



# The Use of Most-Favored-Nation Clauses in Settlement of Litigation

## Citation

Kathryn E. Spier, The Use of Most-Favored-Nation Clauses in Settlement of Litigation, 34(1) RAND J. Econ. 78 (2003).

## Published Version

<http://www.jstor.org/stable/3087444>

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:10611779>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

## Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

# The use of “most-favored-nation” clauses in settlement of litigation

Kathryn E. Spier\*

*Many settlement contracts in litigation involving multiple plaintiffs (or multiple defendants) include “most-favored-nation” (MFN) clauses. If an early settlement includes an MFN and the defendant settles later with another plaintiff for more money, the early settlers receive these terms too. If the defendant knows the aggregate distribution of expected awards but cannot discriminate among the privately informed plaintiffs, then MFNs avoid costly delay. Plaintiffs with weak cases settle early rather than on the courthouse steps. The effects of MFNs on the settlement terms, plaintiffs’ welfare, litigation rates, and the defendant’s ex ante incentives are considered and alternative explanations are explored.*

## 1. Introduction

■ In the spring of 2000, MP3.com faced lawsuits brought by five major record labels: Warner Music Group, BMG Entertainment, EMI Group PLC, Sony Music Entertainment, and Universal Music Group. These five plaintiffs claimed that MP3.com, a service that allows users to listen to music online, had infringed upon their copyrights. By the end of August, MP3.com had settled with four out of the five record labels, paying each a reported \$20 million. Each of these settlement contracts included a so-called most-favored-nation (MFN) provision: If MP3.com settled on better terms with another record label in the future, then the early settler would receive the better terms too.<sup>1</sup> Universal was the only record label that refused to settle on these terms. Early in the fall, a U.S. District Court judge found that MP3.com had deliberately infringed copyrights, and several weeks later, a judgment for Universal was entered for \$50 million.<sup>2</sup> Since the \$50 million was a “judgment” rather than a “settlement,” the MFN clause was not triggered.

MFN clauses are quite common in settlement contracts and, as in the preceding example, typically apply to settlement payments only and not to judgments.<sup>3</sup> This feature raises legiti-

---

\* Northwestern University and NBER; k-spier@kellogg.northwestern.edu.

Thanks to Lucian Bebchuk, David Butz, Jim Dana, Andrew Daughety, Bruce Hay, Louis Kaplow, Alon Klement, Tom Lyons, Leslie Marx, Jennifer Reinganum, Warren Schwartz, several seminar audiences, two referees, and Peter Dicola for the MP3.com example. All errors, of course, remain my own.

<sup>1</sup> The settlement also included a licensing provision that would allow MP3.com to continue providing the service. See “MP3.com Gets Ripped,” *Newsweek*, September 18, 2000, p. 72.

<sup>2</sup> Anna Wilde Mathews and Colleen DeBaise, “MP3.com Deal Ends Lawsuit on Copyrights,” *The Wall Street Journal*, November 11, 2000, p. A3. After liability was determined, Judge Rakoff ruled that the damages would be \$25,000 for each Universal CD that had been digitally copied by MP3.com. A second trial was scheduled to determine how many of Universal’s disks were involved. The “judgment” of \$50 million was actually negotiated in the judge’s chambers on the eve of this second trial.

<sup>3</sup> Florida’s 1997 tobacco settlement had an MFN clause that read: “The Settling Defendants agree that if they enter into any future pre-verdict settlement agreement of other litigation brought by a non-federal governmental plaintiff

mate concerns. In the words of one critic, “because [defendants] are ‘straight-jacketed’ by the most-favored-nations agreements with certain prior settling [plaintiffs], the strong public policies favoring complete settlement are being frustrated.”<sup>4</sup> On the other hand, proponents of these clauses argue that they are a necessary and desirable feature of settlement contracts.<sup>5</sup> This debate is echoed in *The Manual for Complex Litigation*, a reference manual for judges presiding over such cases, which states that “[MFN] clauses can provide an incentive for early settlement as well as an obstacle to later settlement.”<sup>6</sup>

Formally, suppose that many plaintiffs are suing the same defendant. Although the defendant knows the aggregate distribution of damages, each plaintiff has private information about the expected damage award if his or her case goes to trial. Since all plaintiffs look the same to the defendant, he cannot discriminate in his settlement offers. As in Bebchuk (1984), plaintiffs who believe that they have sufficiently strong cases will reject the defendant’s offers to settle and will seek compensation at trial instead.<sup>7</sup> In addition to costly trials, there will be delay: many cases settle on the courthouse steps (Spier, 1992). These “11th-hour” plaintiffs reject the defendant’s early offers because they anticipate, correctly, that the defendant’s offers could only improve with time. Through an MFN clause, the defendant induces these 11th-hour plaintiffs to accept early settlement offers instead.

Although the defendant benefits from the commitment afforded by MFNs, the broader effects of these clauses on the plaintiffs’ welfare, the litigation rate, and the defendant’s incentives for care are more subtle. When the density function representing the plaintiff’s types is constant (as with the uniform distribution) or increasing in a relevant range, then the terms of settlement and the plaintiffs’ payoffs are higher and the overall litigation rate is lower when MFNs are used. In this *ex post* stage, everyone benefits from MFNs. Furthermore, the defendant’s incentives to take precautions to avoid accidents at the *ex ante* stage can be restored at lower social cost through a damage multiplier. When the density function is decreasing in the relevant range, however, the terms of settlement and the plaintiffs’ payoffs are lower and the overall litigation rate is higher with MFNs. Indeed, one can construct examples where the increase in the overall litigation rate could potentially outweigh the cost savings from earlier settlement, leading to higher overall costs of litigation and delay when MFNs are used.

There is a growing economics literature on settlement of litigation with multiple litigants. Che (1996, 2002) looks at the incentives for plaintiffs with private information to consolidate their claims. Miller (1998) gives a survey of the economics of class action litigation, highlighting the conflict of interest between the (often self-appointed) attorney and the dispersed clients he represents. Che and Yi (1993) consider the role of precedent in litigation. Daughety and Reinganum (1999, 2002a) consider settlement negotiations when the defendant may want to keep information about the lawsuit a secret from future plaintiffs. Kornhauser and Revesz (1994a, 1994b) look at multiple-defendant lawsuits under joint and several liability, focusing on the externalities in settlement decisions, and Spier (2002) looks at similar externalities in lawsuits involving multiple plaintiffs and a potentially insolvent defendant.<sup>8</sup> None of these articles, however, discusses MFN clauses.

There is, of course, a large literature on the use of most-favored-customer clauses in supplier

---

on terms more favorable . . .” The full text may be viewed at <http://www.library.ucsf.edu/tobacco/litigation/fl/flsettle.html>. Texas and Mississippi had similar clauses.

<sup>4</sup> In re Chicken Antitrust Litigation, 560 F. Supp. 943 (Ga. 1979). I have reversed the identities of the plaintiffs and defendants in this quote to maintain consistency with the text.

<sup>5</sup> In the 1999 class action settlement of the vitamins price-fixing case, class lawyers who drafted the agreement argued that the MFN was “essential to the settlement and that without it the parties would not have reached agreement.” See re Vitamins Antitrust Litigation, This Document Relates to: All Actions, 1999-2 Trade Cas. (CCH) ¶72, 726.

<sup>6</sup> See *The Manual for Complex Litigation* @ 23.23 at 182 (3d ed., 1995). While this most recent edition of the manual gives a balanced view of MFNs, earlier editions were more critical.

<sup>7</sup> Hay and Spier (1998) and Daughety (2000) give recent surveys of the settlement literature. Spier (1994) shows that trials must arise in every incentive-compatible bargaining mechanism.

