation selection. Part III introduces my model and presents some basic analytical results. In Part IV, I present my results concerning equilibrium Rule 12(b)(6) motion filing and grant rates. In Part V, I consider the effects of switching to more demanding pleading on the parties' behavior and their equilibrium payoffs. There I derive the result that hypotheses concerning judicial behavior cannot be isolated or tested empirically when party selection might occur. Part VI concerns several normative issues—how switching to more demanding pleading affects the parties' welfare, how one might tell measure the quality of cases affected by a switch to more demanding pleading, and how switching to more demanding pleading might affect primary behavior. Finally, I conclude in Part VII.

---

[19] *See Leatherman*, 507 U.S. 163, 168 (1993) (holding, in a civil rights claim brought pursuant to 42 U.S.C. 1983, that "[i]n *Conley v. Gibson* . . . we said in effect that the Rule meant what it said: . . . all the Rules require is a short and plain statement of the claim that will give the defendant fair notice of what the plaintiff's claim is and the grounds upon which it rests.") (quotation marks omitted) (also pointing out that Rule 9(b) requires pleading with particularity when fraud or mistake are involved, so that any requirement of particularity in the § 1983 context is foreclosed by the canon of interpretation *[e]xpressio unius est exclusio alterius*); and *Swierkiewicz*, 534 U.S. 506, 512 (2002) (holding that "imposing [a] more demanding pleading standard in employment discrimination cases conflicts with [Rule] 8(a)(2)" and quoting *Conley v. Gibson*, 355 U.S. 41, 47 (1957), for the proposition that the complaint need only "give the defendant fair notice of what the plaintiff's claim is and the grounds upon which it rests").

[20] Richard L. Marcus, *The Puzzling Persistence of Pleading Practice*, 76 Tex. L. Rev. 1749, 1765 (1998).

[21] Pub. L. 104-52, 109 Stat. 737.

[22] Indeed, the PSLRA imposes conditions considerably stronger than those in Rule 9 in many securities cases. 15 U.S.C. 78u-4(b)(2)(A) ("the complaint shall, with respect to each act or omission alleged to violate this chapter, state with particularity facts giving rise to a strong inference that the defendant acted with the required state of mind").

[23] I should note, as well, that some observers do not believe *Twombly* and *Iqbal* changed very much, either doctrinally or practically. *See, e.g.*, *Evaluating the Supreme Court's Decisions in* Twombly *and* Iqbal*: Hearing Before the S. Comm. on the Judiciary*, 111th Cong. 11 (2009), available at http://judiciary.senate.gov/pdf/12-02-09%20Garre%20Testimony.pdf (statement of Gregory G. Garre, Partner, Latham & Watkins LLP, and former Solicitor Gen. of the United States) ("[T]he Supreme Court's decisions in *Twombly* and *Iqbal* . . . are firmly grounded in decades of prior precedent at both the Supreme Court and federal appellate court level . . . ."); Daniel R. Karon, "'Twas Three Years After Twombly and All Through the Bar, Not a Plaintiff Was Troubled from Near or from Far"—The Unremarkable Effect of the U.S. Supreme Court's Re-Expressed Pleading Standard in Bell Atlantic Corp. v. Twombly, 44 U.S.F. L. REV. 571, 572 (2010) ("*Twombly* is remarkable only for its unremarkability . . . ."); as well as other sources cited in Gelbach, *Locking the Doors to Discovery*, *supra* note 11, at 2285 n.57.

II. Previous Literature

In this Part I review two strands of literature with which the present paper makes direct contact. My discussion will be self-consciously uncomprehensive in the interests of brevity. Given its enormity, I will not review the enormous literature on the history, doctrine, and normative evaluation of *Twombly* and *Iqbal*.[24]

II.A. *The Empirical Literature on Rule 12(b)(6) Motion Grant Rates*

The first strand of literature to discuss concerns comparisons of Rule 12(b)(6) motion grant rates observed before *Twombly* and observed either after *Iqbal*, or in the period between *Twombly* and *Iqbal*. By my count, there have been nine studies whose empirical results were collected by the studies' authors and had not previously been reported.[25] Four of these—Kendall W. Hannon's student Note,[26] Joseph Seiner's two papers on employment discrimination[27] and disability[28] cases, and William Hubbard's paper[29]—compare the pre-*Twombly* period only to parts of the post-*Twombly* period before *Iqbal* was handed down. Of the other five, Patricia Hatamyar Moore's initial[30] and updated[31] studies and Raymond Brecia's[32] each consider the pre-*Twombly* and post-*Iqbal* periods, as well as the period between *Twombly* and *Iqbal*. The

---

[24] Among other sources, interested readers might consult Mark Herrmann, James M. Beck & Stephen B. Burbank, Debate, *Plausible Denial: Should Congress Overrule* Twombly *and* Iqbal?, 158 U. PA. L. REV. PENNUMBRA 141, 145 (2009), http://www.pennumbra.com/debates/pdfs/PlausibleDenial.pdf; Arthur R. Miller, *From* Conley *to* Twombly *to* Iqbal*: A Double Play on the Federal Rules of Civil Procedure*, 60 DUKE L.J.1, 10 (2010); Robert G. Bone, *Plausibility Pleading Revisited and Revised: A Comment on* Ashcroft v. Iqbal, 85 NOTRE DAME L. REV. 849 (2010); Robert G. Bone, Twombly, *Pleading Rules, and the Regulation of Court Access*, 94 IOWA L. Rev. 873 (2009); Brian T. Fitzpatrick, Twombly *and* Iqbal *Reconsidered*, 87 Notre Dame L. Rev. 1621 (2012); and Edward H. Cooper, *King Arthur Confronts TwIqy Pleading*, 90 OR. L. REV. 955 (2012).

[25] *See*, roughly in chronological order, Kendall W. Hannon, Note, *Much Ado About Twombly? A Study on the Impact of* Bell Atlantic Corp. v. Twombly *on 12(b)(6) Motions*, 83 NOTRE DAME L. REV. 1811 (2008); Joseph A. Seiner, *The Trouble with* Twombly: *A Proposed Pleading Standard for Employment Discrimination Cases*, 2009 U. ILL. L. REV. 1011; Joseph A. Seiner, *Pleading Disability*, 51 B.C. L. REV. 95 (2010); Patricia W. Hatamyar, *The Tao of Pleading: Do* Twombly *and* Iqbal *Matter Empirically?*, 59 AM. U. L. REV. 553 (2010); William H.J. Hubbard, *The Problem of Measuring Legal Change, with Application to* Bell Atlantic v. Twombly (Univ. of Chi. Law & Econ., Olin Working Paper No. 575, 2011), available at http://ssrn.com/abstract=1883831; original FJC report, *supra* note 11; in Joe S. Cecil, George W. Cort, Margaret S. Williams & Jared J. Bataillon, FED. JUDICIAL CTR., UPDATE ON RESOLUTION OF RULE 12(B)(6)MOTIONS GRANTED WITH LEAVE TO AMEND: REPORT TO THE JUDICIAL CONFERENCE ADVISORY COMMITTEE ON CIVIL RULES (2011) (hereinafter "the updated FJC report"), available at http://www.fjc.gov/public/pdf.nsf/lookup/motioniqbal2.pdf/$file/motioniqbal2.pdf; Raymond H. Brescia, *The* Iqbal *Effect: The Impact of New Pleading Standards in Employment and Housing Discrimination Litigation*, 100 KY. L.J. 235 (2011-2012); Patricia Hatamyar Moore, *An Updated Quantitative Study of* Iqbal*'s Impact on 12(b)(6) Motions*, 45 U. RICH. L. Rev. 603 (2012)

[26] Hannon, *Much Ado About* Twombly, *supra* note 25.

[27] Seiner, *The Trouble with* Twombly, *supra* note 25.

[28] Seiner, *Pleading Disability*, *supra* note 25.

[29] Hubbard, *The Problem of Measuring Legal Change*, *supra* note 25.

[30] Hatamyar, *The Tao of Pleading*, *supra* note 25.

[31] Hatamyar Moore, *An Updated Quantitative Study*, *supra* note 25.

[32] Brescia, *The* Iqbal *Effect*, *supra* note 25.

original[33] and updated[34] reports conducted by the staff of the Federal Judicial Center, led by Joe Cecil, consider only pre-*Twombly* and post-*Iqbal* periods. Most of the studies consider all or most case types,[35] though three focus on case types related to discrimination or constitutional civil rights.[36] All but the Hubbard and FJC studies construct their data sets using searches of electronic case databases, *e.g.*, Westlaw.[37]

Summarizing such a large number of studies always risks painting with too broad a brush.[38] But in the interests of brevity and consistency, I will simply quote a summary from my own recent paper concerning the studies other than the two FJC reports (I have altered the footnotes included in this summary so that they run sequentially with the present paper and cite to the proper notes *supra*, but the content of these notes is identical to that in my earlier paper):[39]

- They tend to find relatively little difference in MTD grant rates across their pre-*Twombly* and post-*Twombly*/pre-*Iqbal* periods.[40]

- They tend to find differences in the MTD grant or denial rate that range between zero and ten percentage points across their *Conley* and post-*Iqbal*

---

[33] Cecil et al, MOTIONS TO DISMISS FOR FAILURE TO STATE A CLAIM AFTER *IQBAL*, *supra* note 25.

[34] Cecil et al, UPDATE ON RESOLUTION OF RULE 12(B)(6)MOTIONS GRANTED WITH LEAVE TO AMEND, *supra* note 25.

[35] Hannon, *supra* note 25; original FJC report *supra* note 11; updated FJC report, *supra* note 25; Hatamyar, *supra* note 25; Hatamyar Moore, *supra* note 25; Hubbard, *supra* note 25. To be sure, all of these studies exclude some case types, among which are typically those involving PSLRA cases (*see* text surrounding *supra* note 22 for discussion of changes to pleading policy worked by the PSLRA) and *pro se* cases (*see* discussion of *Erickson v. Pardus*, 551 U.S. 89 (2007), original FJC report, *supra* note 25, at 6 n.10).

[36] *See Brescia, supra* note 25; Seiner, *The Trouble with* Twombly, *supra note* 25; Seiner, *Pleading Disability*, *supra* note 25.

[37] For a criticism of this approach , *see* Joe S. Cecil, *Of Waves and Water: A Response to Comments on the FJC Study Motions to Dismiss for Failure to State a Claim after* Iqbal (Mar. 19, 2012 draft), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2026103, at 4 ("missing orders in the Westlaw database are more likely to deny motions to dismiss, thereby building a bias into studies that rely on such sources"); *see also* Cecil et al, MOTIONS TO DISMISS FOR FAILURE TO STATE A CLAIM AFTER *IQBAL*, *supra* note 25, at 37 n.47.

[38] For example, Joe Cecil cries foul in Cecil, *Of Waves and Water, supra* note 37, at 39, writing that I erred in my earlier paper when, characterizing results in the FJC staff's reports (*see* original FJC report, *supra* note 11, and updated FJC report, *supra* note 25) I stated that "the multivariate results are broadly consistent with the simpler differences in grant rates", Gelbach, *Locking the Doors to Discovery*, *supra* note 11, at 2291 n.94. I tried to do too much in one phrase in my earlier paper, as the "broadly consistent" characterization would be accurate concerning the difference in simple and multivariate results for Rule 12(b)(6) motion *filing* rates, which was the context in which I dropped the footnote—and the ball—concerning grant rates. Thus I most readily concede that Cecil is right that the pattern of results in the FJC staff reports' multivariate models for grant rates does differ from the pattern of changes in simple grant rates. I should have provided a separate discussion of the FJC reports' multivariate results for grant rates, noting the most appropriate reason to disregard the multivariate grant rate results—which is that there are good reasons to think the multivariate models in question are importantly mis-specified. I provide a brief discussion of this point *infra*, in *supra* note 86.

[39] The summary, in the bulleted-form quoted herein, is taken verbatim from Gelbach, *Locking the Doors to Discovery*, *supra* note 11, at 2289.

[40] *See, e.g.*, Brescia, *supra* note 25, at 29-30; Hannon, *supra* note 25, at 1836; Hatamyar, *supra* note 25, at 597-99; Hubbard, *supra* note 70, at 28.

periods,[41] with larger differences for cases involving civil rights of one type or another.[42]

- They find either small or no changes in the rate at which MTDs are granted without leave for the plaintiff to amend her complaint, and sizable increases in the rate they are granted with leave to amend.[43]

As for the FJC reports, I discuss their methodology and findings in some detail in my earlier paper,[44] and they have have been the subject of extensive debate elsewhere as well.[45] For present purposes it will be sufficient to let the reports speak for themselves on three points. The original report found that "[t]here was a general increase from 2006 to 2010 in the rate of filing of motions to dismiss for failure to state a claim."[46] Second, it found that "[i]n general, there was no increase in the rate of grants of motions to dismiss without leave to amend," including in "civil rights cases and employment discrimination cases."[47] Third, with the exception of "cases challenging mortgage loans," the authors did not "find an increase in the rate of grants of motions to dismiss without leave to amend."[48]

The relevant aspect of the updated FJC report for my purposes is that the authors followed cases in which a Rule 12(b)(6) motion had been granted with leave to amend, in order to determine how frequently plaintiffs filed amended complaints and how frequently defendants challenged such complaints with subsequent Rule 12(b)(6) motions. The authors "identified cases in which the movant prevailed as those in which the court granted the last motion to dismiss in whole or in part and no opportunity to amend the complaint remained."[49] The authors

---

[41] *See, e.g.*, Hatamyar Moore, *supra* note 25, at 7-8 (finding that the MTD denial rate "fell from 26% under *Conley* to . . . 17% under *Iqbal*").

[42] For example, in her earlier paper, Hatamyar Moore finds that grant rates for MTDs in cases she codes as involving constitutional civil rights rose from 50% under *Conley* to 55% under *Twombly* to 60% under *Iqbal*. Hatamyar, *supra* note 25, at 556. Her later paper shows a much larger *Conley-Iqbal* difference, roughly twenty-three percentage points, when she follows the FJC's methodology and excludes cases with pro se plaintiffs, but her samples in this comparison include just sixty-two cases under *Conley* and fifty-three post-*Iqbal*, so it is unclear what to make of this result. Hatamyar Moore, *supra* note 25, at 11-12.

[43] *See, e.g.*, Brescia, *supra* note 25, at 36-37; Hatamyar, *supra* note 71, at 598 tbl.1; Hatamyar Moore, *supra* note 25, at 7 (showing that the rate at which MTDs are granted without leave to amend was 40% in both her *Conley* and *Iqbal* samples but indicating "[t]he percentage of 12(b)(6) motions granted in full with leave to amend increased from 6% under *Conley* to . . .21% under *Iqbal*").

[44] *See* Gelbach, *Locking the Doors to Discovery*, *supra* note 11, at 2290-2294.

[45] *See* Lonny Hoffman, Twombly *and* Iqbal's *Measure: An Assessment of the Federal Judicial Center's Study of Motions To Dismiss*, 6 FED. CTS. L. REV. 1 (2011); Hatamyar Moore, *supra* note 25, at 632-652; Brescia, *supra* note 25; Cecil, *Of Waves and Water*, *supra* note 37; Lonny Hoffman, *Rulemaking in the Age of* Twombly *and* Iqbal, draft, August 1, 2012, available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2123325.

[46] Original FJC report *supra* note 11, at vii.

[47] Id.

[48] *Id*. Changes in the observed pattern of adjudication in such cases might well "be due to changing economic conditions involving the housing market and [be] unrelated to the recent Supreme Court decisions driven by the financial crisis rather than any change in pleading policy," as the report's authors note. *Id.*, at 21.

[49] Cecil et al, UPDATE ON RESOLUTION OF RULE 12(B)(6)MOTIONS GRANTED WITH LEAVE TO AMEND, *supra* note 25, at 3.

summarize how the qualitative results of the updated study compare to the original one with the helpfully concise statement that "[o]ur conclusions remain the same."[50]

Authors in this literature have argued over various methodological details, as well as over the proper interpretation of studies' results. I believe these debates are largely beside the point. A central thesis of my earlier paper, and a central result of the more rigorous analysis I provide below, is that when there is party selection, changes in observed grant rates simply don't tell us anything meaningful. Thus, arguments over what the studies discussed above really show can't help but miss the boat.

II.B. *The Literature on Selection in Litigation*

The importance of selection in empirical litigation studies has been well understood for nearly three decades, following the publication of George Priest and Benjamin Klein's seminal paper on the topic.[51] If the parties have common beliefs concerning the costs and stakes involved in the case, as well as the likelihood that either party will prevail, then they will have a powerful economic incentive to avoid costly litigation by settling. Priest & Klein's paper might be best known for its suggestion that selection should cause the plaintiff's win rate in litigated cases to be approximate 50%,[52] regardless of the merits of disputes on the whole. William Hubbard has recently suggested, *inter alia*, that similar logic applies to the Rule 12(b)(6) motion grant rate, so that "motions to dismiss will be granted a large fraction (but less than half) the time."[53] Hubbard's paper includes a model, but his suggestion must be regarded as conjectural, since he does not derive any analytical results. My own Proposition 2, *infra*, shows that while Hubbard's conjecture might be correct in some model, it can't be correct generally. It is unclear,

---

[50] *Id.*, at 1.

[51] George L. Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 J. LEGAL STUD. 1 (1984).

[52] On the 50% claim, *see, e.g.,* Priest & Klein, *supra* note 51, at 5 ("plaintiff victories will tend toward 50 percent whether the legal standard is negligence or strict liability, whether judges or juries are hostile or sympathetic") (though, "the relative stakes to the parties will greatly influence the rate of success in litigation and are likely to be the principal reason why success rates differ from the 50 percent baseline"). For a discussion of this result's sensitivity to modifications in modeling details across a class of litigation models, *see* Keith Hylton & Haizhen Lin, *A Formalization And Extension Of The Priest-Klein Hypothesis*, (Boston University School of Law Working Paper No. 11-28, June 14, 2011), http://www.bu.edu/law/faculty/scholarship/workingpapers/documents/HyltonK-LinH061411correctedpaper.pdf. One theme in the literature concerns the distinction between models in which the parties have common knowledge but disagree, as in the model I develop in Part III, *infra*, and models built on asymmetric information. *See, e.g.,* Steven Shavell, *Any Frequency of Plaintiff Victory at Trial Is Possible*, 25 Journal of Legal Studies 493 (1996) for such a model; *see* Keith Hylton & Haizhen Lin, *Trial Selection Theory: A Unified Model*, (Boston University School of Law Working Paper No. 10-12, revision of May 2011), http://www.bu.edu/law/faculty/scholarship/workingpapers/documents/HyltonK-LinH-05-17-2010-may2011rev.pdf for a joint treatment of common knowledge-disagreement and information asymmetry. There have also been numerous empirical papers that consider the 50% hypothesis; examples include Daniel Kessler, Thomas Meites & Geoffrey Miller, *Explaining Deviations from the Fifty-Percent Rule: A Multimodal Approach to the Selection of Cases for Litigation*, 25 J. Legal Stud. 233 (1996); Peter Siegelman & John J. Donohue III, *The Selection of Employment Discrimination Disputes for Litigation: Using Business Cycle Effects To Test the Priest-Klein Hypothesis*, 24 J. LEGAL STUD. 427 (1995); Peter Siegelman & Joel Waldfogel, *Toward a Taxonomy of Disputes: New Evidence Through the Prism of the Priest/Klein Model*, 28 J. LEGAL STUD. 101 (1999); and Theodore Eisenberg, *Testing the Selection Effect: A New Theoretical Framework with Empirical Tests*, 19 J. Legal Stud. 337 (1990).

[53] Hubbard, *supra* note 25, at 18.

at least to me, how Hubbard's model incorporates the dynamic and strategic aspects of case and Rule 12(b)(6) motion filing; differences in our modeling approach might help explain the difference between my result and Hubbard's conjecture.

In another post-*Twombly* paper, Keith Hylton has addressed a different question from mine, the optimal pleading standard.[54] Hylton's focus on normative policy design is self-consciously not targeted at developing empirically testable conclusions, so it is somewhat orthogonal to the present paper. Also, while Hylton's discussion is couched in multi-stage terms, he does not analyze the strategic aspects of the defendant's actions in choosing whether to file Rule 12(b)(6) motions.[55]

This discussion underscores an important distinction between the present paper and the limited previous theoretical work on pleading. Unlike other papers in the literature, the model here focuses on the sequential strategic implications of the way Rule 12(b)(6) works in practice: defendants must decide whether to file Rule 12(b)(6) motions, and if they do, both defendants and plaintiffs must bear some costs of litigating them.[56]

III. A Game-Theoretic Model of Pre-Discovery Procedure, Including Rule 12(b)(6)

In this Part, I introduce two versions of a simple game-theoretic model that captures key aspects of pre-discovery process in federal litigation. I lay out the basic model in section III.A. As with much of the law and economics literature on litigation, disagreement in the parties' beliefs concerning certain parameters of the litigation is what drives these results. Throughout, I will assume that the parties' beliefs are common knowledge: each party knows the other's beliefs, even if she thinks they are mistaken. This assumption rules out the possibility that litigation occurs because of miscommunication or other problems related to information asymmetry, rather than because of disagreement.[57] I will also generally assume that the number of controversies that might become lawsuits is fixed (I briefly discuss effects of pleading policy on primary behavior, and thus the number of controversies, in Part VI, *infra*). Finally, I assume

---

[54] Keith N. Hylton, *When Should a Case Be Dismissed? The Economics of Pleading and Summary Judgment Standards*, 16 SUP. CT. ECON. REV. 39 (2008).

[55] That is, Hylton's analysis seems to take for granted that courts will conduct a *sua sponte* assessment of the complaint.

[56] This aspect of my model also distinguishes it from other important sequential models of litigation, like those in Lucian Arye Bebchuk, *A New Theory Concerning the Credibility and Success of Threats To Sue*, 25 J. LEGAL STUD. 1 (1996) and Joseph A. Grundfest & Peter H. Huang, *The Unexpected Value of Litigation: A Real Options Perspective*, 58 STAN. L. REV. 1267 (2006).

[57] *See, e.g.,* Bone, *Civil Procedure: The Economics of Civil Procedure*, and *Modeling Frivolous Suits*; Avery Katz, *The Effect of Frivolous Lawsuits on the Settlement of Litigation*, 10 International Review of Law and Economics 3 (1990); Shavell, *supra* note 52; Keith N. Hylton**,** *Asymmetric Information and the Selection of Disputes for Litigation*, 22 *Journal of Legal Studies* 187 (1993); Keith N. Hylton and Haizhen Lin, *Trial Selection Theory*, *supra* note 52; and Lucian Arye Bebchuk, *Litigation and Settlement Under Imperfect Information*, 15 RAND Journal of Economics 404 (1984).

there is only one claim in every lawsuit, which simplifies the discussion.[58] In section III.B, I analyze and solve the litigation game using backward induction.

III.A. *The Basic Model*

The stages of litigation I consider in this model are as follows:

Stage 1: The plaintiff decides whether to drop the case (action "Drop") or settle it for an amount the defendant is willing to pay (action "S"); in either event the case terminates. Alternatively, the plaintiff can file suit (action "F").

Stage 1′: Nature moves, choosing a judge at random (for notational simplicity, I omit nature's action from the game history in much of the discussion to come).

Stage 2: If the plaintiff has filed suit, then the defendant has three options. First, she can settle the case for an amount the plaintiff is willing to accept (action "S"), in which event the case terminates. Second, the defendant can file an answer (action "A"), in which event the case continues; as I discuss below, I assign continuation values as parties—best possible expected future—payoffs in the event that the defendant files an answer, which allows me to treat the filing of an answer as terminal for analytical purposes. Finally, the defendant can file a Rule 12(b)(6) motion to dismiss (action "M").[59]

Stage 3: If the defendant files a Rule 12(b)(6) motion, then the judge decides whether to deny the motion (action "Deny") or grant it (action "Grant"). I assume that all grants are with prejudice, *i.e.*, without leave to amend,[60] so the case terminates if the judge grants the motion. I use γ to denote the probability that the judge in a given case will grant a Rule 12(b)(6) motion.

Stage 4: If the judge denies the defendant's Rule 12(b)(6) motion, the defendant must either answer or settle—action "A" or action "S"—since she cannot file a

---

[58] Among other things, this assumption allows me to treat filing a Rule 12(b)(6) motion to dismiss and filing an answer as mutually exclusive in Stage 2 of the litigation game; *see* Gelbach, *Locking the Doors to Discovery*, *supra* note 11, at 2296 n.122 for discussion on this point.

[59] It is important to note that the model here does not map exactly to real-world procedural practice. For example, I assume here that if the defendant files a Rule 12(b)(6) motion, she bears no discovery costs unless the motion is denied. But as Kevin Lynch has pointed out, filing a Rule 12(b)(6) motion does not automatically stay discovery as a general matter—defendants generally must move for a stay, or alternatively a protective order under Rule 26(c); *see* Kevin J. Lynch, *When Staying Discovery Stays Justice: Analyzing Motions to Stay Discovery When a Motion to Dismiss is Pending*, 47 Wake Forest L. Rev. 71 (2012). I do not believe that accounting for such facts would change my qualitative results in any important way.

[60] Using the one-stage deviation principle, Drew Fudenberg and Jean Tirole, *Game Theory*, MIT Press (1992), § 4.2, it is straightforward to solve a generalized version of the infinite-horizon model that allows grants with leave to amend. But incorporating this aspect of realism comes at the price of substantially increased notational and analytical complexity, which are largely divorced from key intuitions about pleading policy and party selection. For these reasons, I assume all Rule 12(b)(6) motion grants are with prejudice.

new Rule 12(b)(6) motion once her original motion has been denied. In either of these events, the game terminates for analytical purposes.

Figure 1 lays out the structure of the game tree just described. Given its size, I use the shorthand approach of drawing Nature's branch as a dashed line, indicating that the judge's identity is not in the plaintiff's information set as of Stage 1.[61]

Figure 1: The Basic Game Tree
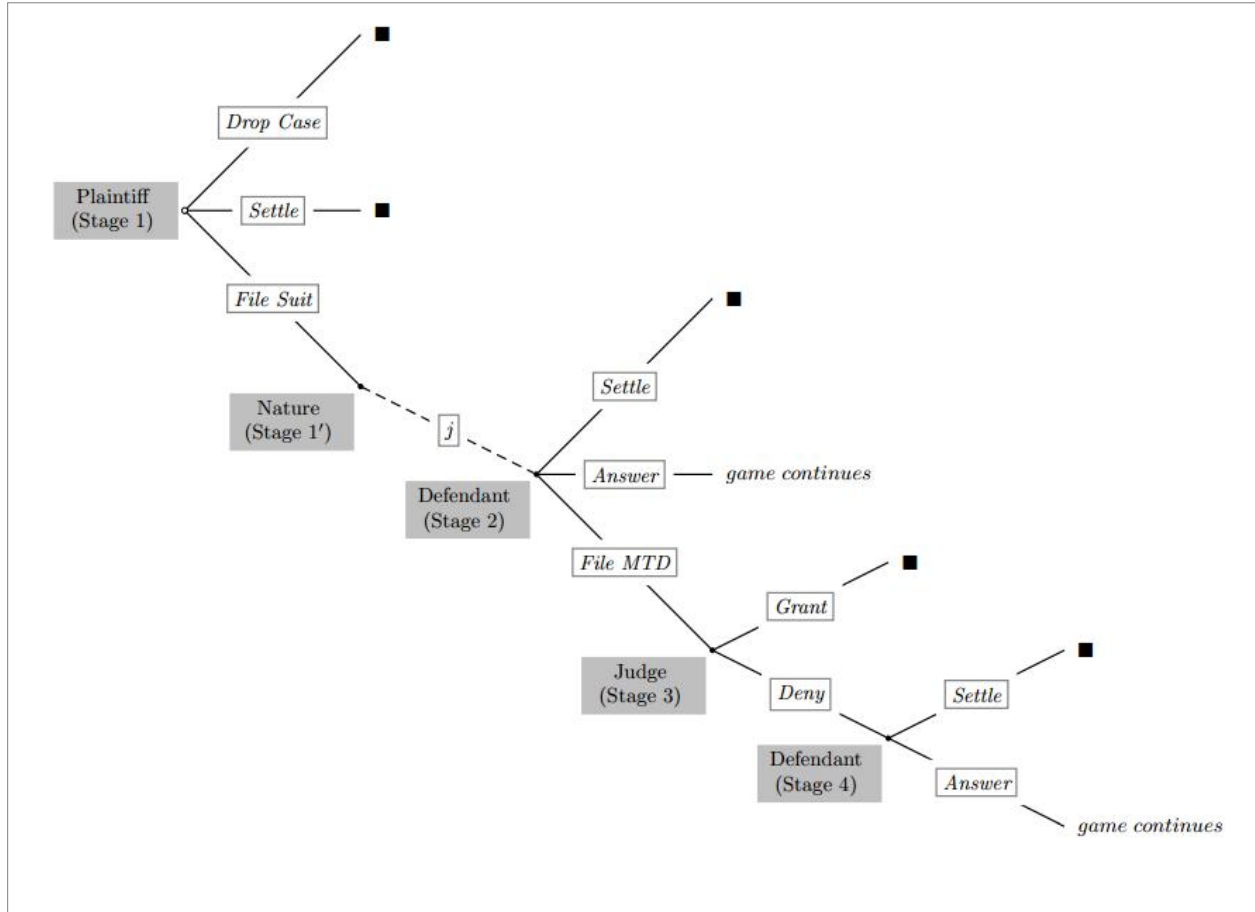
To make the model useful requires that I quantify the parties' payoffs in each of the terminal nodes in Figure 1. Denote the plaintiff's payoff given history h by $u_p(h)$ and the defendant's by $u_d(h)$. I normalize payoffs to zero when the plaintiff drops the case, so $u_p(Drop)=u_d(Drop)=0$. I assume there is no cost to settling in either Stage 1 or Stage 2, and I denote the Stage 1 settlement payment amount by S_1. If the case settles before the plaintiff files suit, the plaintiff's payoff will be $u_p(S)=S_1$ and the defendant's will be $u_d(S_1) = -S_1$, with the sum

---

[61] Technically, Nature's node should have a separate branch for each possible judge to which the case might be assigned, with probability 1/J assigned to the branch representing judge j, with a total of J possible judges. Each of the game's parameters would then have to have a judge index.

of payoffs being zero. I assume that if the plaintiff does file suit, she must pay $c_f > 0$ to cover the cost of filing her complaint.

Let S_2 be the amount of any Stage 2 settlement, one that occurs after the plaintiff files suit but before the defendant files an answer or Rule 12(b)(6) motion. If the parties settle in Stage 2, the plaintiff's payoff will be $u_p(F,S) = S_2 - c_f$, and the defendant's will be $u_d(F,S) = -S_2$. Note that the sum of payoffs in this event is $-c_f$. Thus, with full information at the pre-plaintiff-filing stage, the parties would always be better off settling before the plaintiff files suit than afterward: forcing the plaintiff to file suit chews up c_f in litigation costs, creating surplus from settling in Stage 1 rather than Stage 2. If the parties' information sets did not change as a result of filing, then in equilibrium no case would ever settle after Stage 1. However, the presence of Stage 1′ in the game tree shows that the parties will not generally know which judge will be assigned to the case until after the plaintiff files suit. Since the model allows the parties to wait until they know the judge's identity to form specific beliefs about the key parameters related to litigation costs, gains, and probabilities,[62] it does allow the sub-history "F,S_2" to occur in equilibrium. This flexibility is desirable since the parties may well believe some judges will be more sympathetic than others to their side.[63] To economize on notation, though, I will leave the judge's identity implicit except in analyzing Stage 1.

Let c_{da} be the continuation value of the defendant's total subjective expected cost from filing an answer (mnemonically, c_{da} is the cost to the defendant of answering). This continuation value is the minimum cost associated with filing an answer and then behaving optimally throughout the litigation that follows. Thus c_{da} includes not only the cost of drafting the pleading document itself, but also the expected future cost of settlement, discovery, litigating any summary judgment motions, a trial if there is one, and the cost of any judgment and appeal that ensue. Thus, the defendant's expected payoff if she files an answer is $u_d(F,A) = -c_{da}$. Let v_{pa} be the plaintiff's expected continuation value when the defendant has filed an answer, *i.e.*, the plaintiff's greatest possible expected payoff from the time of the answer's filing forward (mnemonically, v_{pa} is the value to the plaintiff from answering). Then the plaintiff's total payoff in the case when the defendant files a Stage 2 answer is $u_p(F,A) = (v_{pa} - c_f)$.

---

[62] Other types of information can be revealed when a case is filed. For example, the complaint itself might contain information previously unknown either to the defendant or her lawyer. Moreover, when the defendant does not know exactly the quality of the plaintiff's case, the plaintiff can signal that this quality is high by bearing the cost of filing the complaint, even if the words in the complaint itself do not themselves add anything to the defendant's information set. These forms of asymmetric information likely would fit comfortably into my model, since their revelation occurs before the defendant's Stage 2 action. I omit them for expositional simplicity. Other forms of asymmetric information might be more problematic to handle. For an interesting discussion of pleading in the presence of asymmetric information, see Robert G. Bone, *Civil Procedure: The Economics of Civil Procedure*, Foundation Press (2002), as well as Bone, *Modeling Frivolous Suits* 145 University of Pennsylvania Law Review 519, (1997); for discussions of asymmetric information in the literature on win rates and litigation selection, *see, e.g.*, Lucian Arye Bebchuk, *Litigation and Settlement Under Imperfect Information*, *supra* note 57; Avery Katz, *supra* note 57; Keith N. Hylton, *Asymmetric Information and the Selection of Disputes for Litigation*, *supra* note 57; Steven Shavell,*supra* note 52; and Keith N. Hylton and Haizhen Lin, *Trial Selection Theory*, *supra* note 52.

[63] For example, one review of district courts' doctrinal practice in the year following *Twombly* found variation in courts' Rule 12(b)(6) motion practice both within and across district courts. *See* Redish & Epstein, *supra* note 13, Part IV, at 26-30.

In the event that the plaintiff files suit and the defendant files a Rule 12(b)(6) motion, the judge can either grant or deny the motion. I write the defendant's belief about the probability the judge will grant the motion as γ_d, and I write the plaintiff's belief as γ_p; the parties may have different beliefs concerning the grant probability, so that γ_p need not equal γ_d.

Let c_{dm} and c_{pm} be the defendant and plaintiff costs of this part of the litigation. If the judge grants the Rule 12(b)(6) motion, the case terminates, and the defendant pays nothing to the plaintiff, so the parties' payoffs in the event of a granted Rule 12(b)(6) motion are $u_p(F,M,G) = -(c_f + c_{pm})$ for the plaintiff and $u_d(F,M,G) = -c_{dm}$ for the defendant. If the judge denies a filed Rule 12(b)(6) motion, we move on to Stage 4. If the parties settle the case in Stage 4, for a payment in the amount of S_4, then their payoffs will be $u_p(F,M,Deny,S) = S_4 - c_f - c_{pm}$ for the plaintiff and $u_d(F,M,Deny,S) = -(S_4 + c_{dm})$ for the defendant. If, instead, the defendant files a post-denial answer, the plaintiff's payoff will be $u_p(F,M,Deny,A) = v_{pa} - c_f - c_{pm}$, while the defendant's payoff will be $u_d(F,M,Deny,A) = -(c_{da} + c_{dm})$.

Figure 2 augments the basic Figure 1 game tree with the payoffs just discussed.

Figure 2: The Basic Game Tree, With Payoffs

13
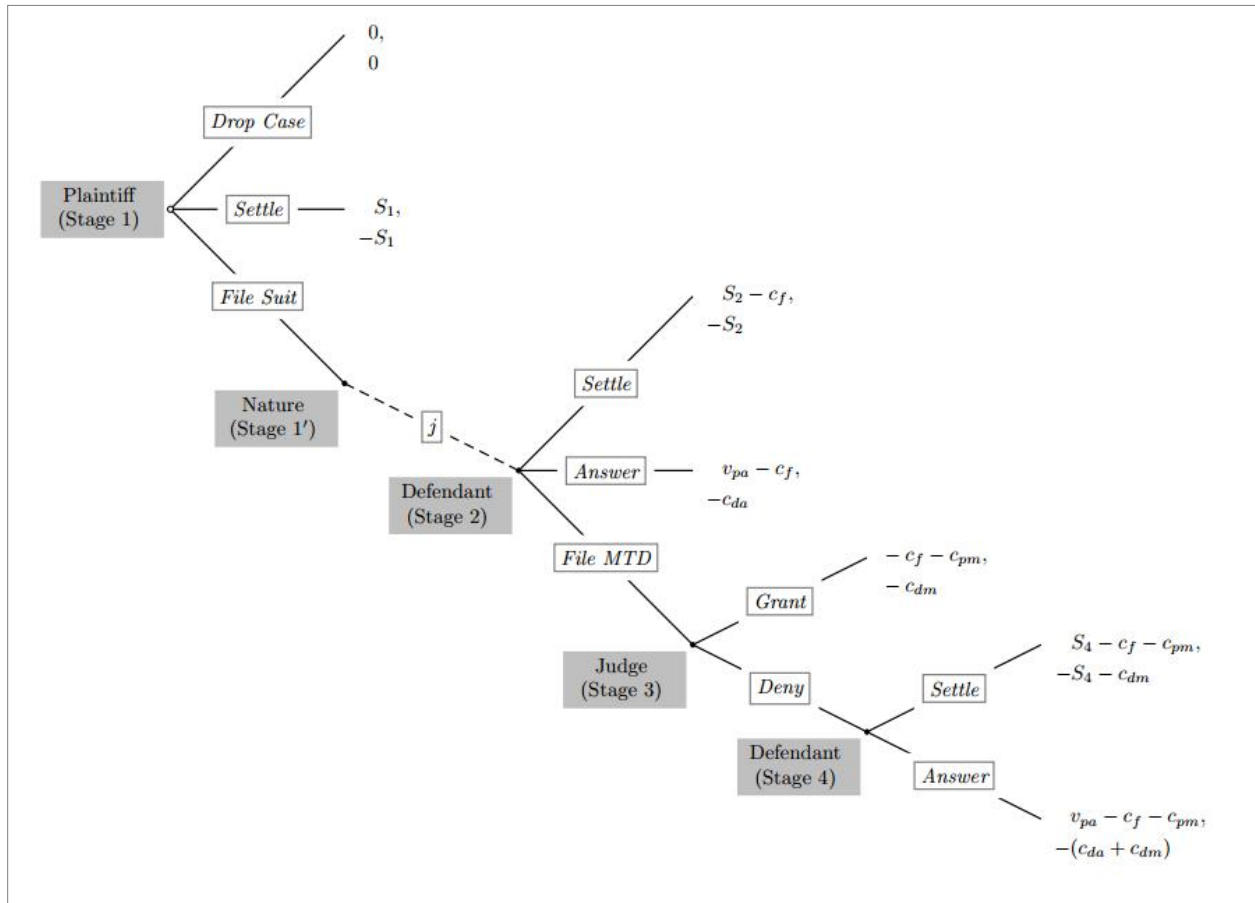
III.B. *Analyzing the Litigation Game Using Backward Induction*

In this section, I solve for the litigation game's subgame perfect Nash equilibrium via backward induction. The primary task in doing so is to determine the equilibrium settlement amounts that would occur if settlement occurred in Stage 2 or Stage 4, because these amounts are required to determine the parties' payoffs along every subgame path. I will adopt the usual approach to settlement: I assume that settlement occurs whenever there is positive surplus, and I assume that each party receives the same fraction of settlement surplus in each settlement state. I begin with Stage 4, under the assumption that the plaintiff has filed suit, the defendant has filed a Rule 12(b)(6) motion, and the judge has denied the motion. I then work my way back through the game to Stage 1.

III.B.1. Stage 4: Following history "F,M,Deny", the defendant chooses between settling and answering

When the judge has denied a filed Rule 12(b)(6) motion, the defendant must decide in Stage 4 whether to file an answer or settle for an amount to which the plaintiff will agree. If the defendant files an answer, then the parties' payoffs will be

$$u_d(F, M, Deny, A) = -(c_{da} + c_{dm}),$$

$$u_p(F, M, Deny, A) = v_{pa} - c_f - c_{pm}.$$

At this stage, the previously incurred litigation costs c_{dm} and c_{pm} are sunk. The parties will settle in Stage 4 if any surplus can be gained from avoiding the further litigation that the defendant would touch off by filing an answer. That will hhappen if and only if the continuation cost to the defendant exceeds the plaintiff's continuation value, i.e., $c_{da} > v_{pa}$.[64] Thhus, the parties will settle when $c_{da} > v_{pa}$, or, equivalently, when the Stage 4 surplus $\sigma_a \equiv c_{da} - v_{pa}$ is positive; the defendant will answer when the reverse holds.

Now let β be the share of the surplus that the plaintiff's bargaining power allows her to capture. Any settlement must transfer to the plaintiff the non-sunk part of her reservation value, v_{pa}, plus the amount $\beta\sigma_a$. The settlement must also leave the defendant paying her cost of answering, c_{da}, minus her part of the surplus, $(1-\beta)\sigma_a$. Therefore, the settlement amount in the event of a Stage-4 settlement is

$$S_4 = v_{pa} + \beta\sigma_a = c_{da} - (1 - \beta)\sigma_a. \tag{1}$$

Since it costs c_{dm} for the defendant to litigate a Rule 12(b)(6) motion, the defendant's payoff when the case reaches Stage 4 and then settles is

$$
\begin{aligned}
u_d(F, M, Deny, S) &= u_d(F, M, Deny, A) + (1 - \beta)\sigma_a \\
&= -c_{dm} - [c_{da} - (1 - \beta)\sigma_a],
\end{aligned} \tag{2}
$$

---

while similar logic shows that the plaintiff's payoff in the event of a Stage 4 settlement is

$$
\begin{aligned}
u_p(F, M, Deny, S) &= u_p(F, M, Deny, A) + \beta \sigma_a \\
&= v_{pa} - c_f - c_{pm} + \beta \sigma_a.
\end{aligned}
\tag{3}
$$

III.B.2. Stage 3: Following history "F,M", the Judge rules on the Rule 12(b)(6) motion.