<sup>8</sup> Spier and Sykes (1998) explore the externalities between settling plaintiffs and debtholders.

relationships.<sup>9</sup> Most of this literature explores how MFCs soften price competition among rivals, an issue very different from the one considered here.<sup>10</sup> The article most related to mine is Butz (1990), where best-price provisions mitigate the time-inconsistency problem that a monopolist faces when selling a durable good.<sup>11</sup> Coase (1972) argued that a monopolist selling a durable good will not capture full monopoly rents for the following simple reason. If customers with high valuations buy early on at a high price, the monopolist will lower the price later to sell to buyers with lower valuations. If customers expect the price to fall over time, they will be inclined to forgo early purchases and wait for the lower price.<sup>12</sup> Butz (1990) observed that the best-price provision is a mechanism by which the monopolist can commit not to lower the price later, and it allows the monopolist to capture the full monopoly rents.

The use of MFNs in settlement of litigation has both important similarities and important differences. The obvious similarity with Butz (1990) is that both articles hinge on the time-inconsistency problem. In the durable goods context, best-price provisions commit the monopolist not to lower his price to make future sales. In my analysis, MFNs commit a defendant not to raise the settlement offer over time. But the policy implications of the two analyses differ dramatically. In Butz (1990), best-price provisions allow the monopolist to limit the quantity sold and therefore unambiguously harm social welfare.<sup>13</sup> Here, by encouraging early rather than late settlement, MFNs reduce the costs of delay and may even benefit plaintiffs and reduce the litigation rate. The two games are fundamentally different. First, the pretrial bargaining game features common values; a plaintiff's private information about his damages affects the payoffs of both the plaintiff and defendant should the case go to trial. Second, the passage of time before the fixed trial date does not "screen" among the different plaintiff types.<sup>14</sup>

Section 2 lays out the basic model where the population of plaintiffs is drawn from a continuous distribution. Section 3 presents comparative statics and evaluates the social welfare consequences of MFNs. Section 4 discusses several important extensions, including multiple offers before trial, learning about the aggregate distribution, a small number of plaintiffs, multiple groups of plaintiffs, and plaintiff bargaining power. Section 5 explores alternative explanations for MFN clauses and discusses avenues for future research. Proofs are in the Appendix.

## 2. The model

■ A continuum of risk-neutral plaintiffs has been injured by a defendant. Their expected damages,  $x$ , are drawn from a differentiable density function  $f(x)$  that is positive on the support  $[0, \infty)$ . I normalize the total volume or mass of lawsuits to be one, or  $\int_0^\infty f(s)ds = 1$ , and assume that  $F(x) = \int_0^x f(s)ds < 1$ .<sup>15</sup> I also assume that  $F(x)/f(x)$  is strictly increasing.<sup>16</sup> Each plaintiff privately observes his or her own expected damages, although the aggregate distribution,  $f(x)$ , is common knowledge. That is, all the plaintiffs initially look the same to the defendant, so he cannot explicitly discriminate in his settlement offers.

<sup>9</sup> See the survey by Lyon (1998).

<sup>10</sup> See, for example, Salop (1986) and Cooper (1986). More recent work along these lines includes Besanko and Lyon (1993), Schnitzer (1994), McAfee and Schwartz (1994), and Marx and Shaffer (2000). Empirical work on this topic includes Crocker and Lyon (1994) and Scott Morton (1997).

<sup>11</sup> P'ng (1991) considers a related problem when the monopolist is learning about demand over time, and DeGraba and Postlewaite (1994) allow for interdependent demands of the buyers, as would be the case when they are rivals in a product market.

<sup>12</sup> These ideas are formalized by Stokey (1981) and Bulow (1982).

<sup>13</sup> Butz does say that the reader should not draw "sweeping policy conclusions" from this result. MFNs may be necessary, for example, to cover the seller's fixed costs of production (see the discussion in the conclusion of Butz (1990)) or to mitigate the holdup problem (see Butz (1995)).

<sup>14</sup> For more discussion of the relationship between settlement of litigation and the durable goods monopoly, see Spier (1992).

<sup>15</sup> This guarantees that a positive mass of plaintiff types will reject any finite settlement offer.

<sup>16</sup> This monotone hazard rate assumption assures the existence and uniqueness of the defendant's optimal settlement offer in the commitment case, and existence in the case without commitment.

There are two rounds of settlement offers before trial. In the first round, the defendant offers to settle with each plaintiff for  $S_1$ . Each plaintiff individually decides whether to accept or reject this offer. Plaintiffs who accept the offer receive their payment,  $S_1$ , and exit the game. If a plaintiff rejects  $S_1$ , however, then the plaintiff and defendant bear costs  $c_p$  and  $c_d$ , respectively, and move to the next round. In the second round, the defendant offers to settle with the remaining plaintiffs for  $S_2$ , and each plaintiff individually decides whether to accept or reject the second offer. Plaintiffs who accept  $S_2$  receive payment and exit the game, while plaintiffs who reject  $S_2$  go to trial.<sup>17</sup> At trial, the plaintiff and defendant bear costs  $k_p$  and  $k_d$ , respectively, and the plaintiff is compensated for his or her damages,  $x$ .<sup>18</sup> We assume that  $c_p$ ,  $c_d$ ,  $k_p$ , and  $k_d$  are strictly positive, so settlement negotiations and trials are costly. These costs would include direct legal costs as well as opportunity costs such as the loss of managerial focus. Finally, I will let  $M_1$ ,  $M_2$ , and  $M_T$  represent the mass of plaintiffs who settle in the first round, the second round, and go to trial, respectively.

This simple model can easily be generalized in several ways without changing the results. First, trials could be risky in this framework. For example, let  $p$  be the probability that a plaintiff will win at trial (which is common knowledge), and let  $y$  be the damages awarded to the plaintiff conditional upon winning (which is privately observed by each plaintiff). Then the expected damages awarded at trial is  $x = py$  and all of my results are maintained.<sup>19</sup> Second, we may interpret the trials for the plaintiffs as taking place either jointly or separately.<sup>20</sup> If all plaintiffs are originally part of the same settlement class action, for example, then plaintiffs who reject the first-round settlement in my model have “opted out” of the class.<sup>21</sup> Finally, the model assumes that the litigants do not discount the future. This assumption is made to simplify the exposition of the article. If the defendant and the plaintiffs discounted time at the same rate, then all the results would hold (the damages and the costs would be in present-value terms).

□ **Settlement with commitment.** Suppose that the defendant *commits* to a sequence of offers,  $S_1$  and  $S_2$ , at the beginning of the game.<sup>22</sup> A plaintiff has three choices: to accept  $S_1$  right away, to wait and settle in round 2 for a payoff of  $S_2 - c_p$ , or to reject both offers and go to trial for a payoff of  $x - c_p - k_p$ .

*Lemma 1.* Given a sequence of offers,  $S_1$  and  $S_2$ , plaintiffs with damages  $x \leq x^*(S_1, S_2)$  settle out of court and plaintiffs with damages  $x > x^*(S_1, S_2)$  go to trial, where

$$x^*(S_1, S_2) = \max\{S_1, S_2 - c_p\} + c_p + k_p. \quad (1)$$

<sup>17</sup> The careful reader may notice that the plaintiffs do not have the option to drop their cases before trial, an option that a plaintiff would want to exercise if both  $S_2 < 0$  and  $x < k_p$ . Additional assumptions on the distribution  $F(x)$  would guarantee that the defendant would never offer  $S_2 < 0$ , making the issue moot. See Nalebuff (1987).

<sup>18</sup> Under the American Rule, each party bears its own litigation costs.

<sup>19</sup> Risk neutrality is important here, however. With risk aversion, uncertainty introduces inefficient risk bearing at trial.

<sup>20</sup> Importantly, though, I have assumed a constant returns to scale technology, where the cost of litigating a case is independent of the number of cases brought. This is clearly not accurate as a representation of class action litigation. Class members are able to take advantage of economies of scale, whereas individual plaintiffs who opt out would forgo these economies.

<sup>21</sup> See re Vitamins Antitrust Litigation, This Document Relates to: All Actions, 1999-2 Trade Cas. (CCH) ¶72, 726. In settlement class actions, as compared with ordinary class actions, the class itself and the settlement agreement are certified by the judge at the same time. The class members therefore retain discretion to opt out at that point if they are dissatisfied with the class settlement.

<sup>22</sup> Although the defendant could commit to randomize, this would never be optimal. It is easy to show that if  $S_2$  were random, then a truncated distribution of plaintiff types would remain in the second round. The assumption that  $F(x)/f(x)$  is increasing in  $x$  implies a unique optimal offer for the defendant given a truncated distribution. The defendant could reduce his total liability by committing to a single offer instead, one that induces the same cutoff.

<sup>23</sup> The tie-breaking assumption is that a plaintiff accepts an offer when he is indifferent between accepting an offer and going to trial. He may, of course, mix between accepting the early and the late offer.

It is interesting (and important) to note that all plaintiffs have the *same preference orderings* over the two settlement offers. If one plaintiff, type  $x$  say, prefers  $S_1$  to  $S_2$ , then another plaintiff, type  $y$ , will also prefer  $S_1$  to  $S_2$ . The plaintiffs differ, of course, in their preferences between settling and going to trial: plaintiffs with weak cases will settle and those with strong cases will go to trial. This property implies that the passage of time before the last round of settlement does not screen among the plaintiffs in the usual sense. In models of bilateral trade, for example, high-valuation buyers strictly prefer to purchase sooner than their low-valuation counterparts.

It is not hard to see that the defendant will commit to a sequence of offers where nobody settles in the second round. To see why this is true, suppose instead that a positive mass of plaintiffs settled in each round  $M_1 > 0$  and  $M_2 > 0$ .<sup>24</sup> Since the plaintiffs are rational, they must be indifferent between the two settlement offers:  $S_1 = S_2 - c_p$ . Using the lemma, there is a cutoff,  $x^*$ , where  $S_1 = x^* - c_p - k_p$  and  $S_2 = x^* - k_p$ . The defendant's total payments are

$$S_1 M_1 + (S_2 + c_d) M_2 + \int_{x^*}^{\infty} (x + c_d + k_d) f(x) dx. \quad (2)$$

The defendant pays  $S_1$  to the plaintiffs who accept in the first round (mass  $M_1$ ), he pays  $x + c_d + k_p$  for those who settle in the second round (mass  $M_2$ ), and pays  $x + c_d + k_d$  for those cases that go to trial. Since  $S_1 = S_2 - c_p < S_2 + c_d$ , the defendant would like to shift the mass of settlement from the second round into the first round. This is easily accomplished with an alternative sequence of offers,  $S'_1$  and  $S'_2$ , where  $S'_1 = S_1$  and  $S'_2 < S_2$ . By committing to a sufficiently unattractive second-round offer, the defendant encourages plaintiffs with low damages,  $x \leq x^*$ , to settle in the first round instead, saving both his own delay cost  $c_d$  and extracting  $c_p$  from the plaintiffs in settlement. The defendant's total payments under this alternative sequence are

$$(x^* - c_p - k_p) F(x^*) + \int_{x^*}^{\infty} (x + c_d + k_d) f(x) dx. \quad (3)$$

The defendant's optimal settlement strategy is characterized by a cutoff,  $\hat{x}$ , and the corresponding settlement offer,  $S_1 = \hat{x} - c_p - k_p$ , that minimizes this expression.

*Proposition 1.* If the defendant can commit to a sequence of settlement offers, then there is a unique early offer  $S_1 = \hat{x} - c_p - k_p$  and a late offer  $S_2 < \hat{x} - k_p$  that minimize the defendant's total payments, where  $\hat{x}$  solves

$$F(\hat{x}) - (c_p + c_d + k_p + k_d) f(\hat{x}) = 0. \quad (4)$$

In equilibrium,  $M_1 = F(\hat{x})$ ,  $M_2 = 0$ , and  $M_T = 1 - F(\hat{x})$ .<sup>25</sup>

Equation (4), the first-order condition defining the full-commitment cutoff  $\hat{x}$ , may be understood intuitively. When the defendant increases the settlement offer  $S_1$  by a small amount,  $\Delta$ , there are both costs and benefits. The cost is that he pays an additional  $\Delta$  to settle with all plaintiff types below  $\hat{x}$ , so the cost is  $\Delta F(\hat{x})$ . The benefit is that plaintiffs in the range  $(\hat{x}, \hat{x} + \Delta]$  will now settle in the first round rather than go to trial. The mass of plaintiffs in this range is approximately  $\Delta f(\hat{x})$ , so the cost savings on these plaintiffs is  $(c_p + c_d + k_p + k_d) \Delta f(\hat{x})$ . The optimal cutoff,  $\hat{x}$ , equates the marginal cost and marginal benefit.

Finally, it is important to note that commitment is critical here. Since a truncated distribution of plaintiffs,  $F(x)$  on  $(\hat{x}, \infty)$ , remain in the second round, the defendant's second-round offer,

<sup>24</sup> The case of  $M_1 = 0$  and  $M_2 > 0$  is very similar and is not presented.

<sup>25</sup> This first-order necessary condition is analogous to the one in Bebchuk (1984). The monotone hazard rate condition assures that the solution is unique and that the second-order condition is satisfied. If  $S_2 = \hat{x} - k_p$ , then the plaintiffs would be indifferent between the two offers. Although it would certainly be an equilibrium for all plaintiffs to accept  $S_1$ , it is not unique. Offering  $S_2 = \hat{x} - k_p$  is weakly dominated by  $S_2 < \hat{x} - k_p$ .

$S_2 < \hat{x} - k_p$ , is not *sequentially rational*. Indeed, in the second round the defendant would prefer to make a second-round offer that *at least some* of the remaining plaintiffs would accept:  $S_2 > \hat{x} - k_p$ .<sup>26</sup> Anticipating that the settlement offer would rise in this way, however, the plaintiffs with damages  $x \leq \hat{x}$  would not be willing to accept  $S_1 = \hat{x} - c_p - k_p$  to begin with.

□ **Settlement without commitment.** Now suppose that the defendant cannot commit to  $S_2$ . We need to consider the *continuation game* following a first-round settlement offer,  $S_1$ , but before the plaintiffs have made their first-round settlement decisions. The strategies of this game include each plaintiff's decision whether or not to accept  $S_1$ , the defendant's second-round offer,  $S_2$ , and the decision of each plaintiff to accept a second-round offer,  $s$ . In general, each of these strategies could depend upon the history of the game. The problem is simplified, however, by three observations.

- (i) A plaintiff's decision to accept a second-round offer,  $s$ , will not depend on  $S_1$  or on the acceptance/rejection decisions in round 1. The plaintiff simply compares the second-round offer to his payoff at trial and accepts the offer if and only if  $s \geq x - k_p$ .<sup>27</sup>
- (ii)  $S_2$  is chosen optimally by the defendant in round 2 given the aggregate equilibrium distribution of the remaining plaintiffs.
- (iii) Since each plaintiff is infinitesimally small, no individual plaintiff's acceptance/rejection decision could possibly alter the aggregate equilibrium distribution of plaintiffs remaining in round 2. At the beginning of round 1, each plaintiff treats the anticipated  $S_2$  as fixed when weighing the three options—accept  $S_1$ , accept  $S_2$ , go to trial—and plans on pursuing the option that yields the highest payoff.<sup>28</sup>

As a benchmark, suppose that the defendant makes a low-ball offer,  $S_1 = 0$ , and that no plaintiff accepts this offer in the continuation equilibrium. The original distribution of plaintiff types,  $f(x)$ , would remain at the beginning of round 2. The defendant would subsequently choose  $S_2$  to minimize  $\int_0^{s+k_p} s f(x) dx + \int_{s+k_p}^{\infty} (x + k_d) f(x) dx$ . By analogy to the full-commitment case in Proposition 1, the first-order condition from this program gives a unique solution,  $S_2 = \underline{x} - k_p$ , where  $\underline{x}$  solves

$$F(\underline{x}) - (k_p + k_d) f(\underline{x}) = 0. \quad (5)$$

This solution,  $\underline{x}$ , is an important threshold in the continuation game following  $S_1$ .