At Stage 3, the judge will either grant or deny a filed Rule 12(b)(6) motion. When the judge grants the Rule 12(b)(6) motion, the defendant pays only the cost of litigating the Rule 12(b)(6) motion, so her payoff is $u_d(F,M,G) = -c_{dm}$. The plaintiff bears both the initial filing cost and her cost of litigating the Rule 12(b)(6) motion, so her payoff is $u_p(F,M,G) = -(c_f + c_{dm})$.

When the judge denies the Rule 12(b)(6) motion, the game will continue to Stage 4, where the defendant will choose to answer or settle according to whether $u_d(F, M, Deny, S)$ or $u_d(F, M, Deny, A)$ is greater. Thus,

$$
u_d(F, M, Deny) = \begin{cases} u_d(F, M, Deny, A), & \text{if } \sigma_a < 0, \\ u_d(F, M, Deny, S), & \text{if } \sigma_a > 0, \end{cases}
$$

which by (2) can be re-written as

$$
\begin{aligned}
u_d(F, M, Deny) &= u_d(F, M, Deny, A) + (1 - \beta)\max\left[\sigma_a, 0\right] \\
&= -[c_{dm} + c_{da}] + (1 - \beta)\max\left[\sigma_a, 0\right].
\end{aligned}
\tag{4}
$$

Similar reasoning shows that the plaintiff's continuation payoff in the event of a Stage 3 denial of the defendant's Stage 2 Rule 12(b)(6) motion may be written

$$
u_p(F, M, Deny) = \begin{cases} u_p(F, M, Deny, A), & \text{if } \sigma_a < 0, \\ u_p(F, M, Deny, S), & \text{if } \sigma_a > 0, \end{cases}
$$

which by (3) can be re-written as

$$
\begin{aligned}
u_p(F, M, Deny) &= u_p(F, M, Deny, A) + \beta \max\left[\sigma_a, 0\right] \\
&= v_{pa} - c_f - c_{pm} + \beta \max\left[\sigma_a, 0\right].
\end{aligned}
\tag{5}
$$

These ways of writing $u_d(F,M,Deny)$ and $u_d(F,M,Deny)$ show that each party's continuation payoff in the event of history "F,M,Deny" equals the payoff she gets when the defendant answers in stage 4, if there is no surplus plus the party's portion of any surplus from settlement.

III.B.3. Stage 2: Following history "F", the defendant chooses to Answer, File a Rule 12(b)(6) motion, or Settle.

Once the case has been filed, the defendant's optimal behavior is to take whichever of her three actions maximizes her payoffs. We have seen that the defendant's payoff if she answers the

complaint in Stage 2 is $u_d(F,A) = -c_{da}$, and her payoff if she settles for amount S_2 is $u_d(F,S) = -S_2$. Her expected payoff when she files a Rule 12(b)(6) motion must take into account her belief concerning the probability, γ_d, that the Rule 12(b)(6) motion will be granted, as well as the payoffs she receives in the alternative events of a grant or a denial. Thus, $u_d(F,M) = \gamma_d u_d(F,M,G) + (1-\gamma_d) u_d(F,M,\text{Deny})$. Plugging in $-c_{dm}$ for $u_d(F,M,G)$ and plugging in $[-c_{dm} - c_{da} + (1-\beta)\sigma_a]$ for $u_d(F,M,\text{Deny})$ (see (4)), the defendant's expected payoff when she files a Stage 2 motion to dismiss is

$$u_d(F,M) = -c_{dm} + (1-\gamma_d)\{-c_{da} + (1-\beta)\max[\sigma_a, 0]\}. \tag{6}$$

Thus, when the defendant files a Rule 12(b)(6) motion, she always pays c_{dm}, and in the $(1-\gamma_d)$ share of the time when she expects such motions to be denied, she must pay her continuation cost when the motion is denied.

To find the defendant's optimal Stage 2 action requires determining which of the three possible payoffs, $u_d(F,A)$, $u_d(F,S)$, and $u_d(F,M)$, is greatest. Since the various model parameters are common knowledge, both parties will be able to correctly predict the action the defendant would take in the absence of settlement. Thus, whether settlement actually would occur in Stage 2 will depend on which action—answer or file Rule 12(b)(6) motion—the defendant would take if settlement weren't an option. The defendant will be better off filing a Rule 12(b)(6) motion than an answer whenever $-u_d(F,M) < -u_d(F,A)$, or, plugging (6) into the left hand side of this inequality, whenever

$$c_{dm} + (1-\gamma_d)\{c_{da} - (1-\beta)\max[\sigma_a, 0]\} < c_{da}. \tag{7}$$

I will now discuss (7) under the two possible scenarios involving the sign of σ_a.

III.B.3.a. Optimal Stage 2 behavior when $\sigma_a < 0$

Consider first the situation when $c_{da} < v_{pa}$, so that $\sigma_a < 0$. In this situation, the parties would not settle in Stage 4 in the event of a denied Rule 12(b)(6) motion. Plugging $\max[\sigma_a, 0] = 0$ into (7), we see that the condition for the defendant to file a Rule 12(b)(6) motion holds whenever $c_{dm} + (1-\gamma_d)c_{da} < c_{da}$, or

$$\gamma_d > \frac{c_{dm}}{c_{da}}. \tag{8}$$

In this situation, the defendant regards the event of a granted Rule 12(b)(6) motion as equivalent to the event of avoiding the costs associated with filing an answer in Stage 4. Filing the Rule 12(b)(6) motion in Stage 2 costs the defendant c_{dm} in up-front litigation costs, but in return she gets probability γ_d of avoiding paying c_{da}. Think of filing a motion to dismiss as a risky investment with price c_{dm} and a return distribution whose support consists of zero and c_{da}. On that interpretation, condition (8) states that the Rule 12(b)(6) motion is worth filing when the probability of the positive payout exceeds the price per unit of return, which is c_{dm}/c_{da}. In general it is reasonable to think that the cost of litigating a motion to dismiss, c_{dm}, will be less than c_{da}, since the latter includes all the expected costs of future litigation. Thus the right hand side of (8) will generally be less than one, so there are some values of $\gamma_d \in (0,1)$ for which defendants would prefer filing a motion to dismiss to answering.

Figure 3 puts this discussion into graphical form. The figure shows the space of possible combinations of the parties' beliefs about the probability a motion to dismiss would be granted, with the defendant's belief on the horizontal axis and the plaintiff's belief on the vertical axis. The vertical line at $\gamma_d = c_{dm}/c_{da}$ separates the part of (γ_d, γ_p) space for which the defendant prefers filing a motion to dismiss from the part where the defendant prefers to answer.

Figure 3: Answer vs. Rule 12(b)(6) motion in Stage 2 When $c_{da} < v_{pa}$
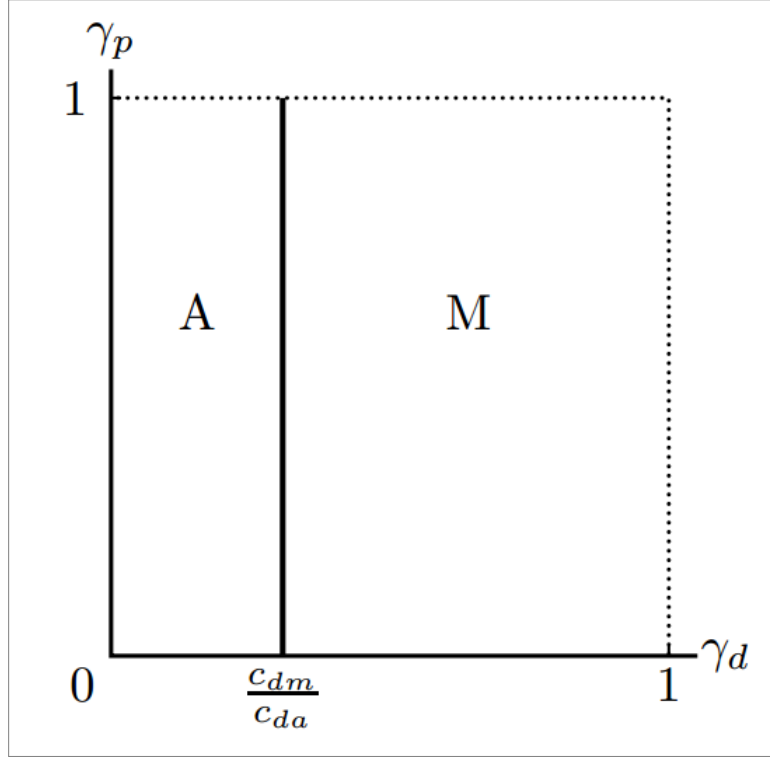
Now I assess whether the defendant would do better with a settlement than with an answer. Since the plaintiff's filing cost c_f is sunk, there will be surplus in Stage 2 if and only if $c_{da} > v_{pa}$, which contradicts our hypothesis that $\sigma_a < 0$. Therefore, when $\sigma_a < 0$ the parties will never settle when γ_d is located to the left of the vertical line in Figure 3. Next suppose that γ_d is located to the right of the vertical line in Figure 3. Then the plaintiff's expected payoff at the moment when the defendant files a Rule 12(b)(6) motion is

$$u_p(F, M) = - \left(c_f + c_{pm} \right) + \left(1 - \gamma_p \right) \{ v_{pa} + \beta \max [\sigma_a, 0] \}, \qquad (9)$$

which reflects the plaintiff's litigation costs—c_f plus c_{pm}—plus the product of the probability that the Rule 12(b)(6) motion will be denied in Stage 3, $1-\gamma_p$, and the plaintiff's payoff in that event, which is $v_{pa} + \beta \max [\sigma_a, 0]$. Plugging values from (6) and (9) into the inequality $-u_d(F,M) > u_p(F,M) + c_f$ and rearranging, we see that the defendant will prefer to settle if and only if

$$\gamma_p > -\frac{\sigma_m + \sigma_a}{v_{pa}} + \frac{c_{da}}{v_{pa}}\gamma_d, \tag{10}$$

where $\sigma_m \equiv c_{dm} + c_{pm}$ is the surplus achieved by avoiding the costs related to litigating the motion to dismiss. Since $c_{da} < v_{pa}$ by hypothesis, the slope of the line that supports the inequality in (10) in (γ_d, γ_p)-space must be less than one. Since the right hand side of (10) is increasing in γ_d and less than one when $\gamma_d = 1$, there will be some logically possible combinations of γ_d and γ_p for which the defendant will prefer to file a Rule 12(b)(6) motion rather than settle. Panel (a) of Figure 4 depicts the case in which v_{pa} is large enough so that the line defined by (10) cuts the vertical line given by $\gamma_d = c_{dm}/c_{da}$ at a positive value of γ_p.[65]

Figure 4: Optimal Defendant's Stage 2 Action in (γ_d, γ_p)-Space

(a) $c_{da} < v_{pa}$

(b) $c_{da} > v_{pa}$

---
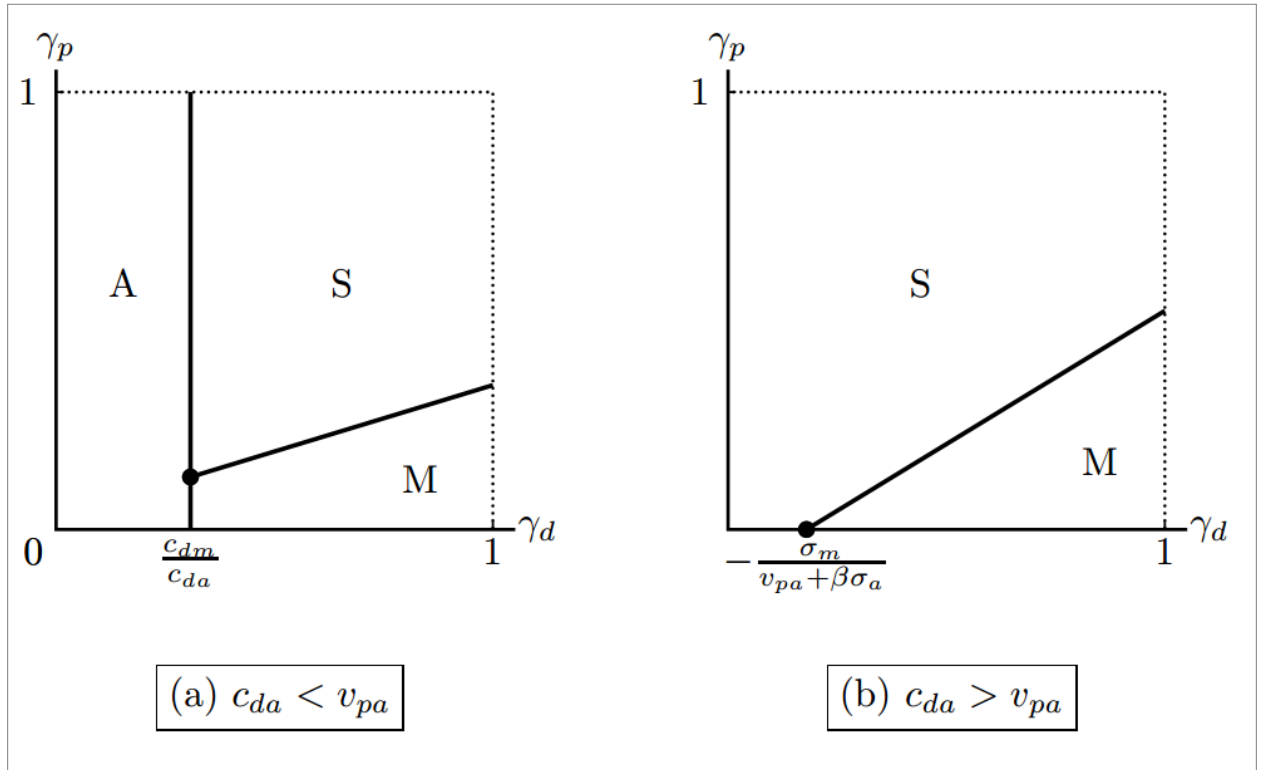
[65] When $\gamma_d = 1$, the right hand side of (10) equals $(v_{pa} - \sigma_m)/v_{pa}$, or $1 - (\sigma_m/v_{pa})$, which is less than one since σ_m and v_{pa} are each positive. The right hand side will also be positive at $\gamma_d = 1$ provided that $\sigma_m < v_{pa}$. Consequently, there exists some logically possible scenario in which the defendant would find it optimal to file a Rule 12(b)(6) motion in Stage 3 provided that the parties' combined costs of litigating the motion do not exceed the plaintiff's continuation value in the event the defendant files an answer (whether in Stage 2 or Stage 4). This condition is not particularly restrictive.

III.B.3.b. Optimal Stage 2 behavior when $\sigma_a > 0$

Now consider the situation when $c_{da} > v_{pa}$, so that $\sigma_a > 0$, and the parties would settle, rather than having the defendant answer, if they found themselves in Stage 4. The surplus from settling in Stage 2 is greater than the surplus in Stage 4, the defendant will not answer in Stage 2 in this situation. The only Stage 2 question when $\sigma_a > 0$ is whether the defendant would prefer to settle or file a Rule 12(b)(6) motion. Once again the defendant would prefer to settle if and only if $-u_d(F,M) > u_p(F,S) + c_f$. Plugging values from (6) and (9) into this inequality, accounting for the fact that $\sigma_a > 0$, and then rearranging shows that the defendant will prefer to settle if and only if[66]

$$\gamma_p > -\frac{\sigma_m}{v_{pa} + \beta\sigma_a} + \gamma_d. \tag{11}$$

Panel (b) of Figure 4 depicts the line defined by condition (11). When $\sigma_a > 0$, this line splits (γ_d, γ_p)-space into regions where settlement is the defendant's optimal Stage 2 action (above the line) and where filing a Rule 12(b)(6) motion is optimal (below the line); we have already seen that the defendant would never answer when σ_a is positive. For any $\gamma_d \in (0,1)$ there will alway exists a set of values of $\gamma_p \in (0,1)$ that satisfy (11), so there will always exist a set of parameter values that make settlement feasible when $\sigma_a > 0$.[67]

III.B.3.c. The Stage 2 settlement amount and associated payoffs

Since the defendant will never settle in Stage 2 when $\sigma < 0$ and γ_d lies to the left of ther vertical line in panel (a) of Figure 4, we can focus only on cases in which the defendant would file a motion to dismiss in Stage 2 if the parties didn't settle. The surplus from settling in Stage 2 is thus $-u_d(F,M|\sigma_a < 0) - [u_p(F,M|\sigma_a < 0) + c_f]$, if this quantity is positive. Plugging in the values in (6) and (9) yields a Stage 2 surplus that can be written

$$\sigma_2 \equiv \sigma_m + \begin{cases} (1 - \gamma_d)c_{da} - (1 - \gamma_p)v_{pa}, & \text{if } \sigma_a < 0, \\ (\gamma_p - \gamma_d)[v_{pa} + \beta\,\sigma_a], & \text{if } \sigma_a > 0. \end{cases} \tag{12}$$

Here, the σ_m term appears because the parties' costs of litigating the Rule 12(b)(6) motion are pure deadweight loss: no one will recoup them once they are spent, so avoiding them creates surplus. When $\sigma_a < 0$, we have seen that the parties would not settle in Stage 4 following the judge's denial of a Rule 12(b)(6) motion. Therefore, the defendant believes that with probability $1 - \gamma_d$, Stage 2 settlement will save her the cost c_{da}. Meanwhile, the plaintiff believes that with probability $1 - \gamma_p$, Stage 2 settlement will cost her the amount v_{pa}. Thus when $\sigma_a < 0$, Stage 2 surplus includes the term $[(1 - \gamma_d)c_{da} - (1 - \gamma_p)v_{pa}]$, whatever its sign.

---

[66] Here I have used the fact that $v_{pa} + \beta\sigma_a = c_{da} - (1 - \beta)\sigma_a$, which immediately implies that the coefficient on γ_d is one.

[67] The right hand side of (11) is always less than γ_d, so even when $\gamma_d = 1$, there exists $\gamma_p < 1$ that satisfies (11). Note in addition that the right hand side intersects the horizontal axis where $\gamma_d = \sigma_m/[v_{pa} + \beta\sigma_a]$. Provided that $\sigma_m < v_{pa} + \beta\sigma_a$, this will occur at a value of γ_d less than one, so that there will exist some space between the line defined by (11) and the point $(\gamma_d, \gamma_p) = (0,1)$. Thus the condition for filing a Rule 12(b)(6) motion in Stage 2 to be optimal when $\sigma_a > 0$ is $\sigma_m < v_{pa} + \beta\sigma_a$. This condition is more forgiving than the one that governs when $\sigma_a > 0$; see footnote 65.

When $\sigma_a > 0$, both parties will expect to settle in Stage 4 if the judge denies the defendant's Rule 12(b)(6) motion. That eliminates c_{da}-v_{pa} from the Stage 2 surplus. However, there is still some of what might be called a "belief surplus" related to the event that the judge grants the Rule 12(b)(6) motion. The defendant believes a grant will happen with probability γ_d, and the plaintiff believes the probability is γ_p. From the perspective of Stage 2, the defendant's expected Stage 4 settlement costs are $(1 - \gamma_d)[c_{da} - (1 - \beta)\,\sigma_a] = (1 - \gamma_d)[v_{pa} + \beta\,\sigma_a]$ (the equality follows from (1)). The plaintiff's expected Stage 4 settlement gains are $(1 - \gamma_p)[v_{pa} + \beta\,\sigma_a]$. Subtracting the plaintiff's expected gain from the defendant's expected costs yields surplus of $(\gamma_p - \gamma_d)[v_{pa} + \beta\,\sigma_a]$, which is positive when the defendant believes it less likely that the Rule 12(b)(6) motion will be granted than the plaintiff does, so that $1-\gamma_d>1-\gamma_p$. That would imply that the Stage 2 surplus from settling is greater than σ_m. If instead $\gamma_d>\gamma_p$, so that the defendant believes the probability of a Rule 12(b)(6) grant is greater than does the plaintiff, then the term $(\gamma_p - \gamma_d)[v_{pa} + \beta\,\sigma_a]$ will be negative. Thhat would reduce the surplus from settlement in Stage 2 to an amount less than σ_m.