*Lemma 2.* Given  $S_1$ , continuation equilibria exist and have the following properties:<sup>29,30</sup>

- (i) If  $S_1 > \underline{x} - c_p - k_p$ , the plaintiffs are indifferent between accepting the early offer  $S_1$  and the late offer  $S_2 = S_1 + c_p$ . Letting  $x^* = S_1 + c_p + k_p$ ,  $M_1 = F(x^*) - (k_p + k_d) f(x^*) > 0$ ,  $M_2 = (k_p + k_d) f(x^*) > 0$ , and  $M_T = 1 - F(x^*)$ . The defendant's payments are

$$(x^* - c_p - k_p)F(x^*) + (c_p + c_d)(k_p + k_d)f(x^*) + \int_{x^*}^{\infty} (x + c_d + k_d)f(x)dx. \quad (6)$$

<sup>26</sup> The defendant would save his litigation costs,  $k_d$ , and extract the plaintiffs' litigation costs,  $k_p$ .

<sup>27</sup> As before, a plaintiff is assumed to accept the offer when indifferent.

<sup>28</sup> We may, without loss of generality, restrict attention to pure strategies for the defendant. If he mixed in his choice of second-round offer, then a truncated distribution of plaintiff types would remain at the beginning of round 2. The assumption that  $F(x)/f(x)$  is increasing in  $x$  would give a unique  $S_2$ .

<sup>29</sup> The continuation equilibrium is not unique. When the early offer is high, as in case (i), there is an equilibrium in mixed strategies where every plaintiff below the threshold randomizes between accepting and rejecting. There is also a pure-strategy equilibrium that features a truncated distribution in each round. The lowest types accept  $S_1$ , the middle types accept  $S_2$ , and the highest types go to trial.

<sup>30</sup> This result is similar to example 2 (the uniform distribution) in Spier (1992). However, Spier (1992) makes a further "skimming" restriction on the plaintiff's settlement strategies in order to characterize the outcome. See assumption (A1) in that article. The current article has proven results for two settlement rounds without this assumption and with more general distributions of plaintiff types.

- (ii) If  $S_1 \leq \underline{x} - c_p - k_p$ , then all plaintiffs reject  $S_1$  and the late offer is  $S_2 = \underline{x} - k_p \geq S_1 + c_p$ .  $M_1 = 0$ ,  $M_2 = F(\underline{x})$ , and  $M_T = 1 - F(\underline{x})$ . The defendant's payments are

$$(\underline{x} + c_d - k_p)F(\underline{x}) + \int_{\underline{x}}^{\infty} (x + c_d + k_d)f(x)dx. \quad (7)$$

Case (i) is of particular interest. When the early offer is high, some plaintiffs will settle early and others will settle late. The intuition is easy. If all plaintiffs below  $x^* = S_1 + c_p + k_p$  accepted  $S_1$ , then the defendant would make a better offer in the second round:  $S_2 > x^* - k_p$ . So the early settlers should have waited. If all the plaintiffs rejected  $S_1$ , then the defendant would make a worse offer in the second round:  $S_2 = \underline{x} - k_p$ . So the 11th-hour plaintiffs should have settled early for  $S_1$ . For  $S_2 = S_1 + c_p$  to be sequentially rational for the defendant, it is necessary that exactly  $(k_p + k_d)f(x^*)$  plaintiffs with damages  $x \leq x^*$  remain at the beginning of the second round. If fewer than  $(k_p + k_d)f(x^*)$  remain, then the defendant would want to raise his offer above  $S_1 + c_p$ . If more than  $(k_p + k_d)f(x^*)$  remain, the defendant would want to lower his offer below  $S_1 + c_p$ .

The defendant's total payments in Lemma 2 are a continuous function of  $S_1$ . Since the defendant's payments in expression (7) are constant when  $S_1 \leq \underline{x} - c_p - k_p$ , the defendant's optimal settlement strategy involves choosing a first-round offer,  $S_1 \geq \underline{x} - c_p - k_p$ , and corresponding cutoff,  $x^* = S_1 + c_p + k_p$ , that minimizes his total payments in expression (6).

*Proposition 2.* If the defendant cannot commit to the sequence of settlement offers and MFNs are not used, then he will offer  $S_1 = \tilde{x} - c_p - k_p$  where  $\tilde{x}$  minimizes the defendant's total payments in expression (6). This value satisfies  $\tilde{x} > \underline{x}$  and

$$F(\tilde{x}) - (c_p + c_d + k_p + k_d)f(\tilde{x}) + (c_p + c_d)(k_p + k_d)f'(\tilde{x}) = 0. \quad (8)$$

The first-order necessary condition may be understood intuitively. In addition to the marginal cost and the marginal benefit described for the commitment case, there is now an additional term. This third term represents the change in the volume of cases that settle in the second round. If the defendant increases the first-round settlement offer (and the corresponding cutoff  $\tilde{x}$ ) by a small amount,  $\Delta$ , then the volume of cases that settle in the second round would rise by  $(k_p + k_d)f(\tilde{x} + \Delta) - (k_p + k_d)f(\tilde{x})$ . Therefore, on the margin, the associated cost rises by  $(c_p + c_d)(k_p + k_d)f'(\tilde{x})$ . Notice that this term may be either positive or negative, depending on the slope of the density function.

In this section we have seen that, without commitment, a positive mass of plaintiffs settle on the courthouse steps. Holding the cutoff  $\tilde{x}$  fixed in Proposition 2, the defendant could strictly reduce his total payments by committing to a low second-round offer:  $S_2 < \tilde{x} - k_p$ . Through this commitment, all plaintiffs with damages below  $\tilde{x}$  would accept in the first round, saving the defendant money. Since delay is costly, the defendant is hurt by his inability to commit to the sequence of settlement offers.

□ **Settlement with MFN clauses.** Suppose a mass of plaintiffs accepts  $S_1$  with an MFN clause in the first round and that the defendant subsequently offers to settle for  $S_2 > S_1$  in the second round. The MFN provision obligates the defendant to pay the early-settling plaintiffs the difference between the offers:  $S_2 - S_1$ . We will see that this obligation has strategic value: the defendant can credibly commit to be "tough" and not raise his offer. Although this commitment destroys value *ex post* (because settlement opportunities are destroyed), it is valuable to the defendant from an *ex ante* perspective because more cases settle early.

*Proposition 3.* The defendant can achieve the full-commitment outcome characterized in Proposition 1 by offering to settle in the first round for  $S_1 = \hat{x} - c_p - k_p$  together with an MFN provision. Plaintiffs with damages  $x \leq \hat{x}$  accept  $S_1$ . No further settlement takes place and plaintiffs with damages  $x > \hat{x}$  go to trial.

Although MFNs are an effective commitment device, their use raises a number of issues. First, the practical distinction between "settlements" and "judgments" is not always clear. Defendants,



for example, may try to disguise their future settlements as judgments in order to avoid triggering an MFN. Plaintiffs, on the other hand, may argue that a judgment was, in fact, a settlement in disguise. In the MP3.com case, Universal's \$50 million "judgment" was negotiated in the judge's chambers on the eve of trial. The other four settling record companies (unsuccessfully) challenged its characterization as a judgment. "It doesn't matter what you call it. This is a settlement," said one record executive. "If it walks like a duck, talks like a duck, smells like a duck, it's a duck."<sup>31</sup>

Second, MFNs are, in theory, just one of many ways of achieving commitment here. A high *fixed* payment to the early settlers should the defendant raise the offer in the future would also work, as would a contractual obligation to pay damages to some third party.<sup>32</sup> The MFN, however, has certain advantages over these other devices. From a practical perspective, MFNs have the "flavor" of expectation damages because they put the early settlers in the same position they would have been in had they only waited.<sup>33</sup> Consequently, judges may feel comfortable enforcing MFNs, whereas these other mechanisms would raise suspicion. MFNs may also facilitate experimentation and learning when the defendant is unsure of the aggregate distribution of plaintiff types, an idea explored in more detail in Section 4.

Finally, it is worth noting that the outcome in Proposition 3 is not renegotiation-proof. Even though it is not in the defendant's private interest to raise the settlement offer in the second round, it may be in the joint interest of the defendant and the early-settling plaintiffs. If they could all get together in the second round and seamlessly rewrite their settlement agreements to allow for more settlement, they could all be made jointly better off *ex post* (although they may be worse off *ex ante*). Renegotiation along these lines may be difficult when there are many plaintiffs, however. Each plaintiff would have an incentive to stick with the original agreement in the hopes of receiving the MFN payout.<sup>34</sup> And even if they could coordinate their actions *ex post*, in many legal situations (a settlement class action, for example) it would require the approval of a judge.

### 3. Comparative statics and social welfare implications of MFNs

■ In the previous section I argued that MFNs credibly commit the defendant not to raise his settlement offers and so plaintiffs settle early rather than wait until the 11th hour. This has an unambiguously positive effect on the defendant: holding the terms of settlement fixed, MFNs economize on the costs of delay. In addition, the defendant may change the terms of settlement (and the associated cutoff), further reducing his total liability. The next lemma characterizes the relationship between  $\hat{x}$ , the cutoff when MFNs are used, and  $\bar{x}$ , the cutoff when MFNs are not permitted, and it serves as a basis for comparing the two regimes.

*Lemma 3.*  $\hat{x} > \bar{x}$  (respectively  $\hat{x} < \bar{x}$ ,  $\hat{x} = \bar{x}$ ) if and only if  $f'(\bar{x}) > 0$  (respectively  $f'(\bar{x}) < 0$ ,  $f'(\bar{x}) = 0$ ).

The difference between  $\hat{x}$  and  $\bar{x}$  comes from the cost of delay. When MFNs are not used and commitment is impossible, the volume of plaintiffs who settle in the second round is  $(k_p + k_d)f(\bar{x})$ , and the associated cost of delay is  $(c_p + c_d)(k_p + k_d)f(\bar{x})$ . This delay cost is a critical ingredient in the defendant's choice of settlement offer. When  $f'(\bar{x}) < 0$  then the defendant has an additional incentive (on the margin) to increase his offer slightly; fewer plaintiffs will settle in the second round, lowering the overall costs of delay. In this case, the defendant will choose a higher settlement offer with a higher associated cutoff,  $\bar{x} > \hat{x}$ . When  $f'(\bar{x}) > 0$ , on the other hand, then the

<sup>31</sup> Derek Caney, "Record Labels Fuming over Universal-MP3.com Ruling," [dailynews.yahoo.com](http://dailynews.yahoo.com), November 17, 2000.

<sup>32</sup> Note that this third party would need to be informed of the original contract and have full information about the future offers. Legally, this third party would also have to contribute something of value in exchange for the right to the MFN payout—there must be "consideration." I am not aware of side contracts of this form, perhaps because they would be very difficult to write and even harder to enforce.

<sup>33</sup> At least they appear to. If all plaintiffs waited, then the equilibrium offer would change.

<sup>34</sup> Since plaintiffs are infinitesimally small, any single plaintiff's refusal to concede would not affect the future outcome. Note that these ideas are similar to the holdout problem with tender offers in corporate takeovers.

defendant faces larger delay costs when he raises his offer slightly and will therefore choose a lower settlement offer and a lower associated cutoff,  $\tilde{x} < \hat{x}$ .

□ **The effect of MFNs on the settlement rate.** The overall rate of settlement is the volume of settlement in the first round,  $M_1$ , plus the volume of settlement in the second round,  $M_2$ . When MFNs are not used, there is settlement in each round,  $M_1 = F(\tilde{x}) - (k_p + k_d)f(\tilde{x})$  and  $M_2 = (k_p + k_d)f(\tilde{x})$ , and the overall settlement rate is  $F(\tilde{x})$  (Proposition 2). When MFNs are used, all settlement takes place in the first round,  $M_1 = F(\hat{x})$  and  $M_2 = 0$ , so the overall rate of settlement is simply  $F(\hat{x})$  (Proposition 3). So we see that the overall settlement rate may either rise or fall, depending on the relationship between  $\hat{x}$  and  $\tilde{x}$ . The following proposition follows immediately from Lemma 3.

*Proposition 4.* MFNs increase (respectively decrease, leave unchanged) the overall settlement rate if and only if  $f'(\tilde{x}) > 0$  (respectively  $f'(\tilde{x}) < 0$ ,  $f'(\tilde{x}) = 0$ ).

□ **The effect of MFNs on plaintiff welfare.** Referring back to Propositions 2 and 3, the defendant's first-round offer with an MFN clause is  $S_1 = \hat{x} - c_p - k_p$ , and the defendant's first-round offer without an MFN clause is  $S_1 = \tilde{x} - c_p - k_p$ . It follows that if  $\hat{x} > \tilde{x}$ , then the plaintiffs prefer the settlement offers under the MFN regime.<sup>35</sup>

*Proposition 5.* MFNs weakly increase (respectively weakly decrease, leave unchanged) the plaintiffs' welfare if and only if  $f'(\tilde{x}) > 0$  (respectively  $f'(\tilde{x}) < 0$ ,  $f'(\tilde{x}) = 0$ ).<sup>36</sup>

□ **The effect of MFNs on the total costs of litigation and delay.** When MFNs are prohibited and commitment is otherwise impossible, the total costs are

$$\tilde{\Omega} = [1 - F(\tilde{x})](c_p + c_d + k_p + k_d) + (k_p + k_d)f(\tilde{x})(c_p + c_d). \quad (9)$$

In the first term,  $1 - F(\tilde{x})$  is the volume of cases that go all the way to trial, and  $(c_p + c_d + k_p + k_d)$  are the total costs borne in these cases. In the second term,  $(k_p + k_d)f(\tilde{x})$  is the volume of cases that settle in the second round, and  $(c_p + c_d)$  are the associated delay costs. The total costs when the defendant uses MFNs, thereby committing to settle in the first round only, are

$$\hat{\Omega} = [1 - F(\hat{x})](c_p + c_d + k_p + k_d). \quad (10)$$

Using these two expressions, the relative cost savings from MFNs is

$$\tilde{\Omega} - \hat{\Omega} = (c_p + c_d)(k_p + k_d)f(\tilde{x}) + [F(\hat{x}) - F(\tilde{x})](c_p + c_d + k_p + k_d). \quad (11)$$

The first term captures the benefit of early (rather than late) settlement. This first term is always positive: a commitment to settle early avoids delay costs,  $(c_p + c_d)$ . The second term is the increase in volume of settlement,  $[F(\hat{x}) - F(\tilde{x})]$ , multiplied by the cost savings from settlement. Since  $\hat{x}$  may be either smaller than or greater than  $\tilde{x}$  depending on the sign of  $f'(\tilde{x})$  (Lemma 3), MFN clauses may either raise or lower the overall volume of settlement.