In any Stage 2 settlement, the plaintiff will receive the share β of the surplus σ_2, plus her continuation payoff from history "F,M". The defendant will receive the share $(1-\beta)$ of σ_2, plus her continuation payoff from history "F,M". Thus, (12), (9), and (1) together imply that the plaintiff's continuation payoff in the event of Stage 2 settlement is

$$u_p(F,S) \quad \equiv \quad -(c_f + c_{pm}) + \left(1 - \gamma_p\right)\left(v_{pa} + \beta\max[\sigma_a, 0]\right) + \beta\sigma_2, \tag{13}$$

while (12), (6), and (1) imply that the defendant's payoff is

$$u_d(F,S) \quad \equiv \quad -c_{dm} - \left(1 - \gamma_d\right)\left(v_{pa} + \beta\max[\sigma_a, 0]\right) + (1 - \beta)\sigma_2. \tag{14}$$

When the surplus from Stage 2 settlement, σ_2, is positive, payoffs in filed cases will be given by $u_p(F,S)$ and $u_d(F,S)$. When σ_2 instead is negative, the relevant payoffs will be based on whether the defendant prefers to answer or to file a Rule 12(b)(6) motion, as discussed above. Working through some very tedious algebra yields the following complicated-looking formula for the defendant's expected payoff when the plaintiff has filed suit:

$$\begin{aligned}
u_d(F) \quad = \quad &-c_{da} + \max\{0, -c_{dm} + \gamma_d c_{da} \\
&+(1 - \gamma_d)(1 - \beta)\max[\sigma_a, 0] \\
&+(1 - \beta)\max[\sigma_2, 0]\}
\end{aligned} \tag{15}$$

Intuitively, the defendant never does worse than $-c_{da}$, since she can get that payoff by answering in Stage 2. She might do better, though, either by filing a Rule 12(b)(6) motion or by settling, which is why the second argument of the first max operator on the right hand side of (15) is there. That argument equals the potential gain to filing a motion to dismiss when the defendant would not settle if the motion were denied, $-c_{dm} + \gamma_d c_{da}$, plus two further terms that involve settlement surplus. The first additional term, $(1 - \gamma_d)(1 - \beta)\max[\sigma_a, 0]$, represents the defendant's share of the surplus from settling in Stage 4, multiplied by the probability that a Rule 12(b)(6) motion would be denied. The second settlement surplus term, $(1 - \beta)\max[\sigma_2, 0]$, represents the defendant's share, of the surplus from settling in Stage 2 rather than filing a Rule 12(b)(6) motion.

The plaintiff's expected payoff when she files suit may be shown to equal

$$
\begin{aligned}
u_p(F) \quad = \quad & -c_f + v_{pa} + (1 - \gamma_p)\beta \max[\sigma_a, 0] \\
& + \; 1(\sigma_a > 0)1(\gamma_d c_{da} - c_{dm} > 0) \times \{-c_{pm} - \gamma_p v_{pa} + \beta max \, [\sigma_2, 0]\}
\end{aligned}
\tag{16}
$$

Intuitively, the plaintiff always pays the filing cost c_f, and she gets the payoff v_{pa} if the defendant answers in Stage 2. If σ_a is positive, then there would be settlement surplus if the case got to Stage 4. In this case, if the defendant filed a Rule 12(b)(6) motion in Stage 2, the plaintiff would expect to receive an additional $\beta\sigma_a$ the fraction $(1-\gamma_p)$ of the timeIn addition, when either $\sigma_a > 0$ or $(\gamma_d c_{da} - c_{dm}) > 0$ holds, the defendant will not answer; that is, in either of these situations, either the defendant will file a Rule 12(b)(6) motion or the parties will settle in Stage 2. This fact is captured by the product of the indicator terms $1(\sigma_a > 0)$ and $1(\gamma_d c_{da} - c_{dm} > 0)$. When the defendant files a Rule 12(b)(6) motion, the plaintiff bears litigation cost c_{pm} and, since the plaintiff expects to lose on the motion the share γ_p of the time, she expects to lose v_{pa} with probability γ_p; this explains the first two terms in curly braces on the second line of (16). Finally, the term $\beta \max[\sigma_2, 0]$ represents the plaintiff's portion of any additional surplus available from settling in Stage 2.

III.B.3.d. Implications for the observed Rule 12(b)(6) motion filing and grant rates

I now discuss two important empirical implications of the foregoing Stage 2 analysis.

Figure 4 shows that a defendant's optimal Stage 2 action can be found by determining where in the five basic subsets of $(c_{da}, v_{pa}, \gamma_d, \gamma_p)$-space the parameters of a given case place it. It is logically possible for all cases to have parameter values that fit in the bottom-right part of either panel of Figure 4. In that case, defendants would file Rule 12(b)(6) motions in 100 percent of cases that plaintiffs filed. We also cannot rule out the possibility that defendants will never file Rule 12(b)(6) motions, because it is also logically possible that no filed cases lie in the bottom-right regions of the graphs in Figure 4. Finally, it is logically possible to observe Rule 12(b)(6) motions being filed in any share between zero and 100 percent of cases, since some cases might lie in one of bottom-right regions of Figure 4 while others might not. Consequently, we have the following result.[68]

---

[68] It might seem premature to state this proposition here, since I have not yet analyzed the plaintiff's Stage 1 filing decision. Perhaps, one might think, something about plaintiffs' filing decisions would systematically eliminate cases that would lie in one or another region of the graphs in Figure 4. In section III.B.4 below, I show that if there is no judicial heterogeneity, then the plaintiff's Stage 1 decision can be characterized using Figure 4. The only modifications needed are those necessary to account for the fact that the plaintiff's filing cost c_f is non-sunk in Stage 1. Accounting for this additional litigation cost shifts the three lines in panel (a) of Figure 4 to the right. Given the discussion in footnotes 65 and 67, there will exist values of the parameters other than γ_d and γ_p such that the bottom-right regions of this modified version Figure 4 have positive area. Thus, there will be a non-empty (and non-knife edge) set of parameter values such that plaintiffs will file suit in Stage 1 knowing that defendants will file Rule 12(b)(6) motions in Stage 2. Since this result holds with no judicial heterogeneity, continuity of all conditions in the game implies that there exists some positive amount of judicial heterogeneity such that the result also will hold.

Proposition 1: Among cases that plaintiffs file, defendants might file Rule 12(b)(6) motions in zero, 100%, or any percentage in between.

What fraction of filed cases have Rule 12(b)(6) motions filed is thus an empirical question. Its answer will depend not just on the fact of case selection, but also on the distribution of across the parties of beliefs concerning the probability a Rule 12(b)(6) motion would be granted, as well as the model's various cost parameters.

Next, consider the grant rate among those cases that have Rule 12(b)(6) motions filed. In these cases, defendants believe they are relatively likely to win a dismissal grant on the motion, while plaintiffs believe defendants are relatively likely to lose. But to say anything about the *actual* share of these cases in which the judge grants dismissal requires one to assume some structure on the relationship between the parties' beliefs and the actual grant probability, γ, that characterizes the judge's behavior in any case. Assume, for instance, that the probability that a Rule 12(b)(6) motion would be granted, given γ_d and γ_p, can be written as:

$$\gamma \quad \equiv \quad P(\text{Judge actually grants Rule 12(b)(6)motion}|\gamma_d, \gamma_p)$$
$$= \quad \alpha\gamma_d + (1 - \alpha)\gamma_p.$$

With this structure, the actual grant probability in a case is a weighted average of the parties' beliefs, with weight α placed on the defendant's belief. Suppose that $\alpha=1$, and that all filed cases have $\gamma_d=1$, so that the defendant's belief (i) is always correct and (ii) is that a Rule 12(b)(6) motion would be granted with certainty. Now suppose instead that $\alpha=0$, and that all filed cases have $\gamma_p=0$, so that the plaintiff is always correct and always believes the Rule 12(b)(6) motion will be denied with certainty. In these polar situations, 100 percent and zero percent of Rule 12(b)(6) motions will be granted. And nothing rules either situation out as a logical matter. Finally, for values of α strictly between zero and one, we will have grant rates greater than zero but less than 100 percent. We thus have the following result.

Proposition 2: Among cases that actually have Rule 12(b)(6) motions filed, the grant rate can be zero, 100%, or any percentage in between.

Propositions 1 and 2 teach an important lesson: the presence of party selection is insufficient to yield any empirical restrictions on the observed values of either the Rule 12(b)(6) motion filing or grant rate. These propositions are inconsistent with Hubbard's (2011) conjecture that selection effects will necessarily force the Rule 12(b)(6) motion filing rate to be low, and also with his conjecture that selection effects will necessarily force the Rule 12(b)(6) motion grant rate to be close to 50 percent. [69]

---

[69] *See* Hubbard, *supra* note 25, at 17-18. It is not clear to me how plaintiffs' filing decisions in Hubbard's model are linked to defendants' potential future Rule 12(b)(6) motion filing decisions, because he does not analyze the parties' behavior sequentially. Thus, I am not sure whether his model allows dynamically inconsistent behavior; as an example, I am not sure whether Hubbard's model rules out the possibility that a plaintiff files suit under the assumption that the defendant would not file a Rule 12(b)(6) motion, even though the plaintiff can see that the defendant will indeed file such a motion.

III.B.4. Optimal plaintiff's behavior in Stage 1

In Stage 1, the plaintiff must decide whether to drop the case, settle for an amount the defendant is willing to pay, or file suit. Because the judge's identity is random, and because I allow the possibility of judge heterogeneity, the parties in Stage 1 do not know the exact payoff values in the event that the plaintiff files the case. Thus, the payoff functions for Stage 1are expected values taken with respect to each party's subjective probability distribution over the future actions the parties believe each judge would take.

Let $u_p^j(h)$ and $u_d^j(h)$ be the parties' payoffs conditional on game history h and on Nature's selection of judge j in Stage 1′. These functions may vary with j because, as discussed above, the parties may believe different judges handle procedural and substantive aspects of the case differently. This means that the plaintiff will think γ_p differs across judges, the defendant will think γ_d does, and so on. Thus, with the exception of β, all the parameters introduced above—γ_d, γ_p, c_{da}, v_{pa}, c_{dm}, c_{pm}—might need a "j" index. However, it will be enough to define

$$\overline{u_p}(F) = \frac{1}{J}\sum_{j=1}^{J} u_p^j(F) \tag{17}$$

and

$$\overline{u_d}(F) = \frac{1}{J}\sum_{j=1}^{J} u_d^j(F), \tag{18}$$

where $u_d^j(F)$ and $u_p^j(F)$ are given by inserting "j" subscripts on all parameters of (15) and (16) except for β. Since I normalized to zero the payoff from dropping the case, we have $u_p^j(F, Drop) = u_d^j(F, Drop) = 0$ for all judges, so that $\overline{u_d}(Drop) = \overline{u_d}(Drop) = 0$. It follows that the plaintiff will not drop the suit if and only if her expected payoff from filing, $\overline{u_p}$(F), is positive.

As usual, the parties will settle if there is surplus from doing so.[70] There will be surplus when the defendant's expected cost given that the plaintiff files, -u_d(F), is less than the plaintiff's expected gain. We have the following conditions for each plaintiff action to be optimal in Stage 1:

---

[70] Note that the defendant will not agree to settle with a plaintiff whose expected payoff from filing suit, $\overline{u_p}$(F), is negative, because such a plaintiff's threat to file in the absence of settlement is not credible. Because the backward-induction solution concept ensures that the players behave optimally along all subgames, the plaintiff would only make herself worse off by sinking the filing cost. This result is actually consistent with those in papers like Bebchuk (1996) and Grundfest and Huang (2006), even though those papers show that what they define as negative expected value (NEV) suits can be worth filing. Those authors define suits to be NEV when the plaintiff's expected value is negative *conditional on pursuing the case to final judgment*. But parties in my model will pursue the case to the answering stage only when it is subjectively rational to do so along the equilibrium path. Thus, a negative value of $\overline{u_p}$(F) indicates not simply that the plaintiff's expected value of pursuing the case as far as the court will allow is negative, but rather that there is a negative expected value of filing the case and *then pursuing the best possible course of action in the future*, given that the defendant will do likewise. Thus the apparent difference with Bebchuk's and with Grundfest and Huang's models is just terminological, in that their definitions of NEV refer to behavior that ventures off the subgame-perfect equilibrium path.

Plaintiff drops case:	$\overline{u_p}(F) < 0$
Plaintiff files suit:	$\overline{u_p}(F) > |\overline{u_d}(F)| > 0$ (19)
Parties settle:	$|\overline{u_d}(F)| > \overline{u_p}(F) > 0$

I illustrate these conditions in Figure 5. When the payoffs are located in the lower half of the figure, the plaintiff's expected gain is negative, so she drops the case. Above the 45-degree line in the top half of the figure, the plaintiff files suit, because her expected payoff from doing so exceeds the defendant's expected costs in that event. The parties settle in the space in between.

Figure 5: Optimal Stage 1 Plaintiff Actions in $(\overline{u_d}(F), \overline{u_p}(F))$-Space
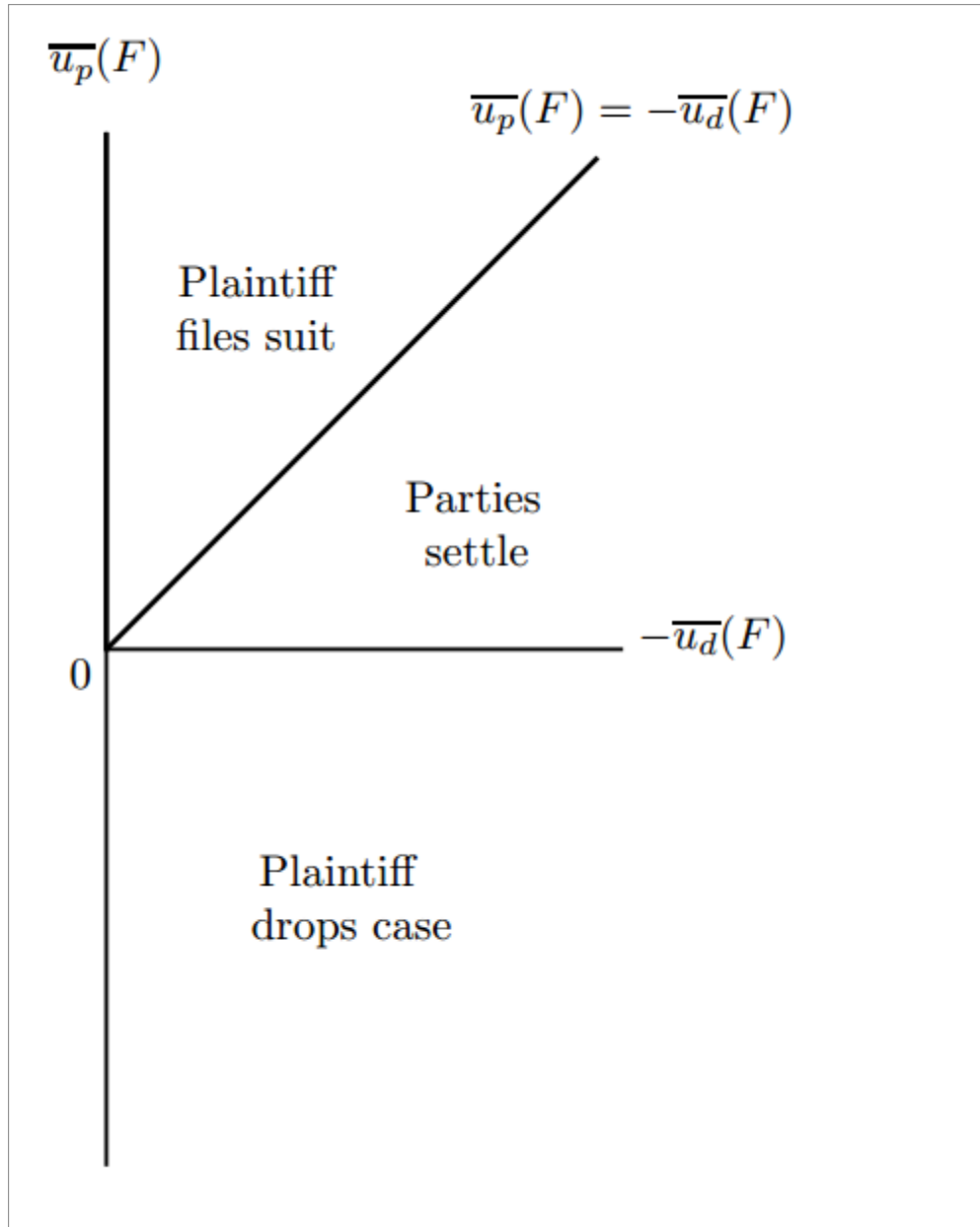
IV. An Evaluation of Rule 12(b)(6) and Judicial Administration Costs

With both plaintiffs' filing behavior and defendants' Rule 12(b)(6) motion filing behavior now analyzed, I can state some basic facts concerning the effects of Rule 12(b)(6) on judicial administration and on the parties' welfare. To evaluate Rule 12(b)(6)'s effects, I will compare equilibria of the litigation game as discussed in Part III to the equilibrium that would occur if

$\gamma_d^j = \gamma_p^j = 0$ for every judge in every case, since that situation is equivalent to the elimination of Rule 12(b)(6). For convenience, I will assume in this Part that there is no judicial heterogeneity in the game's various parameters; this assumption is inessential to the logic of the discussion here, but it allows me to avoid using judge indexes.

The discussion in Part III implies that when $\gamma_d = \gamma_p = 0$, the necessary and sufficient condition for cases to settle is $c_{da} < v_{pa}$, or, equivalently, $\sigma_a > 0$. Thus, cases would never settle when the parameters' values are located in panel (a) of Figure 4. Cases would always settle in Stage 2, at the latest, when the parameter values are located in panel (b). By comparison, I showed in Part III that when Rule 12(b)(6) exists, some cases in panel (a) of Figure 4 will settle, while some cases in panel (b) will *not* settle. We thus have the following result.

> **Proposition 3:** The existence of Rule 12(b)(6) might either increase or reduce the number of cases that settle in Stage 2.

There are some interesting intuitive facts behind this proposition. First, Rule 12(b)(6) gives defendants a credible threat when the parties' beliefs diverge too much to support the existence of settlement surplus in the absence of the Rule. By (i) reducing plaintiffs' expected gains and (ii) increasing both parties' expected litigation costs, the Rule transforms some cases from negative- to positive-surplus cases, along the equilibrium path. On the other hand, Rule 12(b)(6) also can work in the opposite direction, because the defendant's option to file a Rule 12(b)(6) motion creates an additional issue on which the parties' beliefs might diverge. The more the parties disagree about the likelihood a Rule 12(b)(6) motion would be granted, the greater the range of cases that will be caused *not* to settle as a result of Rule 12(b)(6).

In addition, there are judicial administration costs from litigating the Rule 12(b)(6) motions themselves. In sum, the existence of Rule 12(b)(6) might either increase or reduce the costs of judicial administration. Which result occurs depends on the relative numbers of cases that the Rule causes to settle rather than being litigated, and vice-versa.

V. Effects of Switching to More Demanding Pleading

In this Part, I consider the effects of switching pleading regimes. By "pleading regime," I mean a set of rules for assessing whether a complaint meets Rule 8(a)'s requirements. Each pleading regime generates a particular distribution of (γ_d^j, γ_p^j) parameters across cases. In a given pleading regime, the defendant in each case has a fixed set of beliefs $\{\gamma_d^j\}_{j=1}^J$, and the plaintiff has a fixed set $\{\gamma_p^j\}_{j=1}^J$.

I model a switch to more demanding pleading as a change in legal rules that satisfies two properties. First, it causes some parties to increase the subjective probability that they put on the event that a Rule 12(b)(6) motion would be granted if filed. Second, it does not cause any parties to reduce the probability they put on this event. In some cases—especially those that are difficult to plead due, say, to the plaintiff's need for discovery to make factual allegations—both the defendant and the plaintiff might believe that the switch in pleading regimes increases the probability of a Rule 12(b)(6) grant, should a motion actually be filed. In other cases—especially those that are easy to plead, say, because the plaintiff already has substantial information about the controversy—one or both party might have the same belief, concerning the probability of a Rule 12(b)(6) grant, before and after the switch. Letting subscript "TI" denote the *Twombly/Iqbal*

pleading regime, and letting subscript "C" denote the *Conley* regime, a switch to more demanding pleading occurs when $\gamma_{dTI}^{j} \geq \gamma_{dC}^{j}$ and $\gamma_{pTI}^{j} \geq \gamma_{pC}^{j}$ for all judges j, with strict inequality for at least one party-judge combination.