*Proposition 6.* If  $f'(\tilde{x}) \geq 0$ , then the total costs of litigation and delay fall when MFNs are used. If  $f'(\tilde{x}) < 0$ , then the total costs may either rise or fall when MFNs are used.

□ **Examples.** *The uniform distribution.* Suppose  $f'(x) = 0$ . Comparing (4) and (8) we see that the plaintiff type who is indifferent between settling and litigating in the two regimes is exactly the same:  $\hat{x} = \tilde{x}$ . It follows that the overall rate of settlement is unchanged and the plaintiffs are

<sup>35</sup> More specifically, all plaintiffs who accept this offer are better off, and those who reject the offer and go to trial instead are no worse off than before. Without MFNs, plaintiffs would mix between  $S_1$  and  $S_2$  (they were indifferent), and so we need only consider  $S_1$ .

<sup>36</sup> "Weakly" because plaintiffs who go to trial under both regimes receive exactly the same payoff in each.

all equally well off under the two regimes.<sup>37</sup> The total costs of litigation and delay unambiguously fall, since cases that would have settled in the second round settle in the first round instead.

*The exponential distribution.* MFNs may reduce the total costs of delay and litigation even when the distribution of plaintiff types is everywhere decreasing. Consider the exponential distribution,  $f(x) = (1/\lambda)e^{-x/\lambda}$ . A straightforward evaluation of equations (4) and (8) gives us closed-form solutions for  $\hat{x}$  and  $\tilde{x}$ :

$$\begin{aligned}\hat{x} &= \lambda \ln \left[ 1 + \frac{c_p + c_d + k_p + k_d}{\lambda} \right]; \\ \tilde{x} &= \lambda \ln \left[ 1 + \frac{c_p + c_d + k_p + k_d}{\lambda} + \frac{(c_p + c_d)(k_p + k_d)}{\lambda^2} \right].\end{aligned}\quad (12)$$

It is, of course, true that the overall litigation rate rises when MFNs are adopted:  $1 - F(\hat{x}) > 1 - F(\tilde{x})$ . But it is easy to verify that the additional cost from more litigation is outweighed by the benefit of earlier settlement.

*The normal distribution.* When the plaintiff's damages are normally distributed, total costs may either rise or fall, depending upon the parameter values. Suppose that the plaintiffs' damages are normally distributed with a mean of 100 and a standard deviation of 15. The sum of the litigation costs is  $k_p + k_d = 40$ , and the sum of the delay costs is  $c_p + c_d = 10$ . When the defendant can commit with an MFN,  $\hat{x} = 115$  and 83% of all cases settle in the first round. Without an MFN,  $\tilde{x} = 120$  and 91% of the cases settle, 49% in the first round and 42% in the second round. The social cost of MFNs here is that the litigation rate rises by 8% and costs  $c_p + c_d + k_p + k_d = 50$  are borne per case. The benefit is that the 42% of cases settle in the first round instead of the second, saving costs of 10 per case. The overall cost,  $(.08)(50) = 4$ , is smaller than the benefit,  $(.42)(10) = 4.2$ , so MFNs reduce total costs. But, unlike the exponential distribution, this is not always the case. Reducing the standard deviation from 15 to 5, for example, or changing the costs to a more equal allocation where  $k_p + k_d = c_p + c_d = 25$  both will lead to higher costs with MFNs. The comparative statics are not monotone, however, and no general results for the normal distribution have been obtained.

*The binary distribution.* Spier (forthcoming) considers an example with two types of plaintiffs, weak and strong. Their expected damages are  $x_L$  and  $x_H$ , respectively, where  $x_H > x_L$ . When the proportion of weak plaintiffs in the population is in an intermediate range, the defendant will adopt an aggressive low-offer strategy when MFNs are used, leading to higher overall costs of litigation and delay. (Note that this example violates the monotone hazard rate assumption in this article.)

□ **The effect of MFNs on primary behavior.** Up until this point, the normative analysis has focused on the costs of delay and litigation. This may be too short-sighted. Imagine that the defendant chooses precautions to avoid harm at an *ex ante* stage, where higher precautions reduce the probability of harm (but do not change the distribution of damages conditional upon an accident). Since MFNs reduce the defendant's total expected payments, they will tend to dilute his *ex ante* incentives for care.<sup>38</sup> If the defendant has appropriate incentives for care when MFNs are prohibited, then permitting MFNs may reduce social welfare.<sup>39</sup> One should certainly not infer that MFNs should be prohibited on these grounds, however. If MFNs reduce the overall costs of delay—and delay involves a true waste of resources—then policy makers should adopt other

<sup>37</sup> The first round settlement offers are the same:  $S_1 = \hat{x} - c_p - k_p = \tilde{x} - c_p - k_p$ .

<sup>38</sup> If the defendant was overdeterred to begin with, however, then MFNs would be a social improvement.

<sup>39</sup> There could be a tradeoff: MFNs reduce care levels but typically save delay and litigation costs. Note that these concerns may be raised about the desirability of settlement more generally. See Shavell (1997) and Spier (1997) for related discussions.

policy instruments to punish careless defendants. Policies might include a tax on defendants when cases are filed, or the prudent adoption of damage multipliers.<sup>40</sup>

*Proposition 7.* Suppose that  $f'(\tilde{x}) \geq 0$ . Then there exists a damage multiplier,  $m^* > 1$ , that when used in conjunction with MFNs gives the defendant the same total payments as in the no-commitment case but with lower total costs of litigation and delay.

The case where the plaintiffs' damages are uniformly distributed illustrates the desirability of allowing MFNs and restoring incentives for care with a damage multiplier. Recall that, in this special case, MFNs do not affect plaintiff welfare or the litigation rate, although the delay costs are certainly lower. Consequently, the defendant's total payments (payments to the plaintiffs *plus* the litigation and delay costs) are lower with MFNs. Now suppose that a damage multiplier is chosen so that the defendant's total payments are exactly equal to his total payments when MFNs are prohibited. Since the total payments to the plaintiffs rise when a damage multiplier is applied, the plaintiffs are better off in the MFN regime (with the damage multiplier) than they are in the regime where MFNs are prohibited (and there is no multiplier). It follows that the total litigation and delay costs are lower in the MFN regime than they would be in a regime with no MFNs.

## 4. Extensions

■ **Multiple rounds of settlement offers.** For tractability, the model assumed that there were only two opportunities for settlement: "early" and "late." A more realistic framework would allow a sequence of  $T$  settlement offers before trial,  $S_1, \dots, S_T$ , with associated positive delay costs  $c_{i1}, \dots, c_{iT}$ ,  $i = p, d$ . As before, the defendant would like to commit to a sequence where all settlement takes place in the very first round and commitment is readily achieved through an MFN. While a full characterization of the no-commitment equilibrium of this game is beyond the scope of this article, the insights of Spier (1992) are relevant here. With no costs of delay, the defendant would wait until the very last round to make an offer: all settlement would take place on the courthouse steps. With positive costs of delay, there would be settlement in each round and the defendant would benefit from commitment. A closed-form solution is obtained when the plaintiffs are uniformly distributed. In this case, the plaintiffs' welfare and the litigation rate are the same with and without MFNs, confirming the results obtained earlier.

□ **Uncertainty about the aggregate distribution of plaintiff types.** An important assumption in the model was that the aggregate distribution of plaintiff types was known. In practice, there is often uncertainty about this aggregate distribution. This implies that the defendant may be able to learn about the aggregate distribution after observing the volume of settlement in the first round and adjust his settlement offer accordingly.<sup>41</sup> In this environment, MFNs can facilitate learning by the defendant and provide the flexibility to raise his offer to the optimal level in the future.<sup>42</sup>

Formally, suppose that there are  $N$  states of nature with densities  $f_i(x)$  and cumulative distribution functions  $F_i(x)$ ,  $i = 1, \dots, N$ , and assume that  $F_i(x) = F_j(x)$  only if  $i = j$ .<sup>43</sup> There are  $T$  offers before trial, and assume that the costs of delay between rounds,  $c_{pt}$  and  $c_{dt}$ , are negligible (although trial is costly). If the defendant knew the state of nature with certainty, he would commit to offer  $S_1 = \hat{x}_i - k_p$ , where  $\hat{x}_i$  is defined in Proposition 1. In other words, the

<sup>40</sup> An additional lump-sum payment by defendants at trial would have no effect on the asymmetric information here and so would not change the litigation costs or the plaintiffs' welfare. Damage multipliers will, in general, change the bargaining game in a meaningful way.

<sup>41</sup> See the analysis of the durable-goods monopolist in Butz (1990).

<sup>42</sup> Learning could also involve new information arising exogenously over time. A previous version of this manuscript, dated October 5, 2001, and available on SSRN, shows that the main results of the article are unchanged when the defendant's liability is established at an interim stage.

<sup>43</sup> One can make a similar argument for an infinite number of states indexed by  $\theta$ , where  $F_i(x, \theta) = F_j(x, \theta')$  if and only if  $\theta = \theta'$ .

defendant would “price discriminate” among the different states of nature. With the help of an MFN, the defendant can implement this price-discrimination strategy when there is uncertainty as well.

To see this, suppose the defendant offers to settle in round 1 for  $S_1 = x^* - k_p$  with an MFN, where  $x^* = \min\{\hat{x}_i\}_{i=1}^N$ . The MFN makes it a dominant strategy for plaintiffs with damages  $x < x^*$  to accept this offer.<sup>44</sup> It is also a dominant strategy for plaintiffs with damages  $x > x^*$  to reject it.<sup>45</sup> The volume of plaintiffs who accept the first offer,  $F_i(x^*)$ , perfectly reveals the true state of nature to the defendant.<sup>46</sup> Modifying (A6) from the Appendix, the defendant will choose  $S_2$  (and the corresponding cutoff  $x'$ ) to minimize his continuation payoff:

$$F_i(x')(x' - k_p) + \int_{x'}^{\infty} (x + k_d)f_i(x)dx - (x^* - k_p)F_i(x^*). \quad (13)$$

Since  $\hat{x}_i \geq x^*$ , the defendant offers to settle for  $S_2 = \hat{x}_i - k_p \geq S_1$  in round 2. There will be no settlement after round 2; the MFN serves to commit the defendant not to raise his offer further once the optimal offer has been reached.

This simple illustration is interesting for a couple of other reasons. First, it is consistent with the observation of MFN payouts. When the tobacco industry settled on favorable terms with Minnesota in 1998, for example, previous MFN clauses were triggered and payments to Florida, Texas, and Mississippi were increased \$1.8 billion, \$2.3 billion, and \$.6 billion, respectively.<sup>47</sup> Second, this example addresses the earlier point that there are other ways for the defendant to achieve commitment—a large lump-sum payment to a third party, for example. Here, the defendant does not want to fully commit. Instead, he wants to retain the flexibility to raise the offer in the future in response to the information learned in round 1.

□ **A small number of plaintiffs.** The previous sections assumed that plaintiffs were drawn from a continuum, an assumption that is clearly unrealistic. Even large class action lawsuits involve a finite number of individuals, and many lawsuits in which MFNs are adopted involve small numbers of plaintiffs, such as the MP3.com case. This subsection shows that MFNs can be useful in cases involving small numbers of similar plaintiffs, although they will not always implement the full-commitment outcome.

To start, suppose that there is a single plaintiff whose damages are drawn from the probability density function  $f(x)$ , and these damages are private information. (In the previous sections,  $f(x)$  described an entire *population* of plaintiffs.) The defendant would like to commit to the very same sequence of settlement offers specified in Proposition 1, but cannot do so. The reason why MFNs worked so beautifully earlier was that they served as a valuable strategic commitment not to raise the offer in the future. With a single plaintiff, MFNs are an ineffective commitment device.

It is not hard to see that MFNs can be effective with a finite number of plaintiffs. Here is a stark example. Suppose there are  $N$  plaintiffs in total. Only one of the plaintiffs has private information: his damages are drawn from distribution  $f(x)$ . The remaining  $N - 1$  plaintiffs all have damages of  $\hat{x}$ , the full-commitment cutoff, and this is common knowledge. The defendant would, of course, like to commit to offer  $S_1 = \hat{x} - c_p - k_p$  to each of the  $N$  plaintiffs. MFNs achieve the full-commitment outcome in this special case: After  $N - 1$  plaintiffs have accepted

<sup>44</sup> A plaintiff with  $x < x^*$  is strictly better off accepting if the defendant's offer were to fall in the future, and is no worse off accepting if the defendant's offer rises. Note that no plaintiff would be willing to accept  $S_1 = x^* - k_p$  without an MFN. The plaintiffs would all anticipate, correctly, that the offers would only rise with time. So learning would not be achieved.

<sup>45</sup> With positive delay costs it would no longer be a dominant strategy.

<sup>46</sup> The result that the volume of plaintiffs who accept the offer perfectly reveals the state of nature to the defendant is, of course, driven by the assumption that there is an infinite number of plaintiffs. With a finite number, the defendant would update his beliefs but would not learn the state perfectly.

<sup>47</sup> See “States Suffer Tobacco Settlement Envy,” *The National Law Journal*, May 25, 1998, p. A6, and “Tobacco Companies Sweeten Florida's Settlement Pot,” *Tobacco Industry Litigation Reporter*, September 25, 1998, p. 6.

$S_1$ , it becomes prohibitively expensive for the defendant to raise the offer to the privately informed plaintiff (assuming  $N$  is not too small).

The analysis is much more complicated, of course, when the  $N$  plaintiffs' damages are independently drawn from the probability density function  $f(x)$ . MFNs certainly do not implement the full-commitment outcome in this case. If all plaintiffs happened to have damages above the threshold,  $x > \hat{x}$ , then no plaintiff would accept  $S_1 = \hat{x} - c_p - k_p$  and then nothing would deter the defendant from raising his offer in the second round. Indeed, the defendant's best second-round offer,  $S_2$ , would depend critically on the number of plaintiffs who accepted in the first round. This was not true in the earlier analysis. With a continuum of plaintiffs, the volume of plaintiffs who accepted  $S_1$  is deterministic (by the law of large numbers). Although the defendant cannot implement the full-commitment outcome, MFNs may be valuable in creating partial commitment here. A formal analysis of this important case is beyond the scope of this article and remains an important topic for future research.