V.A. *Effects of Switching to More demanding pleading on the Defendant's Stage 2 Decision*

In Figure 6, I modify the graphs in Figure 4 to illustrate how switching to more demanding pleading affects the defendant's optimal Stage 2 action. Consider panel (a) first. Each dot in this graph corresponds to a pair (γ_d, γ_p) that logically could represent the parties' subjective beliefs concerning Rule 12(b)(6) motion grant rates as of Stage 2, in a given pleading regime, when $c_{da} < v_{pa}$ (equivalently, $\sigma_a > 0$). I represent the switch to more demanding pleading with arrows from one dot to another.[71]

Figure 6: How A Switch to More Demanding Pleading Affects the Defendant's Stage 2 Choice
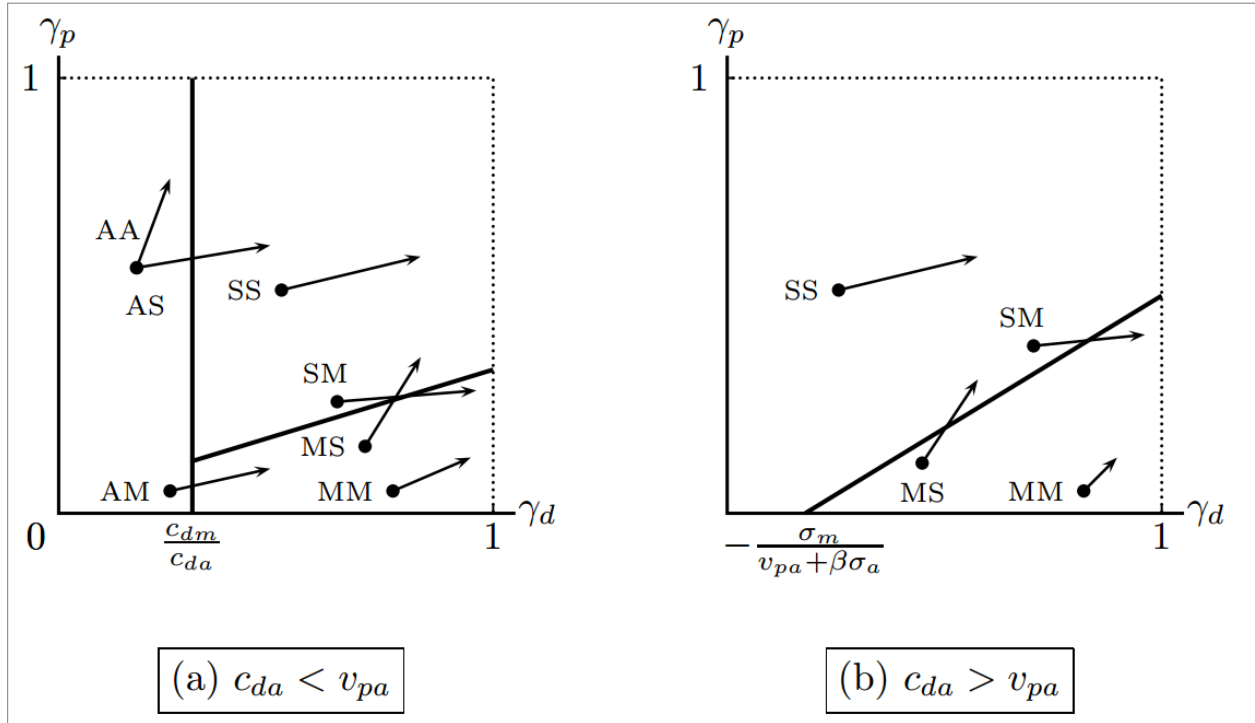
Consider the arrow labeled "AA." This arrow emanates from a dot inside the region of panel (a) where the defendant would answer a filed complaint, and it lands on a dot that is also in this region. Thus, this pair of dots represents a case whose defendant would choose to answer in Stage 2 regardless of the pleading regime. Similarly, the arrows labeled SS and MM represent

---

[71] For clarity, I drew Figure 6 so that all moves following the switch to more demanding pleading are to the northeast; recall, though, that moves due north or due east would also be consistent with a switch to more demanding pleading.

cases that would be settled in Stage 2 or would face Rule 12(b)(6) motions under both pleading regimes. The SS and MM case types also appear in panel (b) of the figure (the AA type does not, since we have seen that it is never optimal to answer in Stage 2 when $c_{da} > v_{pa}$).

Consider next the other arrow, labeled "AS", emanating from the dot spawning the AA arrow in panel (a). The defendant here would answer under *Conley*, but the parties would settle under more demanding pleading. Type AM cases also get an answer under *Conley*, but they face Rule 12(b)(6) motions under more demanding pleading. The structure of panel (a) makes it obvious that there can be neither MA nor SA cases: Since γ_d cannot fall as a result of the switch to more demanding pleading, the switch will never cause a leftward move in panel (a). As such, no cases ever flow into the set that will be answered as a result of the switch to more demanding pleading.

Both panels of Figure 6 show that there can also be MS cases—those that would face a motion to dismiss under *Conley* but be settled under more demanding pleading, were the cases to be filed in Stage 1. Finally, SM cases will exist if the switch to more demanding pleading makes the defendant sufficiently more optimistic about her chances of having her Rule 12(b)(6) motion granted, by comparison to the change in the plaintiff's beliefs. In that event, the switch to more demanding pleading eliminates all perceived surplus from settlement. Recall that the lines separating the regions in which the defendant would optimally settle from the regions in which she would optimally file a Rule 12(b)(6) motion—defined by (10) and (11)—have slope less than or equal to one. Therefore, if the switch to more demanding pleading causes γ_p to increase more than γ_d does, a case cannot be a type-SM case.[72] The following proposition collects the foregoing results.

> **Proposition 4**: The following statements are true of any case that reaches Stage 2 under both *Conley* and more demanding pleading:
>
> 2. There might be AA, SS, and MM cases—those whose Stage 2 outcomes are unaffected by switching to more demanding pleading.
>
> 3. Defendants might file Rule 12(b)(6) motions under more demanding pleading in some cases that they would answer under *Conley*: there might be AM cases.
>
> 4. The parties might settle some cases under more demanding pleading that defendants would answer under *Conley*: there might be AS cases. The parties might settle some cases under more demanding pleading in which defendants would file Rule 12(b)(6) motions under *Conley*: there might be MS cases.
>
> 5. Cases that are not answered under *Conley* are also not answered under more demanding pleading: there are no SA or MA cases.

---

[72] I call this condition the "differentially more pessimistic plaintiff condition", since it implies that the plaintiff is more pessimistic about the effects of switching to more demanding pleading than the defendant is optimistic. Note that this condition is only sufficient, and not necessary, to rule out type-SM cases. For example, suppose switching to more demanding pleading causes belief changes that violate the differentially more pessimistic plaintiff condition, but suppose that the magnitude by which γ_d increases is relatively small. Then if the case would be settled before the switch to more demanding pleading, it will still be settled after the switch, since γ_d and γ_p will stay above the line defined by whichever of (10) or (11) applies.

6. If for all cases γ_p increases more than γ_d does, then there will be no SM cases.

Next I consider the effects of switching to more demanding pleading on Stage 1 of the litigation game discussed in Part III.

V.B. *Effects of Switching to More demanding pleading on the Plaintiff's Stage 1 Behavior*

I now turn to the effects of switching to more demanding pleading on the plaintiff's Stage 1 choice—whether to drop the case, file suit, or settle for an amount the defendant would be willing to pay. To understand these effects, the following proposition will be helpful.

Proposition 5:

1. The plaintiff's expected payoff from filing suit when nature subsequently schooses judge j, $u_p^j(F)$, is unaffected by changes in γ_d^j and γ_p^j when the defendant answers in Stage 2. This payoff is strictly decreasing in these parameters when the defendant does not answer in Stage 2. The defendant's expected payoff when the plaintiff files suit, $u_d^j(F)$, is also unaffected when the defendant would answer in Stage 2, and it is strictly increasing in γ_d^j and γ_p^j otherwise.

2. Therefore, the parties' expected Stage 1 payoffs from filing suit, $\overline{u_p}(F)$ and $\overline{u_d}(F)$, are unaffected by changes in γ_d^j and γ_p^j when the defendant would answer in Stage 2 for every judge assignment. When there exists at least one judge j for which the defendant would not answer in Stage 2, the plaintiff's expected Stage 1 payoff is strictly decreasing, and the defendant's is strictly increasing, in γ_d^j and γ_p^j.

The proof of the Proposition's first part involves straightforward but tedious calculations, so I relegate it to Appendix VIII. The second part of the Proposition follows immediately from the first part, together with the fact that $\overline{u_p}(F)$ and $\overline{u_d}(F)$ are monotonic transformations of the set of $u_p^j(F)$ and $u_d^j(F)$ values. Proposition 5 tells us that switching to more demanding pleading either doesn't affect where the parties' payoffs are located in Figure 5's $\left(-\overline{u_d}(F), \overline{u_p}(F)\right)$-space or causes moves down, to the left, or both. Figure 7 demonstrates such possibilities visually using arrows to illustrate the pairs of Stage 1 outcomes that are possible under *Conley* and under more demanding pleading.

29

Figure 7: Effect of Pleading Regime on Plaintiff's Stage 1 Action
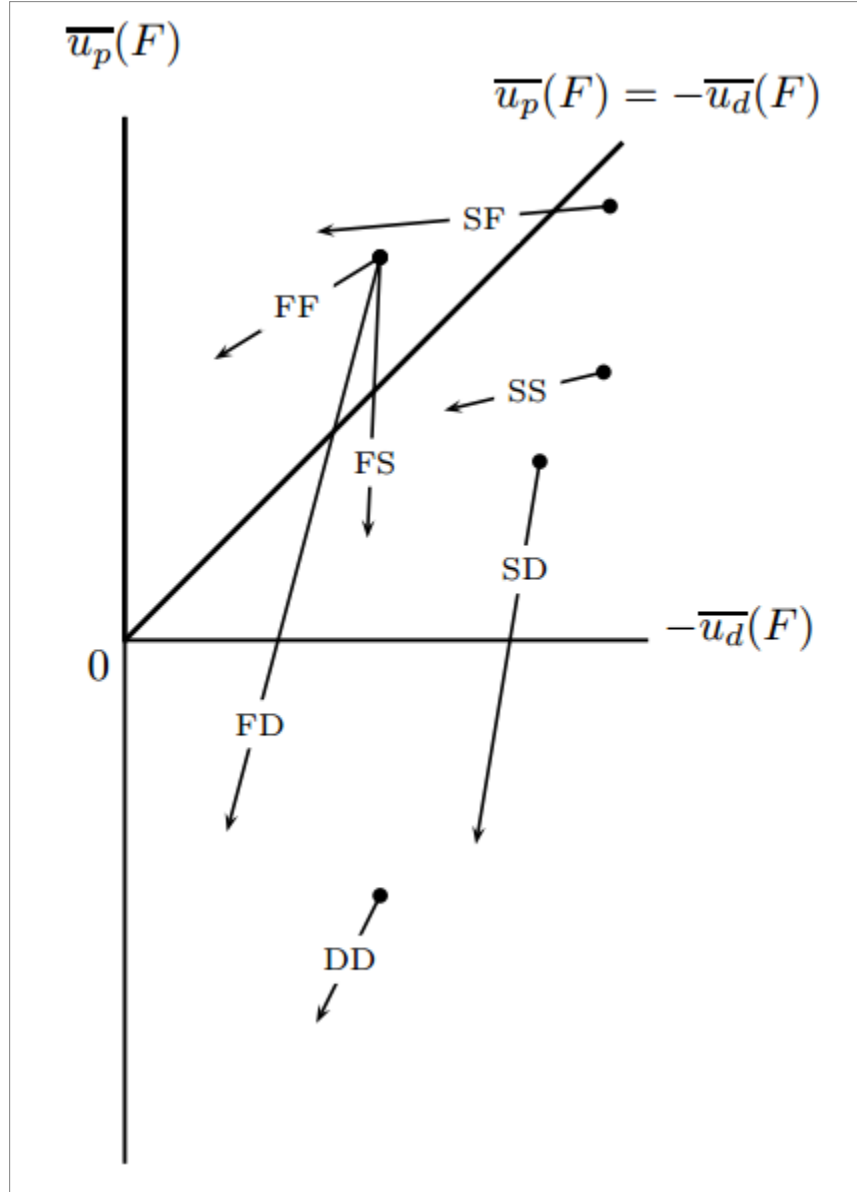
The main points to take away from the figure are that any pair of outcomes is possible for cases that would be filed or settled under *Conley*, and that any case dropped under *Conley* must also be dropped under more demanding pleading. Though, as with SM cases in Stage 2, one can rule out SF cases in Stage 1 with the assumption that γ_p increases more than γ_d as a result of the switch to more demanding pleading.[73] I sum up these results in the following proposition.

---

[73] Recall that under the condition that γ_p^j increases more than γ_d^j for each judge j, any case settled in Stage 2 before a switch to more demanding pleading would also settle in Stage 2, assuming it got to that stage, after the switch. Therefore, the plaintiff's expected Stage 2 payoff under judge j, $u_p^j(F)$, must fall at least as much after a switch to

Proposition 6

1. *Any case dropped in Stage 1 under* Conley *will be dropped in Stage 1 under more demanding pleading: there are no DF or DS cases.*

2. *There may be FF, SS, DD, FS, FD, SD, and SF cases.*

3. *If for all cases γ_p increases more than γ_d does, then there will be no SF cases.*

V.C. *An "Unidentification" Result on Judicial Behavior Effects*

As I discussed in section II.A, there has been considerable debate over empirical work measuring changes in the observed Rule 12(b)(6) grant rate following *Twombly* and *Iqbal*. This literature conveys the impression that if the observed grant rate were to have increased, then there must have been what I have called judicial behavior effects—changes in the way judges adjudicate a given set of cases. Conversely, a reader of the literature might think, an absence of any change in the observed grant rate would suggest no change in judicial behavior. In this section I show that when the parties respond to perceived changes in pleading policy, such conclusions are untenable.

To economize on notation, it will help to group cases with different game histories into common types. I do this in Table 1, each row of which concerns one possible game history. For example, the first row concerns the history "Drop", representing a case that the plaintiff drops in Stage 1; the last row concerns history "F,M,G", representing a case that the plaintiff files in Stage 1, the defendant challenges via Rule 12(b)(6) motion in Stage 2, and the judge terminates by granting the motion in Stage 3. The table's second column characterizes each possible history in terms of its status through the end of Stage 3, using "D" to characterize cases that are dropped, "S" to represent those that are settled, "A" to represent those that are answered in Stage 2, "M_D" to represent those that have Rule 12(b)(6) motions filed in Stage 2 and *denied* in Stage 3, and "M_G" to represent those that have Rule 12(b)(6) motions filed in Stage 2 and *granted* in Stage 3.

---

more demanding pleading as the defendant's expected costs, $|u_d^j(F)|$, fall. Thus, $\left|u_d^j(F)\right| - u_p^j(F)$ can never fall after a switch to more demanding pleading. Summing over judges and multiplying by J^{-1} shows that $|\overline{u_d}(F)| - \overline{u_p}(F)$ can never fall, either. If the case would have settled in Stage 1 under *Conley*, then this difference was originally positive. Therefore if the difference in Stage 1 payoffs does not fall, then it cannot be negative under more demanding pleading. So if γ_p^j rises more than does γ_d^j, then cases that are settled under *Conley* would never be filed under more demanding pleading. That rules out the existence of SF cases under the differentially more pessimistic plaintiff condition.

Table 1: Notation Relating Game Histories to
Observable Case Facts Through Stage 3

Litigation Game History	Type Through Stage 3
Drop	D
S	S
F,S	S
F,A	A
F,M,Deny	M_D
F,M,G	M_G

In Table 2, I use this case characterization to catalogue the different pairs of case history that might occur under *Conley* and under more demanding pleading. The rows of this table list the possible game histories under *Conley*, while the columns list the possible histories under more demanding pleading. Each cell of the table indicates the pair of abbreviations from Table 1. For example, a case that would be dropped under both pleading standards is a DD case, while one that would be settled under *Conley* but dropped under more demanding pleading is an SD case. This table has 25 cells, because the taxonomy in Table 1 allows five possible case types.

Table 2: Potential Outcome Pairs in Terms of Observable Case Facts

		Game History Under More demanding pleading				
		(Drop)	(S)/(F,S)	(F,A)	(F,M,Deny)	(F,M,Grant)
Game History Under *Conley* Pleading Standard	(Drop)	DD	DS	DA	DM_D	DM_G
	(S)/(F,S)	SD	SS	SA	SM_D	SM_G
	(F,A)	AD	AS	AA	AM_D	AM_G
	(F,M,Deny)	M_DD	M_DS	M_DA	M_DM_D	M_DM_G
	(F,M,Grant)	M_GD	M_GS	M_GA	M_GM_D	M_GM_G

The analysis in Part IV implies that not all 25 of these case types can occur when my model accurately characterizes behavior. First, there cannot be any M_GM_D cases. Such cases would have Rule 12(b)(6) motions filed and granted under *Conley*, but filed and denied under more demanding pleading. For these cases, then, switching pleading regimes would have to *reduce* the pleading standard; I am unaware of a single commentator who has suggested that *Twombly* and *Iqbal* would have such an effect on any case.

Of the remaining 24 cells in Table 2, there are six others whose case types cannot occur if the plaintiff and defendant behave rationally. As part 1 of Proposition 6 shows, the plaintiff will always drop a case under more demanding pleading that she would drop under *Conley*. That rules out the possibility that there are any DS, DA, DM_D, or DM_G cases. Moreover, part 4 of Proposition 4 rules out the existence of any M_DA or M_GA cases.[74] All told, then we can rule out

---

[74] It also rules out the existence of any cases that follow history "F,S" under *Conley* and history "F,A" under more demanding pleading. But it is possible that there are cases with history "S" under *Conley* and history "F,A" under more demanding pleading, so we cannot rule out the existence of some SA cases. I claimed otherwise in Gelbach, *Locking the Doors to Discovery*, *supra* note 11, at 2308. This error is a minor one in the context of that paper,

the existence of any cases of type M_GM_D, DS, DA, DM_D, DM_G, M_DA, or M_GA. I illustrate this fact using Table 3, which modifies Table 2 by blacking out the cells corresponding to these six case types.

Table 3: Potential Outcome Pairs Consistent with
Imposing Economic Rationality in the Litigation Game

		Game History Under More demanding pleading				
		(Drop)	(S)/(F,S)	(F,A)	(F,M,Deny)	(F,M,Grant)
Game History Under *Conley* Pleading Standard	(Drop)	DD	███	███	███	███
	(S)/(F,S)	SD	SS	SA	SM_D	SM_G
	(F,A)	AD	AS	AA	AM_D	AM_G
	(F,M,Deny)	M_DD	M_DS	███	M_DM_D	M_DM_G
	(F,M,Grant)	M_GD	M_GS	███	███	M_GM_G

Now suppose momentarily that we could identify those cases that would have Rule 12(b)(6) motions filed under either pleading standard—those cases having type M_DM_D, M_GM_G, and M_DM_G. By construction, the first two of these case types are unaffected by the switch, since Rule 12(b)(6) motions in these cases would either be denied under both pleading standards or granted under both pleading standards. By contrast, M_DM_G cases involve some "treatment effect": by construction, Rule 12(b)(6) motions in these cases would be denied under *Conley* and granted under more demanding pleading. It is thus meaningful to think of M_DM_G cases as involving what I have elsewhere called a *judicial behavior effect*.[75]

As a share of cases that have Rule 12(b)(6) motions filed under both pleading standards, this judicial behavior effect equals the number of M_DM_G cases divided by the total number of M_DM_D, M_GM_G, and M_DM_G cases. This ratio equals the increase in the share of MM cases that would have Rule 12(b)(6) motions granted under more demanding pleading, divided by the share that would have grants under *Conley*. If we could separately identify these three types of cases, then we could measure the judicial behavior effect among all MM cases. In fact, for the subset of MM cases, the Rule 12(b)(6) motion grant rate under *Conley* minus the grant rate under more demanding pleading equals the relative judicial behavior effect just defined. This fact would justify the use of differences in observed grant rates to measure the judicial behavior effects of switching to more demanding pleading, were there no party selection. Unfortunately, the discussion above suggests that party selection is likely to be endemic. Consequently, before-and-after comparisons generally will conflate judicial behavior effects among MM cases with compositional changes in the set of cases that are observed facing Rule 12(b)(6) motions.

Table 3 casts some light on the scope of the party selection problem. Even when we do not distinguish between settlement in Stage 1 and Stage 2, there are 18 different case-type cells

though, since such cases play no role in any of the analysis or calculations relad to the empirical work I conducted there *Locking the Doors to Discovery*.

[75] *See* Gelbach, *Locking the Doors to Discovery*, *supra* note 11, section II.B, at 2298 for a discussion of judicial behavior effects.

to consider when we evaluate the effects of switching to more demanding pleading. Of these 18 case types, 11 involve cases that would have a Rule 12(b)(6) motion filed under at least one pleading regime, and the $M_D M_D$, $M_G M_G$, and $M_D M_G$ case types account for only three of these 11. Unless one is willing to assume away party selection, then, comparing the pre-*Twombly* and post-*Twombly* or -*Iqbal* grant rates will unavoidably conflate judicial behavior effects on MM cases and party selection involving other case types.

Perhaps one might think grant rate comparisons could still be useful. For example, perhaps the nature of party selection is such that the observed grant rate must rise if the judicial behavior effect defined above is positive. In my earlier paper, I showed that positive judicial behavior effects can exist even when the observed grant rate falls or stays the same.[76] The following Proposition generalizes this result substantially, showing that even information on the direction of change in three variables—the Rule 12(b)(6) motion filing rate, the Rule 12(b)(6) motion grant rate, and the number of cases filed—is insufficient to reveal anything about the existence of judicial behavior effects (see Appendix VIII.B for proof).

Proposition 7: As a result of a switch to more demanding pleading:

1. It is possible to observe any combination of increases, unchanged values, or decreases in the number of cases filed, the Rule 12(b)(6) motion filing rate among cases that are filed, and the Rule 12(b)(6) motion grant rate among the set of cases that have Rule 12(b)(6) motions filed.

2. Point 1 holds even if the parties' belief that judicial behavior changes is mistaken.

Proposition 7 highlights the difficulty of drawing useful conclusions about the impact of switching to more demanding pleading by observing "only" the direction of change in the Rule 12(b)(6) motion grant rate, the Rule 12(b)(6) motion filing rate, and the number of cases filed. Among other things, Proposition 7 shows that the presence of party selection is fully consistent with a failure to observe a drop in the number of cases that plaintiffs file.[77] More generally, Proposition 7 shows that by itself, no combination of observed changes in the Rule 12(b)(6) motion filing rate grant, Rule 12(b)(6) motion grant rate, or number of cases filed can refute the proposition that *Twombly* and *Iqbal* have affected judicial behavior when one allows parties to alter their litigation behavior. Nor can any observed pattern of changes *establish* that fact, as part 2 of the Proposition establishes. If we want to learn something about the effects of switching to more demanding pleading, we must find another way to do so.

---

[76] *See* discussion surrounding Table 1 of *id.*, at 2312.

[77] This fact contradicts a suggestion, in Joe Cecil's response to my earlier work, that "one would expect plaintiffs to file fewer cases after *Twombly* and *Iqbal* if they perceive the pleading standards to be more demanding." Cecil, *Of Waves and Water*, *supra* note 37, at 42. While Cecil does qualify this statement with the prefatory phrase "[a]ll other things being equal," the other things thhat have to be equal are simply those factors that invalidate Cecil's substantive claim—case filings involving party selection. Thus, Cecil's claim is correct only if he assumes away all the reasons we shouldn't expect it to be.

V.D. *Identifying Bounds on Twombly and Iqbal's Negative Effects on Certain Plaintiffs*

Despite the apparently draconian nature of Proposition 7, I showed in my earlier paper that it is still possible to learn something important concerning the impact of switching to more demanding pleading. My focus there concerned the concept of negatively affected plaintiffs. A plaintiff is negatively affected by a switch to more demanding pleading if the plaintiff's realized payoff is reduced as a consequence of the switch. Such a payoff reduction could occur because of a change in judicial behavior, or because of the changes in the parties' litigation choices I discuss in Part III. The full set of cases with negatively affected plaintiffs is too broad to be observably measured,[78] but we can learn something about negatively affected plaintiffs among those cases that would have Rule 12(b)(6) motions filed under the more demanding pleading regime.[79] Among these cases, the set of negatively affected cases includes:

- SM_G cases: under *Conley* the plaintiff would receive a settlement payment without having to litigate a Rule 12(b)(6) motion, whereas she has to litigate the motion and then receives no payment under more demanding pleading;

- AM_G cases: for similar reasons to SM_G cases (except that under *Conley* the plaintiff gets the continuation value associated with the defendant's answer, rather than a settlement payment);

- $M_D M_G$ cases: the defendant's Rule 12(b)(6) motion would be denied under *Conley* but granted under more demanding pleading, terminating the claim;

- AM_D cases: under *Conley* the plaintiff gets the continuation value associated with an answer in Stage 2, but under more demanding pleading she must litigate a Rule 12(b)(6) motion to get this value in Stage 4;

- and possibly SM_D cases.[80]

I show in my earlier paper that the number of cases with Rule 12(b)(6) motions granted under more demanding pleading minus the corresponding number under *Conley* is a lower bound on the total number of SM_G, AM_G, and $M_D M_G$ cases.[81] It follows that this difference is also a lower bound on the total number of negatively affected plaintiffs among all cases that would have Rule 12(b)(6) motions filed under *Twombly* and *Iqbal*, since the latter category includes AM_D and, possibly, SM_D cases as well.[82] The ratio of the difference in the number of granted Rule 12(b)(6) motions to the number of cases with Rule 12(b)(6) motions filed under more

---

[78] For example, any case that would be settled under either pleading standard is potentially negatively affected by a switch to more demanding pleading, since settlement payments in the model above depend partly on γ_d and γ_p.

[79] *See* Part V of Gelbach, *Locking the Doors to Discovery*, *supra* note 11, at 2315-2324.

[80] Whether plaintiffs in SM_D cases are better off under *Conley* or under more demanding pleading can be shown to depend on the values of model parameters.

[81] Gelbach, *Locking the Doors to Discovery*, *supra* note 11, at 2319.

[82] Joe Cecil has rightly pointed out that in *Locking the Doors to Discovery*, I failed to note that AM_D cases involve negative effects on plaintiffs (Cecil doesn't mention SM_D cases, though I also failed to note that these cases also might be negatively affected, depending on parameter values). This failure has no impact on any of my empirical calculations or their interpretation. The presence of such cases simply increases the distance between my lower bound on the number of negatively affected plaintiffs and the actual number, among cases with Rule 12(b)(6) motions filed under more demanding pleading.

demanding pleading is thus a lower bound on the share of these cases whose plaintiffs are negatively affected by *Twombly* and *Iqbal*.

To make all this concrete, observe that the number of negatively affected cases among those with Rule 12(b)(6) motions granted under more demanding pleading is $\eta_G \equiv N(SM_G, AM_G, M_D M_G)$, where the function N equals the sum of the number of cases having the type of each of the function's arguments. The number of negatively affected cases among those with Rule 12(b)(6) motions filed (whether granted or not) under more demanding pleading is $\eta_F \equiv N(SM_G, AM_G, M_D M_G, AM_D, SM_D^-)$, where SM_D^- is the number of SM_D cases whose plaintiff's realized payoff is lower under more demanding pleading than under *Conley*. Since $\eta_F = \eta_G + N(AM_D, SM_D^-)$, any lower bound on η_G is a lower bound on η_F. Finally, define λ_G to be the negatively affected share of cases that have Rule 12(b)(6) motions granted under more demanding pleading, and define λ_F to be the negatively affected share of cases that have Rule 12(b)(6) motions filed under more demanding pleading:

$$
\begin{aligned}
\lambda_G &\equiv \frac{\eta_G}{Grants_{TI}} \\
\lambda_F &\equiv \frac{\eta_G}{MFiled_{TI}},
\end{aligned}
\tag{20}
$$

where $Grants_{TI} = N(SM_G, AM_G, M_D M_G, M_G M_G)$ and $MFiled_{TI} = N(SM_G, AM_G, M_D M_G, M_G M_G, SM_D, AM_D, M_D M_D)$ are the numbers of Rule 12(b)(6) motions filed under more demanding pleading.

The number of cases with Rule 12(b)(6) motions granted under *Conley* is $Grants_C = N(M_G D, M_G S, M_G M_G)$.[83] Since $M_G M_G$ cases appear in both Grants_C and Grants_TI, they drop out of the cross-pleading regime difference in the number of grants, which equals

$$
\Delta Grants = N(SM_G, AM_G, M_D M_G) - N(M_G D, M_G S) \leq \eta_G \leq \eta_F.
$$

It follows immediately that $\lambda_G^* \equiv \frac{\Delta Grants}{Grants_{TI}}$ is a lower bound on λ_G, while $\lambda_F^* \equiv \frac{\Delta Grants}{MFiled_{TI}}$ is a lower bound on λ_F.

There are convenient ways to re-write these lower bounds in terms of the observed Rule 12(b)(6) motion grant rates under *Conley* and under more demanding pleading, as well as the numbers of cases with Rule 12(b)(6) motions filed under each pleading regime. Let $g_C = \frac{Grants_C}{MFiled_C}$ be the observed grant rate under *Conley*, and let $g_{TI} = \frac{Grants_{TI}}{MFiled_{TI}}$ be the observed grant rate under more demanding pleading. We can re-write the lower bound λ_F^* as follows:

$$
\lambda_F^* \equiv (g_{TI} - g_C) + g_C \times \left[\frac{MFiled_{TI} - MFiled_C}{MFiled_{TI}} \right].
\tag{21}
$$

The form of (21) shows that λ_F^* can be written as the increase in the observed Rule 12(b)(6) motion grant rate, $g_{TI} - g_C$, plus a correction term. The correction term is the product of the Rule 12(b)(6) motion grant rate under *Conley* and the increase in the number of Rule

---

[83] Recall that Proposition 4 establishes that there are no $M_G A$ or $M_G M_D$ cases.

12(b)(6) motions filed, as a share of the number of Rule 12(b)(6) motions filed under more demanding pleading. The correction term will be positive when the number of Rule 12(b)(6) motions filed increases after a switch to more demanding pleading. The correction term will be positive and *large*, relative to the change in the grant rate, when this increase is large as a share of the number of Rule 12(b)(6) motions filed under *Conley*. Also, (20) and the definition of g_{TI} make it clear that $\lambda_G^* = \frac{\lambda_F^*}{g_{TI}}$.

This discnn shows that the increase in the Rule 12(b)(6) motion grant rate plays a role in determining the lower bounds on my negatively affected shares, but its part it need not be large. The data in Table 4, which ultimately come from the FJC reports and which I have assembled from figures provided in my earlier paper,[84] illustrate this point well. In the table's first two columns, I provide the share of Rule 12(b)(6) motions in which defendants prevail in Rule 12(b)(6) motions.[85] This share is the appropriate empirical analogue to the Rule 12(b)(6) motion grant rate as I used it in my analytical discussion above. During the pre-*Twombly* period, the share of Rule 12(b)(6) motions in which defendants prevailed was 56.3% for Contracts, Torts, and Other cases and roughly 60% for Employment Discrimination and Civil Rights cases. For the post-*Iqbal* period, the share was 55.2% for Contracts, Torts and Other cases, 61.1% for Employment Discrimination cases, and 68.1% for Civil Rights cases. For Employment Discrimination cases and Contracts, Torts, and Other cases, the increases in the defendant's rate of prevailing—0.2 and 1.1 percentage points—are essentially indistinguishable from zero. For Civil Rights cases, the increase of 7.8 percentage points is more substantial.[86]

---

[84] *See* the notes to **Table 4** for details on the various data elements' source.

[85] The defendant prevails when either (i) the motion is granted without leave to amend, or (ii) the motion is granted with leave to amend but the plaintiff does not file an amended complaint, or (iii) the motion is granted with leave to amend but the plaintiff files an amended complaint, which leads to a chain of events ending in either (i) or (ii). For more detail on the rate at which defendants prevail, *see* either Gelbach, *Locking the Doors to Discovery*, *supra* note 11, at 2293-2294, or the updated FJC report, *supra* note 25, at 3.

[86] The updated FJC report provides a table showing some results from estimating a binary logit model in which the outcome variable equals one if the defendant prevails and zero otherwise. Updated FJC report, at 8, tbl. A-2. The regressors are district court dummies, case-type dummies, a post-*Iqbal* period dummy, and interactions of the case-type and post-*Iqbal* period dummies. The report indicates that the estimated coefficients on these interaction terms are not statistically significant (except for cases involving financial instruments, which I exclude due to the financial crisis). *Id.* In his critique of my earlier paper, Cecil points this fact out, claiming that "corrections for factors unrelated to *Twombly* and *Iqbal* often account for the statistically significant differences that appear in the simple comparison between the pre-*Twombly* and post-*Iqbal* periods." Cecil, *Of Waves and Water*, *supra* note 37, at 39. This is not the place to respond in detail to this criticism (though I plan to do so soon), a couple points are worth noting. First, the differences in the grant rate I use here are essentially zero for the Employment Discrimination and Contracts, Torts, and Other case categories, so any argument plausibly related to my use of the simple grant rate figures is limited in relevance to the Civil Rights case category. Second, it is hardly clear that the logit model in question "correct[]s for factors unrelated to *Twombly* and *Iqbal*." A logit model with only case-type dummies, a post-*Iqbal* period dummy, and interactions of these variables would be fully saturated and so would just re-generate the simple differences in mean rates at which defendants prevail. Therefore, the district dummies are doing all the work to which Cecil points. It can be shown that the only way district dummies could be statistically relevant in a repeated cross-section is if there are important changes in the pattern of Rule 12(b)(6) motion filing across districts. Cecil evidently regards such changes as "factors unrelated to *Twombly* and *Iqbal*." But the whole point of my argument is that parties can be expected to change their behave in response to a change in the pleading regime. There is no reason to doubt that such changes might vary across districts. So where Cecil sees "factors unrelated to

Table 4: Lower Bounds on the Negatively Affected Shares

Case Type[a]	$(1)^a$ g_C	$(2)^b$ g_{TI}	$(3)^c$ $\dfrac{MFiled_{TI} - MFiled_C}{MFiled_{TI}}$	$(4)^d$ Correction Term for λ_F^*	$(5)^e$ Δg	$(6)^f$ λ_F^*	$(7)^g$ λ_G^*
Employment discrimination	60.9	61.1	0.25	15.2	0.2	15.4	25.2
Civil Rights	60.3	68.1	0.17	10.4	7.8	18.2	26.7
Contracts, Torts, and Other	56.3	55.2	0.37	20.4	1.1	21.5	38.9

[a] *Source:* Gelbach, *Locking the Doors to Discovery*, Table 4, first column; expressed in percentage terms.
[b] *Source:* Gelbach, *Locking the Doors to Discovery*, Table 4, second column; expressed in percentage terms.
[c] *Source:* Gelbach, *Locking the Doors to Discovery*, Table 5, column (3); expressed in share terms.
[d] Product of columns (1) and (3); expressed in percentage points.
[e] Equals column (2) minus column (1); expressed in percentage points.
[f] Column (4) plus column (5); expressed in percentage points.
[g] Column (6) divided by ratio of column (2) to 100; expressed in percentage points.

In column (3) of Table 4, I report the increase in the number of Rule 12(b)(6) motions filed between the FJC's pre-*Twombly* and post-*Iqbal* study periods, as a share of the number filed in the post-*Iqbal* period.[87] This increase was non-trivial for all three groups of cases, accounting for 17% of Rule 12(b)(6) motions filed in Civil Rights cases post-*Iqbal*, 25% of these motions in Employment Discrimination cases, and 37% of them in Contracts, Torts, and Other cases. Consequently, the correction term $\left(g_C \times \left[\frac{MFiled_{TI} - MFiled_C}{MFiled_{TI}} \right] \right)$ from the lower bound formula in (21) is quite substantial for all three types of cases. Column (4) shows that the correction term is 10.4 percentage points for Civil Rights cases, 15.2 percentage points for Employment Discrimination cases, and 20.4 percentage points for Contracts, Torts, and Other cases.

Column (6) of Table 4 reports my lower bound on the share of cases in which plaintiffs were negatively affected among cases with Rule 12(b)(6) motions *filed* under more demanding pleading. Each lower bound equals the sum of the correction term in column (4) and the grant rate increase in column (5). The lower bounds are 15.4% for Employment Discrimination cases, 18.2% for Civil Rights cases, and 21.5% for Contracts, Torts, and Other cases.

In column (7), I use the number of cases with Rule 12(b)(6) motions *granted* under more demanding pleading, rather than all those filed, as the reference population in calculating my lower bound. (Recall that this figure can be calculated by dividing the first lower bound, in column (6), by the Rule 12(b)(6) motion grant rate under more demanding pleading, in column (2).) In cases that had a defendant prevail in a Rule 12(b)(6) motion under more demanding pleading, at least one in four had a plaintiff who was negatively affected by *Twombly* and *Iqbal*.

Twombly and *Iqbal*," I see factors that more plausibly result from *Twombly* and *Iqbal*. On that interpretation, the results from the FJC report's logit model have no relevance.

[87] *See* Gelbach, *Locking the Doors to Discovery*, *supra* note 11, section VI, beginning at 2324, for a detailed discussion of how I used the data in the FJC reports; *see also* the original FJC report, *supra* note 25, at 8-12.

This figure is even greater for Contracts, Torts, and Other cases, for which it is roughly two out of five.

VI. Some Normative Issues Related to More demanding pleading

Parts III-V above analyze how Rule 12(b)(6) and changes in the pleading regime affect both party behavior and payoffs. In this Part, I consider some normative issues related to those results, as well as some open questions.

VI.A. *How do* **Twombly** *and* **Iqbal** *Affect the Distribution of Payoffs Across the Parties, and is that a Good Thing?*

The discussion in section V.D makes clear that some plaintiffs are negatively affected by switching to more demanding pleading in the sense that the actual payoffs they receive—their realized payoffs—will be lower under more demanding pleading than under *Conley*. Intriguingly, switching to more demanding pleading also increases realized payoffs for some plaintiffs. An obvious example involves a plaintiff whose case would have history "F,M,G" under *Conley* and history "Drop" under more demanding pleading. Since the plaintiff would file suit and then litigate and lose a Rule 12(b)(6) motion under *Conley*, her realized payoff would be $-c_f - c_{pm}$. This payoff is negative, so it is necessarily less than the realized payoff of 0 that she achieves under more demanding pleading, since she does not file suit.[88] This discussion shows that it is possible for a plaintiff's realized payoff to either increase or decrease as the pleading regime becomes more demanding; it is also possible for a defendant's realized payoff to either increase or decrease.

Now I change perspective, and ask: Could the plaintiff's expected payoff, as of the moment the dispute begins, ever rise as a result of switching to more demanding pleading? Could the defendant's expected payoff as of this moment ever fall? The answer to both questions is no, as the following Proposition states.

---

[88] Cases in which Rule 12(b)(6) motions would be granted under more demanding pleading are not the only types in which the plaintiff's realized payoff rises as a result of switching to more demanding pleading. Consider a case that would be settled in Stage 2 under *Conley* (history "F,S") but that, under more demanding pleading , would have a Rule 12(b)(6) motion filed in Stage 2 and denied in Stage 3, with the defendant answering in Stage 4 (history "F,M,Deny,A"). In such a case we must have $\sigma_a < 0$, so (16) implies that the plaintiff's realized payoff under *Conley* is $\{-c_f + [(1 - \gamma_{pC})v_{pa} - c_{pm}] + \beta\sigma_{2C}\}$, where "C" subscripts indicate that parameters take on their values under *Conley*. Under more demanding pleading, the plaintiff's realized payoff is $-c_f - c_{pm} + v_{pa}$. The plaintiff will be better off with this payoff than with the Stage 2 settlement amount if $\gamma_{pC}v_{pa} > \beta\sigma_{2C}$. Fixing all other parameters, this condition will hold if β is small enough (i.e., the plaintiff has low enough bargaining power), or if either γ_{pC} or v_{pa} is large enough. In sum, plaintiffs might do better winning a Rule 12(b)(6) motion and obtaining an answer than they do with a Stage 2 settlement when the plaintiff has low bargaining power or believes that the judge would be sufficiently likely to grant a Rule 12(b)(6) motion under the *Conley* pleading standard. The same analysis can be used to show that defendants in such cases do better under *Conley* than under more demanding pleading—the reverse of what one might expect.

Proposition 8: Let $U_p \equiv [\overline{u_p}(F), u_p(S), 0]$ be the plaintiff's expected payoff for the game, and let $U_d \equiv [\overline{u_d}(F), u_d(S), 0]$ be the defendant's expected payoff for the game. Then:

1. If the game's equilibrium history involves "Drop" or "F,A" in the first two stages under *Conley*, and if the equilibrium history does not change as a result of switching to more demanding pleading, then U_p and U_d are unaffected by the switch to more demanding pleading. The same holds if under both pleading regimes (i) the equilibrium history is "S", and (ii) the defendant would answer in Stage 2 were the game to reach that stage.

2. For any other pair of equilibria under the two pleading regimes, U_p is strictly lower under more demanding pleading, while U_d is strictly greater.

Proposition 8 has an intuitive basis. Switching to more demanding pleading hhas two important implications: (i) it sometimes allows the defendant to avoid liability and some litigation costs, and (ii) it never increases the defendant's liability. Suchh a policy change could hardly make the defendant worse off. It is also obvious that such a rule should make plaintiffs worse off in cases that would be litigated both with and without the rule. The result might seem a bit less obvious for cases that would be litigated in the absence of the rule but settled in its presence, since the economic logic of settlement is that there is some surplus to split. But more demanding pleading creates Stage 2 settlement surplus by increasing the costs the parties would bear in the absence of settlement. Thhese costs necessarily come out of the reservation payoffs the parties would have if more demanding pleading did not exist. Therefore, when more demanding pleading induces settlement, it does so not by creating any additional surplus, but instead by redistributing reservation payoffs: more demanding pleading simply reduces the plaintiff's threat point while raising the defendant's.

Thus, there is a definite sense in which more demanding pleading can be said to redistribute from defendants to plaintiffs. But is that a bad thing, a good thing, or a sometimes-bad thing and sometimes-good thing?

One view might be that redistribution from plaintiffs to defendants is a bad thing, since (i) defendants are wealthier than plaintiffs, and (ii) redistribution from the wealthy to others is socially desirable. Even if one accepts this dual-pronged premise, the question of whether courts are the best venue for redistribution to occur has provoked an interesting and important debate,[89] though one that is beyond the scope of the present work to address.

VI.B. *The Merit of Affected Cases*

A second normative vantage point concerning *Twombly* and *Iqbal* involves a practical question: Which types of cases are actually affected by switching to more demanding pleading?

---

[89] *See, e.g.*, LOUIS KAPLOW & STEVEN SHAVELL, FAIRNESS VERSUS WELFARE (2002); Daniel A. Farber, *What (If Anything) Can Economics Say about Equity?*, 101 Michigan L. Rev. 1791 (2003) (reviewing Louis Kaplow & Steven Shavell, *Fairness versus Welfare* (2002)); Louis Kaplow & Steven Shavell, *Why the Legal System Is Less Efficient than the Income Tax in Redistributing Income*, 23 J. Legal Stud. 667 (1994); Chris William Sanchirico, *Taxes Versus Legal Rules as Instruments for Equity: A More Equitable View*, 29 J. Legal Stud. 797 (2000).

One important question concerns whether *Twombly* and *Iqbal* disproportionately affect cases inn which detailed pleading is difficult due to asymmetric availibility of information. For example, plaintiffs in contract disputes between two businesses are relatively likely to have sufficient information to plead with considerable factual detail. But plaintiffs in employment discrimination cases might well not be privy to discussions or decisions involving supervisors. Discovery would be the only way for such plaintiffs to obtain the information necessary to meet the more demanding pleading standard, setting up a Catch 22.[90] Drawing normative conclusions is complicated even here, though. Since discovery is expensive and employers who lose an employment discrimination case face the possibility of fee shifting,[91] simply filing an employment discrimination complaint might convey substantial settlement value to a plaintiff, even to a plaintiff who has no basis to suspect unlawful discrimination. The overall normative effect of more demanding pleading's impact on asymmetric-information cases thus depends crucially on the relative numbers of meritorious and non-meritorious cases in such categories (and the relative weights one places on avoiding errors of under- and over-inclusiveness).

This discussion highlights the normative importance of what I will call the case quality question. For all the sturm and drang over *Twombly* and *Iqbal*, there has been very little scholarship concerning the quality of cases affected by *Twombly* and *Iqbal*.[92] An obvious potential reason why is the one suggested by A. Benjamin Spencer: "it is unknowable whether a dismissed claim was nonetheless meritorious in an absolute sense."[93]

One author who has tried to circumvent this problem is Alex Reinert. In his interesting study,[94] Reinert analyzed a set of cases, decided between 1990 and 1999, when *Conley* still applied. He focused on cases in which (i) Rule 12(b)(6) motion grants were reversed on appeal, and (ii) based on his own reading of the cases, he believed "the pleadings would likely be subject to dismissal under an *Iqbal/Twombly* standard".[95] He then examined the post-remand record in these cases "to determine their ultimate resolution, generating an estimate of the 'success' of thinly pleaded cases during this time period."[96] Finally, he "compared the rate of success in the thinly pleaded cases [he] identified with the success of all cases litigated during the same time

---

[90] Suppose supervisor A wrote an email to Supervisor B stating that the company should "fire Veronica because she's a woman." Veronica might not be able to allege anything more than that the believes she was the victim of sex discrimination, since she would have no access to the e-mail at the time she files suit.

[91] *See* 42 U.S.C.A. § 2000e-5 ("In any action or proceeding under this subchapter the court, in its discretion, may allow the prevailing party . . . a reasonable attorney's fee (including expert fees) as part of the costs").

[92] There has been some interesting work on statutory changes to pleading rules implemented via the Private Securities Litigation Reform Act of 1995 (Pub. L. 104-67, 109 Stat. 737). The PSLRA made multiple changes to securities litigation, including imposing conditions considerably stronger than those in Rule 9 in many securities cases. 15 U.S.C. 78u-4(b)(2)(A) ("the complaint shall, with respect to each act or omission alleged to violate this chapter, state with particularity facts giving rise to a strong inference that the defendant acted with the required state of mind"). A number of authors have considered the PSLRA and the merits of securities litigation. *See, e.g.,* Marilyn F. Johnson, Karen K. Nelson, and A. C. Pritchard, *Do the Merits Matter More? The Impact of the Private Securities Litigation Reform Act*, 23 J. L. Econ. Org. 627 (2007); Stephen J. Choi, *Do the Merits Matter* Less *After the Private Securities Litigation Reform Act?* 23 J. L. Econ. Org. 598 (2007); and Stephen J. Choi, Karen K. Nelson, and A. C. Pritchard, *The Screening Effect of the Private Securities Litigation Reform Act*, 6 J. Empirical Leg. Stud. 35 (2009).

[93] A. Benjamin Spencer, *Understanding Pleading Doctrine*, 108 Michigan Law Review 1, at 24 (2009).

[94] Alex A. Reinert, *The Costs of More demanding pleading,* 86 Indiana Law Journal 119 (2011).

[95] *Id.*, at 134.

[96] *Id.* (footnote deleted).

period for which there are records supplied by theAdministrative Office of the United States Courts." ("Administrative Office").[97] While I credit Reinert's inventiveness, my analytical discussion suggests that even within a fixed pleading regime, the set of all cases litigated is likely to differ importantly from the set of cases selected to have Rule 12(b)(6) motions filed against them. Even holding other things equal, my analysis suggests that the event that a Rule 12(b)(6) motion will be filed is tied to the parties' beliefs about the probability that a Rule 12(b)(6) motion would be granted if filed. If the parties' beliefs are at all correlated with judicial behavior, then—even leaving aside the other methodological issues that Reinert forthrightly acknowledges[98]—the comparison Reinert draws might not be an appropriate one.

In *Locking the Doors to Discovery*, I suggested a measure, the adjudication of defense summary judgment motions, that could help measure the quality of cases that *Twombly* and *Iqbal* affect. As I wrote there:

> One can view *Twombly*, and *Iqbal* by extension, as asking judges to forecast the results of discovery and thus the likely outcome of defense summary-judgment motions. If judges are successful at this task, then cases that would have gone through discovery and been dismissed pursuant to defense summary-judgment motions under Conley will now be dismissed before discovery at the MTD stage. This culling of weak cases will tend to reduce defendants' win rate in summary judgment motions. Thus, determining whether this win rate has fallen, and if so by how much, might provide important evidence on the quality of cases affected by *Twombly* and *Iqbal*.[99]

I am working on an extension of the present paper's model that will make these ideas precise. I anticipate that this extension will appear in a future paper, together with some detailed and newly collected empirical evidence on defense summary judgment adjudication rates. With luck this evidence will help make at least some progress on the case quality question.

VI.C. *Primary Behavior*

A final normative issue concerns effects of *Twombly* and *Iqbal* on primary behavior.[100] Throughout my analytical discussion, I assumed that switching to more demanding pleading has no impact on the number of underlying controversies. This assumption is a reasonable baseline from which to study the effects of switching to more demanding pleading, but of course it could be wrong. Consider two examples that might generate disputes related to putative employment discrimination against members of groups protected from discrimination via disparate treatment or disparate impact.[101]

---

[97] *Id.*

[98] *Id.* (acknowledging that "[e]ach of these steps has specific methodological considerations and difficulties.").

[99] Gelbach, *Locking the Doors to Discovery*, *supra* note 11, at 2337 (footnote omitted).

[100] *See, e.g., Hana v. Plumer*, 380 U.S. 460, 475 (1965) (drawing the procedure-substance distinction in terms of "those primary decisions respecting human conduct which our constitutional system leaves to state regulation") (J. Harlan).

[101] *See* 42 U.S.C.A. § 2000e-2(a) for Title VII's basic prohibition of disparate treatment. Since 1971, Title VII has been held to prohibit disparate impact as well, described as "practices that are fair in form, but discriminatory in operation." *Griggs v. Duke Power Co.*, 401 U.S. 424, 431 (1971).

In the first example, a legally savvy, self-consciously sexist employer decides whether to pay a female employee less than her male peers. He knows that if he does, the employee will be suspicious that she has been discriminated against, and he believes that she would file a lawsuit against him if she thought she could win. The employer also knows that in confidential communications he has had with the employee's supervisor, the supervisor has called her the firm's best worker. And he recalls that he once sent an e-mail to that supervisor asking whether the employee wasn't too "emotional" to handle the stress of her job. He believes the supervisor would tell the truth in any deposition and comply fully with any subpoena *duces tecum*[102]—in other words, he believes that if a lawsuit got to discovery, the employee would find damning evidence. Under the *Conley* pleading regime, the employer quite possibly would feel compelled to settle if the employee filed a lawsuit against him; the threat of a Title VII suit might deter this employer from discriminating. But under *Twombly* and *Iqbal*, the employer has at least some chance of avoiding discovery via a Rule 12(b)(6) motion. If the chance is great enough, the employer might well decide to discriminate after all. This example illustrates the risk that switching to more demanding pleading might eliminate the protective effect of Title VII, subjecting an employee to substantively unlawful sex discrimination.

Second, consider a large firm whose top-level managers are deciding whether to adopt an up-or-out policy for some of its employees. Under the policy, any junior supervisor who isn't promoted within two years would be fired. The firm believes the repolicy will increase workforce productivty and profit but it will obviously also lead the firm to fire some employees. The top-level managers know that at least some of the fired employees will be African American, but the managers correctly believe that the incidence of firing will not vary across workers' race. Even if the firm is confident that it has no exposure on disparate treatment grounds, it faces some risk of disparate impact liability. The policy would be facially race-neutral and the firm's top-level managers genuinely believe it would be in the firm's best business interests, but they might hesitate to institute the policy if they thought they would have to either defend or settle a lawsuit with a reasonable chance of getting to discovery. So under *Conley*, the managers might not implement the new policy; under *Twombly* and *Iqbal*, they might push forward. As I have stated the facts in this example, the policy does not violate Title VII: if it adopted the policy, the firm would not be doing so because it intended to treat African Americans worse than other employees, and the policy would not actually treat them worse (and even if it did, the firm would have a genuine business-judgment reason to adopt the up-or-out policy). As such, switching to more demanding pleading is more likely to vindicate the substantive law than would be the *Conley* pleading regime.

These examples show that either pleading regime will affect primary behavior by employers in ways inconsistent with the goals of the underlying substantive law. The first example shows that switching to more demanding pleading can lead to more disparate treatment, because employers can reasonably expect more demanding pleading to shield them from inculpating discovery. Like Holmes's bad man,[103] such bigots are, *de facto*, free to flout the law. But the second example shows that switching to more demanding pleading can also increase

---

[102] Fed. R. Civ. P. 45.

[103] Justice Oliver Wendell Holmes, *The Path of the Law*, 10 Harv. L. Rev. 457, 459 (1897) ("If you want to know the law and nothing else, you must look at it as a bad man, who cares only for the material consequences which such knowledge enables him to predict, not as a good one, who finds his reasons for conduct, whether inside the law or outside of it, in the vaguer sanctions of conscience.").

employers' freedom to adopt lawful policies, by allaying fears, well-founded in the example, that the policies will induce costly litigation for lawful behavior.[104]

These examples illustrate the fact that reducing the pleading regime's overinclusiveness—its tendency to deter lawful behavior due to rational fears of litigation—might well come at the cost of increasing the pleading regime's underinclusiveness—its tendency to protect unlawful behavior due to rational confidence that meritorious litigation can be stopped in its tracks via Rule 12(b)(6). I believe these twin errors are unlikely to be jointly unavoidable: any pleading regime will be subject to at least some of each type. Consequently, choosing the "optimal" pleading regime necessarily involves accounting for a wide array of litigation's effects not only on parties, but also on anyone whose status as a litigant might be affected by the pleading regime.[105] That task is a daunting one, and it is certainly beyond the scope of the present work.[106]

VII. Conclusion

In this paper, I construct a sequential game-theoretic model of pre-discovery litigation that is rich enough to capture the key pleading standard issues related to *Twombly* and *Iqbal*. The model is also tractable enough to generate a number of important substantive results.

My first set of results concerns whether the fact of party selection, by itself, is sufficient to yield testable implications concerning the Rule 12(b)(6) motion filing rate among filed cases, or the grant rate among cases that face Rule 12(b)(6) motions. The answer is no. Proposition 1 shows that in equilibrium, the litigation game I study here is capable of generating Rule 12(b)(6) motions in any fraction of filed cases—including none or all. Proposition 2 shows the same result for the Rule 12(b)(6) motion grant rate: any fraction of filed Rule 12(b)(6) motions between zero and 100% might be granted in the litigation game's equilibrium. Consequently, as Proposition 3 states formally, the existence of Rule 12(b)(6) might either increase or reduce the number of cases that settle. It follows that screening at the pleading stage might either increase or decrease the costs of judicial administration, even (wrongly) ignoring the costs of litigating the Rule 12(b)(6) motions themselves.

---

[104] To be clear, two slight variations on the second example's theme would make it more like the first: (i) the top-level managers in question might have no business purpose at all for adopting the up-or-out policy, and (ii) the nature of the firm's work might be such that a disparate impact on African Americans both would occur and is entirely predictable. Under these circumstances, adopting the policy would violate Title VII. These variations convert the second example into one in which the switch to more demanding pleading might increase unlawful behavior by removing the threat of litigation from the employer's calculus. I take no position on the relative likelihood of the two scenarios for the second example. Its main point is simply to show the possibility that switching to more demanding pleading can both increase unlawful primary behavior and deter lawful primary behavior. To be sure, it is important to note that deterring lawful behavior and inducing unlawful primary behavior are not identical: in my original example for the second firm, its primary behavior is lawful whether or not it adopts the workforce policy.

[105] *See, e.g.*, § 4.2.1 of Robert G. Bone, Civil Procedure: The Economics Of Civil Procedure 128 (2003).

[106] For a more detailed recent discussion of tradeoffs in designing procedural rules, including related to pleading, *see* Alan B. Morrison, *The Necessity of Tradeoffs in a Properly Functioning Civil Procedure System*, 90 Oregon L. Rev. 993 (2012). *See also* Hylton, *supra* note 54.

My second set of results concern party selection—changes in parties' litigation behavior—following a switch to more demanding pleading. I define a switch to more demanding pleading as a change in the standard for meeting Rule 8(a)(2)[107] that satisfies two conditions. First, at least one party must believe that there has been an increase in the probability a Rule 12(b)(6) motion would be granted, if filed, for at least one judge who might be assigned the case. Second, no party can believe that there has been a reduction this probability for any judge who might be assigned the case. Propositions 4 and 6 show a switch to more demanding pleading causes rampant selection in terms of both plaintiffs' and defendants' behavior. Proposition 7 shows that the array of possible changes in party behavior is very broad—so broad that even with empirical evidence on the Rule 12(b)(6) motion grant rate, the Rule 12(b)(6) motion filing rate, *and* the number of cases filed, it is logically impossible to detect whether a perceived switch to more demanding pleading resulted in any changes in *judges'* behavior. As such, a reasonable observer could conclude that the heated debates over the empirical evidence on Rule 12(b)(6) motion grant rates[108] haven't—couldn't—shed any light at all on the actual effects of *Twombly* and *Iqbal*.

The fundamental issue that generates this uber-selection problem is compositional. Following a switch to more demanding pleading, there will be systematic changes in both the set of cases that will be filed and the subset of these that will face Rule 12(b)(6) motions. These changes wreak havoc on the analytical validity of any effort to use changes in the Rule 12(b)(6) motion grant rate to measure changes in judicial behavior. If compositional change is the problem, then finding a way to focus on a fixed set of cases should be a solution. That idea was the linchpin of the heuristic discussion in my earlier paper, *Locking the Doors to Discovery*,[109] and in section V.D of the present paper I reprise its analytical and empirical results in light of the more rigorous discussion here.

The analytical results show that it is possible to use empirical evidence on grant rates and Rule 12(b)(6) motion filings to identify a lower bound on what I call the negatively affected share of plaintiffs among a particular set of cases—those that are observed facing a Rule 12(b)(6) motion under *Twombly* and *Iqbal* (and separately, the subset of these cases in which the motion is granted). The negatively affected share is the share of a given set of cases in which the plaintiff's actual payoff is lower as a result of the switch to more demanding pleading.

The empirical results, originally reported in *Locking the Doors to Discovery*[110] and based on data from the original[111] and updated[112] FJC reports, show that the fraction of negatively affected plaintiffs is substantial. Among cases with Rule 12(b)(6) motions filed under *Twombly/Iqbal*, at least 15%−more than one in seven—had negatively affected plaintiffs. Among cases with Rule 12(b)(6) motions *granted* under *Twombly/Iqbal*, the negatively affected share

---

[107] *See* Fed. R. Civ. P. 8 ("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").

[108] *See* discussion in section II.A, *supra*.

[109] Gelbach, *Locking the Doors to Discovery*, *supra* note 11.

[110] *Id.*, at 2324-2332.

[111] *See supra* note 11.

[112] *See supra* note 25.

was at least 25% for Civil Rights and Employment Discrimination case categories, and two in five for the Contracts, Torts, and Other category.

Finally, in Part VI, I discuss some normative implications of switching to more demanding pleading. Proposition 8 shows that from the perspective of the moment a controversy develops, the switch to more demanding pleading (i) never causes the plaintiff's expected payoff to rise, nor the defendant's to fall, and (ii) does cause the plaintiff's expected payoff to fall, and the defendant's to rise, in a wide array of situations. As I discuss in section VI, switching to more demanding pleading can thus be seen as inducing a redistribution of resources, on the whole, away from plaintiffs and toward defendants.

This result raises two further normative issues. First, have *Twombly* and *Iqbal* struck their declared targets—have they eliminated primarily low-merit suits? Or has their impact been more scattershot, knocking out meritorious suits while missing lower quality ones? As I suggested in my earlier work,[113] I believe there are conditions under which adjudication results from defense summary judgment motions can be useful in addressing these questions. My future work will include an extension of the present paper's model into the summary judgment phase of litigation and empirical methods for relating adjudication of defense summary judgment motions to case quality.

The second additional normative issue concerns primary behavior. I offer an example in which *Twombly* and *Iqbal* would induce unlawful behavior by shielding it from discovery. But I also offer another in which *Twombly* and *Iqbal* induce behavior that both is consistent with substantive law and would be chilled by the threat of ultimately meritless litigation under *Conley*. These examples show that pleading policy's complexity extends past any tradeoff among a given set of controversies and into the tradeoff between preventing unlawful actions and deterring lawful behavior. If pleading "is the gateway to the federal courts,"[114] then it should be no surprise that changes in pleading policy throw open the door to far-flung changes in litigation.

---

[113] Gelbach, *Locking the Doors to Discovery*, *supra* note 11.

[114] Fairman, *supra* note 7, at 988.

VIII. Appendix: Proofs of Propositions 5, 7, and 8

VIII.A. *Proof of Proposition 5*

I will first prove the Proposition under the assumption that changes in the γ parameters do not affect the game's equilibrium history. Differentiating (16) partially with respect to γ_p yields $\{-\beta \max[\sigma_a, 0] + 1(\sigma_a > 0)1(\gamma_d c_{da} - c_{dm} > 0)[-v_{pa} + \beta \times 1(\sigma_2 > 0)\partial\sigma_2/\partial\gamma_p]\}$. Since $\frac{\partial\sigma_2}{\partial\gamma_p} = v_{pa} + \beta\max[\sigma_a, 0]$, the overall derivative is $-\beta \max[\sigma_a, 0][1 - \beta \times 1(\sigma_a > 0)1(c_{da} - c_{dm} > 0)1(\sigma_2 > 0)] - v_{pa}1(\sigma_a > 0)1(\gamma_d c_{da} - c_{dm} > 0)[1 - 1(\sigma_2 > 0)]$, which is negative whenever $\sigma_a > 0$ or $\gamma_d c_{da} - c_{dm} > 0$, and zero otherwise (i.e., the derivative equals zero when the defendant answers in Stage 2 and is negative otherwise). Differentiating (16) partially with respect to γ_d yields $1(\sigma_a > 0)1(\gamma_d c_{da} - c_{dm} > 0)1(\sigma_2 > 0)(-c_{da} + (1 - \beta) \max[\sigma_a, 0])$. This term is zero when the defendant answers and, since $(1 - \beta)\sigma_a$ is always less than c_{da}, negative otherwise. This establishes that $u_p(F)$ is weakly decreasing in the γ parameters.

Differentiating (15) partially with respect to γ_d yields zero when the defendant answers and $[c_{da} - (1 - \beta) \max[\sigma_a, 0] + (1 - \beta)1(\sigma_2 > 0)\partial\sigma_2/\partial\gamma_d]$ when she files a Rule 12(b)(6) motion or the parties settle. Since the partial derivative of the Stage 2 surplus with respect to γ_d is $-c_{da} + (1 - \beta) \max[\sigma_a, 0]$, the full partial derivative equals $\{c_{da} - (1 - \beta)\max[\sigma_a, 0]\} \times [1 - (1 - \beta)1(\sigma_2 > 0)]$. Since $(1 - \beta)\sigma_a$ is always less than c_{da}, the term in the curly braces is positive, and so is the term in square brackets, since $\beta < 1$. Differentiating (15) partially with respect to γ_p yields zero when the defendant files an answer and $v_{pa} + \beta\max[\sigma_a, 0] > 0$, otherwise. This establishes that $u_d(F)$ is weakly increasing in the γ parameters.

It remains to show that the Proposition also holds when switching to more demanding pleading affects the game's equilibrium history. One way to do this would be to grind out the parties' payoffs under each pleading regime's equilibrium and show that they must change in the claimed way. A much shorter way is to observe that at points in the parameter space where a party is indifferent between any pair of possible actions, both parties' payoffs are continuous in all parameters. This fact can be seen by simply inspecting the various inequality conditions for preferring one action to another, inserting an "=" sign where either "<" or ">" appears, and then observing the the payoffs must be continuous at these points. We can then decompose the total effect, on the parties' various payoff functions, of a change in each γ_d^j and γ_p^j into components reflecting (i) the change in payoffs that results from changes in the γ parameters necessary to move the action-taking party from her preferred action under *Conley* to indifference between that action and the one taken under more demanding pleading, and (ii) the change in payoffs that results from changing the γ parameters "the rest of the way". The first part of the proof of the present Proposition, just above, establishes that within each of components (i) and (ii), the parties' payoffs either don't change at all or change in the way claimed in the Proposition. Since payoffs are continuous at the boundaries between preferred actions—between components (i) and (ii)—the Proposition must hold when switching to more demanding pleading affects the game's equilibrium history. ∎

VIII.B. *Proof of Proposition 7*

For this proof only, let S_1 refer to cases that are settled in Stage 1 and S_2 refer to cases that are settled in Stage 2 (rather than to settlement amounts in stages 1 and 2). Let $r \in \{C, TI\}$ index the pleading regime. Let Z_r be the number of cases whose types have Rule 12(b)(6) motions *filed and granted* under pleading regime R only. Let Y_r be the number of cases whose types have Rule 12(b)(6) motions *filed* under pleading regime R only. Let X_r be the number of cases whose types are filed by plaintiffs under pleading regime R only. Notice that $Y_r \geq Z_r$. Define W to be the number of cases whose types are filed regardless of the pleading regime.

It will be helpful to collect together case types represented by Z_C, Y_C, X_C, Z_{TI}, Y_{TI}, X_{TI}, and W, which I do in the first two columns of the following table.

Group	Cases Represented	Types of Cases with Positive Numbers in Example Constructed Here
Z_C	M_GD, M_GS_1, M_GS_2	M_GS_2
Y_C	Z_C cases, M_DD, M_DS_1, M_DS_2	M_GS_2, M_DS_2
X_C	S_2D, AD, M_DD, M_GD, S_2S_1, AS_1, M_DS_1, M_GS_1	S_2D, AD, S_2S_1, $_{AS1}$
Z_{TI}	S_1M_G, S_2M_G, AM_G, M_DM_G	M_DM_G
Y_{TI}	Z_{TI} cases, S_1M_D, S_2M_D, AM_D	S_2M_D, AM_D
X_{TI}	S_1S_2, S_1A, S_1M_D, S_1M_G	S_1S_2, S_1A
W	AA, AS_2, AM_D, AM_G, S_2S_2, S_2M_D, S_2M_G, M_DS_2, M_DM_D, M_DM_G, M_GS_2, M_GM_G	AA, AS_2, AM_D, S_2S_2, S_2M_D, M_DS_2, M_DM_D, M_DM_G, M_GS_2, M_GM_G

For the analysis below, I will set to zero the number of cases of types M_GD, M_GS_1, M_DD, M_DS_1, S_1M_G, S_1M_D, S_2M_G, and AM_G. The third column of the table above shows the case types with possibly non-zero numbers of cases in the example here.

By construction, cases represented by Z_C, Y_C, and X_C are all absent under more demanding pleading, and cases represented by Z_{TI}, Y_{TI}, and X_{TI} are all absent under *Conley*. Thus, no case represented in Z_C, Y_C, or X_C is represented in either Z_{TI}, Y_{TI}, or X_{TI}, and vice-versa.

Now, the third column of the table above shows that Z_C is the number of M_GS_2 cases, and Y_C equals Z_C plus the number of M_DS_2 cases. Thus we can set Z_C and Y_C to any values we like with appropriate choices of M_GS_2 and M_DS_2, provided each is a whole number. Moreover, observe that no cases represented by either Z_C or Y_C are included in X_C. Therefore, we can set the triple (Z_C, Y_C, X_C) to any triple of whole numbers satisfying $Y_C \geq Z_C \geq 0$ and $X_C \geq 0$. The third column of the table above shows that the case types represented by Z_{TI}, Y_{TI}, and X_{TI} are all

mutually disjoint. Therefore, we can set (Z_{TI}, Y_{TI}, X_{TI}) to any triple of whole numbers satisfying $Y_C \geq 0$, $Z_C \geq 0$, and $X_C \geq 0$. Finally, observe that

$$W = Y_C + Y_{TI} + AA + AS_2 + S_2S_2 + M_DM_D + M_DM_G + M_GM_G, \tag{A-22}$$

so that any value of W is feasible provided that it exceeds the sum (Y_C+Y_{TI}). Moreover, we can vary $(AA+AS_2+S_2S_2)$ to set the value of W to any whole number at least as great as $W \geq (Y_C+Y_{TI}+M_DM_D+M_DM_G+M_GM_G)$.

So far, then, I have shown that any integer-valued $(Z_C, Y_C, X_C, Z_{TI}, Y_{TI}, X_{TI}, W)$ is feasible provided that the following conditions are satisfied: $Y_C \geq Z_C \geq 0$, $X_C \geq 0$, $Y_{TI} \geq 0$, $Z_{TI} \geq 0$, $X_{TI} \geq 0$, and $W \geq (Y_C+Y_{TI}+M_DM_D+M_DM_G+M_GM_G)$.

Now let ε be any fixed positive whole number. Set $M_DM_D=2\varepsilon$, $M_GM_G=6\varepsilon$, and $M_DM_G=\varepsilon$ (notice that this last choice means that there is at least some judicial behavior effect). Since the number of cases with Rule 12(b)(6) motions granted under *Conley* equals the number of M_GM_G cases plus Z_C, this number is $G_C \equiv 6\varepsilon+Z_C$. The total number of cases with Rule 12(b)(6) motions granted under more demanding pleading is $G_{TI} \equiv 7\varepsilon+Z_{TI}$, since M_DM_G cases have Rule 12(b)(6) motions granted under more demanding pleading. Now define

$$\begin{aligned} \mathcal{S}_{Z_C} &\equiv \{\varepsilon, 2\varepsilon, 3\varepsilon, 4\varepsilon, 5\varepsilon, 6\varepsilon, 7\varepsilon\} \\ \mathcal{S}_{Z_H} &\equiv \{0, \varepsilon, 2\varepsilon, 3\varepsilon, 4\varepsilon, 5\varepsilon, 6\varepsilon\} \end{aligned} \tag{A-23}$$

Let Z_C have domain $\mathcal{S}_{Z_C}$, and let Z_{TI} have domain $\mathcal{S}_{Z_H}$. Then the domain of G_C, the number of cases with Rule 12(b)(6) motions granted under *Conley*, is given by

$$\mathcal{S}_G \equiv \{7\varepsilon, 8\varepsilon, 9\varepsilon, 10\varepsilon, 11\varepsilon, 12\varepsilon, 13\varepsilon\}, \tag{A-24}$$

as is G_{TI}, the number of cases with Rule 12(b)(6) motions granted under more demanding pleading.

Next, observe that the number of cases with Rule 12(b)(6) motions filed under *Conley* equals the total number of cases with Rule 12(b)(6) motions filed and granted under *Conley*, plus the total number with Rule 12(b)(6) motions filed and denied. The number of cases with Rule 12(b)(6) motions filed and denied under *Conley* equals the number of M_DM_D and M_DM_G cases, plus Y_C, minus Z_C. Since there are 3ε M_DM_D and M_DM_G cases, the total number of cases with Rule 12(b)(6) motions filed under *Conley* is $MFiled_C \equiv G_C+3\varepsilon+(Y_C-Z_C)$. Plugging in $G_C=6\varepsilon+Z_C$, we have $MFiled_C=9\varepsilon+Y_C$. Similar reasoning shows that the total number of cases with Rule 12(b)(6) motions filed under more demanding pleading is $MFiled_{TI} \equiv G_{TI}+2\varepsilon+(Y_{TI}-Z_{TI})=9\varepsilon+Y_{TI}$.

Now define the set

$$\mathcal{S}_Y \equiv \{7\varepsilon, 9\varepsilon, 11\varepsilon, 13\varepsilon, 15\varepsilon\}. \tag{A-25}$$

Letting Y_C and Y_{TI} each have domain $\mathcal{S}_Y$, it follows that $MFiled_C$, the number of cases with Rule 12(b)(6) motions filed under *Conley*, has domain

$$\mathcal{S}_{MF} \equiv \{16\varepsilon, 18\varepsilon, 20\varepsilon, 22\varepsilon, 24\varepsilon\}, \tag{A-26}$$

49

Plugging in the fixed values of $M_D M_D$, $M_G M_G$, and $M_D M_G$ into (A-22) shows that $W=9\varepsilon+Y_C+Y_{TI}+AA+AS_2+S_2 S_2$. Now set $(AA+AS_2+S_2 S_2)=30\varepsilon-(Y_C+Y_{TI})$, which is always possible since $(Y_C+Y_{TI})\leq30\varepsilon$ for every (Y_C,Y_{TI}). It follows that $W=39\varepsilon$ always holds. Now define the set

$$\mathcal{S}_X \equiv \{\varepsilon, 5\varepsilon\}, \tag{A-27}$$

and let $\mathcal{S}_X$ be the domain for each of X_C and X_{TI}. Then the total number of cases filed under *Conley*, $CFiled_C$, has domain

$$\mathcal{S}_{CF} \equiv \{40\varepsilon, 44\varepsilon\}, \tag{A-28}$$

as does the total number of cases filed under more demanding pleading, $CFiled_{TI}$.

Since each element of $\mathcal{S}_{CF}$, $\mathcal{S}_{MF}$, and $\mathcal{S}_G$ can be observed as a number of cases filed, number of Rule 12(b)(6) motions filed, and number of Rule 12(b)(6) motions granted, the domain of $(G_C, MFiled_C, CFiled_C)$ is $\mathcal{S}_G \times \mathcal{S}_{MF} \times \mathcal{S}_{CF}$; likewise for $(G_{TI}, MFiled_{TI}, CFiled_{TI})$. Since the value of $(G_C, MFiled_C, CFiled_C)$ does not restrict the possible values of $(G_{TI}, MFiled_{TI}, CFiled_{TI})$, and vice-versa, it follows that the domain of $(G_C, MFiled_C, CFiled_C, G_{TI}, MFiled_{TI}, CFiled_{TI})$ is $\mathcal{S}_G \times \mathcal{S}_{MF} \times \mathcal{S}_{CF} \times \mathcal{S}_G \times \mathcal{S}_{MF} \times \mathcal{S}_{CF}$, or, equivalently, the domain of $(G_C, G_{TI}, MFiled_C, MFiled_{TI}, CFiled_C, CFiled_{TI})$ is $\mathcal{S}_G \times \mathcal{S}_G \times \mathcal{S}_{MF} \times \mathcal{S}_{MF} \times \mathcal{S}_{CF} \times \mathcal{S}_{CF}$.

Rcall that the set of possible values for the total number of cases filed is $\mathcal{S}_{CF} = \{40\varepsilon, 44\varepsilon\}$. From (A-26), we know that 20ε and 22ε are both elements of $\mathcal{S}_{MF}$. Therefore, for each $C \in \mathcal{S}_{CF}$, it is possible to observe a Rule 12(b)(6) motion filing rate equal to ½. Moreover, the least element of $\mathcal{S}_{MF}$ is 16ε, and 40ε is the least element of $\mathcal{S}_{CF}$, so for each $C \in \mathcal{S}_{CF}$, it is also possible to observe a Rule 12(b)(6) motion filing rate less than ½. Since 24ε is an element of $\mathcal{S}_{MF}$ and 44ε is the greatest element of $\mathcal{S}_{CF}$, it is possible to observe a Rule 12(b)(6) motion grant rate greater than ½ for each $C \in \mathcal{S}_{CF}$. Thus for each pleading regime, it is possible to observe a Rule 12(b)(6) motion filing rate less than, equal to, or greater than ½. Repeating this same argument with the grant rate in place of the filing rate, and with $\mathcal{S}_{MF}$ in place of $\mathcal{S}_{MF}$ and $\mathcal{S}_G$ in place of $\mathcal{S}_{MF}$, we see that the Rule 12(b)(6) motion grant rate can be less than, equal to, or greater than ½ for each pleading regime, regardless of the value of the Rule 12(b)(6) motion filing rate or the number of cases filed.

Now define $\mathcal{R}$ to be the range of the function v, which maps from possible values of $(G_r, MFiled_r, CFiled_r)$ to possible values of $(g_r, m_r, CFiled_r)$, where r again indexes the pleading regime. Since I have constructed this example so that the *Conley* and more demanding pleading values of $(G_r, MFiled_r, CFiled_r)$ are mutually non-restricting, it follows that $\mathcal{R} \times \mathcal{R}$ is the set of possible values for $(g_C, m_C, CFiled_C, g_{TI}, m_{TI}, CFiled_{TI})$. Observe from the discussion above that $\mathcal{R}$ includes elements of the form $(a, b, CFiled_C)$ and $(a, b, CFiled_{TI})$, with each of the nine possible combinations of sgn$(a-\frac{1}{2})$ and sgn$(b-\frac{1}{2})$ included for each fixed value of the pair $(CFiled_C, CFiled_{TI})$. This establishes that for any pair $(CFiled_C, CFiled_{TI})$ in the set $\{40\varepsilon, 44\varepsilon\} \times \{40\varepsilon, 44\varepsilon\}$, it is possible to observe the grant rate rising, falling, or staying the same, regardless of the change in the filing rate; the same is true for the filing rate, regardless of the change in the grant rate. Now observe that by varying $(CFiled_C, CFiled_{TI})$ appropriately, we can make the difference in the number of cases filed rise, fall, or stay the same, regardless of the directions of change of the

filing and grant rates. Thus, any of the 27 possible combined changes in these three variables is possible.

That establishes part 1 of the Proposition, since this example includes a positive judicial behavior effect, with $M_D M_G = \varepsilon > 0$. To establish part 2, put $M_D M_G = 0$ instead, and $M_G M_G = 7\varepsilon$ (rather than 6ε). Then let Z_C have domain $\mathcal{S}_{Z_{TI}}$ rather than $\mathcal{S}_{Z_C}$, which preserves $\mathcal{S}_G$ as the domain of G_{TI}. Since nothing else changes, the rest of the argument through the preceding paragraph goes through, establishing part 2 of the Proposition and completing the proof. ∎

VIII.C. *Proof of Proposition 8*

Expected payoffs are 0 for both parties if the equilibrium involves the plaintiff dropping the case in Stage 1, and realized payoffs are unaffected by changes in γ_p^j and γ_d^j when the game gets to Stage 2 and the defendant answers. Now suppose that the equilibrium actually involves Stage 1 settlement, with the proviso that the defendant would always choose to answer in Stage 2 were settlement not possible. In that situation, the parties' expected and realized payoffs as of Stage 1 are the same, and these payoffs do not depend on either party's subjective probability that a Rule 12(b)(6) motion would be granted (since filing a Rule 12(b)(6) motion would be off the equilibrium path even in the absence of Stage 1 settlement). Therefore the surplus from settling in Stage 1 would not depend on any γ_d^j or γ_p^j parameter, so changes in these parameters cannot affect the settlement amount, nor, therefore, the parties' expected Stage 1 payoff. Thus I have established part 1 of the Proposition.

To establish part 2, I first show the claimed result holds whenever the plaintiff would file suit or the parties would settle in Stage 1 under *Conley*, but when the plaintiff would drop the case under more demanding pleading. By revealed preference, the payoff under *Conley* must be positive, since the plaintiff could drop the suit and receive a payoff of zero. Thus if the switch to more demanding pleading causes the plaintiff to drop the case, her expected payoff for the game must fall as a result; the defendant's payoff must rise since she receives a payoff of zero rather than bearing at least some costs from Stage 1 settlement or litigation/future settlement. Thus I have established the result whenever the plaintiff drops the case under more demanding pleading (Proposition 6 shows that the plaintiff will always drop the case under more demanding pleading if she drops it under *Conley*), and I have disposed of all situations involving a history of "Drop" under either pleading standard.

That leaves only the pairs of equilibrium game histories in which either the parties settle in Stage 1 or the plaintiff files suit in Stage 1, with the defendant not choosing to answer in either situation were the game to reach Stage 2. First, consider the situation when the plaintiff files suit under either equilibrium game history and the defendant would not answer in Stage 2 if the game got that far. Part 2 of Proposition 5 immediately implies that part 2 of the present Proposition holds in this situation.

Second, consider the situation in which the parties settle in Stage 1 under both pleading standards, with the defendant preferring not to answer in Stage 2 in the absence of the Stage 1 settlement. We know that the plaintiff's Stage 1 payoff will be her threat point, which is $\overline{u_p}(F)$, plus the share β of the surplus from settlement, which is $-\overline{u_d}(F) - \overline{u_p}(F)$, so the plaintiff's

Stage 1 payoff under each pleading regime will be given by $U_p = (1 - \beta)\overline{u_p}(F) + \beta[-\overline{u_d}(F)]$. By hypothesis, the defendant would not answer in Stage 2 if the plaintiff did file suit, so part 2 of Proposition 5 implies that $\overline{u_p}(F)$ is strictly decreasing in γ_d^j and γ_p^j, while $\overline{u_d}(F)$ is strictly increasing. It follows immediately that U_p must be strictly increasing in γ_d^j and γ_p^j.

Third, consider the case when the plaintiff files suit under *Conley* and the parties settle in Stage 1 under more demanding pleading. Let $\overline{u_{ir}}(F)$ be party i's expected payoff when the plaintiff files suit under pleading regime r, for $i \in \{p, d\}$ and $r \in \{C, TI\}$. The change in the plaintiff's expected payoff as of Stage 1 due to the switch to more demanding pleading is $\Delta U_p = (1 - \beta)\overline{u_{pTI}}(F) + \beta[-\overline{u_{dTI}}(F)] - \overline{u_{pC}}(F) = \{\overline{u_{pTI}}(F) - \overline{u_{pC}}(F)\} + \beta[-\overline{u_{dTI}}(F) - \overline{u_{pTI}}(F)]$. By hypothesis, the defendant would not answer in Stage 2 if the game reached that stage, so part 2 of Proposition 5 implies that $\overline{u_{pTI}}(F) < \overline{u_{pC}}(F)$, which establishes that the term in curly braces in ΔU_p is negative. Part 2 of Proposition 5 also implies and $\overline{u_{dTI}}(F) > \overline{u_{dC}}(F)$, so $-\overline{u_{dTI}}(F) < -\overline{u_{dC}}(F)$. Subtracting $\overline{u_{pTI}}(F)$ from both sides of this inequality yields $-\overline{u_{dTI}}(F) - \overline{u_{pTI}}(F) < -\overline{u_{dC}}(F) - \overline{u_{pTI}}(F)$. Now add and subtract $\overline{u_{pC}}(F)$ to the right hand side, which then equals the sum of $-\overline{u_{dC}}(F) - \overline{u_{pC}}(F)$ and $\overline{u_{pC}}(F) - \overline{u_{pTI}}(F)$. The first of these terms is negative since there would be positive settlement surplus in Stage 1 under *Conley* if it were positive, and by hypothesis the parties do not settle in Stage 1. The second term is negative by part 2 of Proposition 5. Therefore $-\overline{u_{dTI}}(F) - \overline{u_{pTI}}(F)$ must be less than a negative number, so it is negative. That establishes that both components of ΔU_p are negative, proving that the plaintiff's expected payoff as of Stage 1 must fall. A similar argument establishes that the defendant's expected payoff as of Stage 1 must rise.

Finally, consider the case when the parties settle under *Conley* and the plaintiff files suit under more demanding pleading. Then $\Delta U_p = \overline{u_{pTI}}(F) - (1 - \beta)\overline{u_{pC}}(F) - \beta[-\overline{u_{dC}}(F)] = \{\overline{u_{pTI}}(F) - \overline{u_{pC}}(F)\} + \beta[\overline{u_{dC}}(F) + \overline{u_{pC}}(F)]$. We saw above that the term in curly braces is negative. Observe that the second term can be written $-\beta[-\overline{u_{dC}}(F) - \overline{u_{pC}}(F)]$, which will be negative since the term in brackets must be positive for Stage 1 settlement to occur under *Conley*. Thus ΔU_p must be negative. A similar argument again establishes that the defendant's expected payoff as of Stage 1 must rise.

I have thus established both part 1 and part 2 of the Proposition. ∎