□ **Multiple groups of plaintiffs.** The main analysis assumed that plaintiffs all looked the same to the defendant. Consequently, the defendant was not able to directly price discriminate among the different plaintiffs in the first round—the same settlement terms were offered to each plaintiff. Suppose instead that the defendant can observe the individual case characteristics that identify each plaintiff as belonging to one group or another. The damages of the plaintiffs in Group A are distributed  $f_A(x)$  and the damages of the Group B plaintiffs are distributed  $f_B(x)$ .<sup>48</sup>

At first glance, it would appear that the defendant would not want an MFN to apply across the two groups. Since the defendant can tell these two groups apart, he would like to price discriminate and commit to a separate sequence of settlement offers for each group, a sequence that is fine-tuned to the particular characteristics of each group.<sup>49</sup> This strategy can be achieved through separate MFNs, one for each group. The first-round offer to Group A might read as follows:

We offer to settle with you for  $S_1 = \hat{x} - c_p - k_p$ . If we offer to settle with anyone else from Group A in the future for more than this, then early settlers will receive the higher amount as well. This only applies to Group A, however. We retain full discretion to settle with plaintiffs from Group B on better terms.<sup>50</sup>

The problem with this strategy is that distinguishing between multiple groups of plaintiffs is often tricky. First, even if the defendant could privately distinguish between a Group A plaintiff and a Group B plaintiff, this information may be difficult or impossible to contract upon. It is common for exceptions to be broadly specified as litigants who are not “similarly situated” with the early settlers.<sup>51</sup> Consequently, the litigants may try to take advantage of the “similarly situated” clause *ex post*. First, Group A plaintiffs who reject the settlement offer in the first round may argue that they are not in fact members of Group A and should therefore be exempt.<sup>52</sup> Second, suppose the defendant gives members of Group B a better deal than Group A. If Group A can convince a court that the Group B plaintiffs are similar to them, then they would be entitled to the MFN payout!<sup>53</sup> Third, the defendant may have an incentive to misrepresent the truth *ex post* as well. At

<sup>48</sup> More generally, these two groups may have different litigation costs or other distinguishing features.

<sup>49</sup> The desire to distinguish between groups appears in many other contexts as well. In natural gas markets, for example, MFNs typically apply within the boundaries of each field and not across fields. See Butz (1986).

<sup>50</sup> The recent vitamins antitrust settlement agreement, for example, explicitly excluded vitamin B9 (folic acid) from the agreement, allowing the defendant to settle in the future on better terms on these issues. See section 22(c) of the Vitamins Antitrust Litigation Settlement Agreement. (United States District Court for the District of Columbia).

<sup>51</sup> In the Corrugated Container price-fixing case, the MFN clause was held not to apply to defendants who were financially constrained (so-called hardship settlements) or to defendants who were not indicted in related criminal cases (who were presumably less culpable than their indicted counterparts). See *In re Corrugated Container Litigation*, 1983-1 Trade Cas. (CCH) ¶65, 451 (1983).

<sup>52</sup> This situation frequently arises in practice. In the vitamins case, for example, plaintiff Nutra-Blend argued that it was in a unique position as a blender and suffered greater damages than the other class members and therefore should be permitted to receive more in settlement. In *re Vitamins Antitrust Litigation*, This Document Relates to: All Actions, 2000-1 Trade Cas. (D.D.C. CCH) ¶72, 862 (2000).

<sup>53</sup> In the Corrugated Container case, the defendants who settled early argued that the lesser payments made later by insolvent defendants (the “hardship settlements” mentioned earlier) should trigger the MFN provision and entitle them to a refund. The court disagreed. In *re Corrugated Container Litigation*, 1983-1 Trade Cas. (CCH) ¶65, 451 (1983).

the beginning of round 2, the defendant may try to argue that Group A plaintiffs who rejected the first-round offer are actually members of Group B. In that way, the defendant could get around his earlier commitment not to raise the offer and extract some additional value in settlement.

In summary, the attempt to distinguish between groups can lead to opportunistic behavior. If anticipated in the first round, opportunism will tend to undermine the positive effects of MFNs identified earlier. In addition, these rent-seeking activities may be costly in and of themselves, as is evidenced by the follow-on litigation described in the last paragraph. For these reasons, MFNs that attempt to distinguish between different groups of plaintiffs using broad language may destroy rather than create value for the defendant.

Finally, MFNs may be useful across groups (in addition to within them) when the groups have small numbers of plaintiffs.<sup>54</sup> As discussed earlier, MFNs are completely ineffective when a group has exactly one plaintiff and are less than fully effective with groups of finite size. If the second group is not too different from the first, then the defendant may want to pool the two groups together and have the MFN apply across the two groups. This may be seen most obviously when the first group is very large (so the full-commitment outcome is essentially achieved) and the second group has exactly one plaintiff. If the MFN with the first group applies to the second group, then it will commit the defendant not to raise the offer to this lone plaintiff. If the lone plaintiff is not too different from the earlier plaintiffs, then this is a valuable strategy.

□ **Plaintiff bargaining power.** The previous analysis gave the defendant the power to make a sequence of offers to the plaintiffs. What would happen if the defendant were not so powerful and negotiated the settlement terms instead? Consider a situation where the plaintiffs are jointly represented by an attorney who seeks to maximize their aggregate payoffs. The attorney representing the settlement class action and the defendant are both uninformed about the strengths of the individual claims, and they realize that the strongest claims will later “opt out” and pursue independent cases.<sup>55</sup> Since both the attorney and the defendant have every incentive to avoid the delay of the settlement opt-outs, they would find it in their mutual interest to include an MFN clause in the class settlement contract.

If instead of negotiating jointly through a common attorney the plaintiffs negotiated bilaterally, two important issues would arise. First, the formal analysis would need to incorporate signalling on the part of the informed plaintiffs (Reinganum and Wilde, 1986). Second, the issue of externalities would arise because an individual plaintiff’s decision to sign a contract with an MFN will typically harm other plaintiffs in the future. These externalities are another reason why the defendant may find it in his private interest to use MFNs, an issue that will be discussed further in the conclusion.

## 5. Conclusion

■ This article has argued that MFNs are privately valuable because they commit the defendant not to raise his settlement offers over time and therefore encourage early settlement. There may be other rationales for MFNs, too. Some commentators have argued that MFNs lead to better risk sharing. This may be true if MFNs lead to settlement that wouldn’t otherwise have occurred (since trials are risky). But it is not true if the case would otherwise have settled for a fixed amount. In this latter case, the defendant and plaintiffs face less risk if they settle for a fixed amount, and are worse off if the settlement will be adjusted over time.<sup>56</sup> Another rationale is that MFNs may serve to discourage future suits. Since plaintiffs would be forced to bring their cases all the way

---

<sup>54</sup> This argument would also apply if there were many plaintiffs represented by a small number of decision makers such as lawyers.

<sup>55</sup> In a settlement class action, the class is simultaneously certified for settlement and trial. The individual plaintiffs later have the opportunity to opt out if they do not like the settlement terms.

<sup>56</sup> Notice that these inefficiencies would be greater if MFNs applied to judgments at trial. See MacAvoy (1962) and Broadman and Montgomery (1983) for discussions of risk sharing and most-favored-customer clauses in natural gas contracts.

to trial, it may discourage them from bringing the case at all. This does beg the question of why this plaintiff would be able to extract a settlement offer to begin with, if his case is too weak to pursue.

It is also plausible that MFNs may allow an informed defendant to signal his type to the plaintiffs.<sup>57</sup> The defendant might argue: “I am not liable for your damages, so I will settle with you for a token amount. If it is ever determined that I am lying to you, and I settle with someone else for more, then you will receive the difference.” While this signalling story may have some appeal, it is not consistent with an important feature of most of the MFN clauses that we observe in practice. The clause is typically applied to settlement only, and not to awards made at trial. With this type of private information, a defendant could signal more credibly by making the MFN apply more broadly to awards if some cases proceed to trial.

An interesting exception to the practice of restricting MFNs to future settlements and not to trial outcomes is the antitrust settlement between Microsoft and the state of New Mexico, wherein New Mexico will receive the *judgment or settlement* received by the other states. Commentators are referring to this as a “free-rider” clause, implying that New Mexico will “reap the benefits from a battle that others will continue to fight.”<sup>58</sup> While New Mexico is surely free-riding, this by itself does not explain the use of the MFN clause. After all, the MFN puts New Mexico’s ultimate recovery at risk. Why not settle for a fixed amount instead? One answer may be that the MFN commits Microsoft to fighting harder in the future, a commitment that has good strategic effects and may, in the long run, reduce Microsoft’s overall liability. Another may be that officials in New Mexico are concerned about “looking bad” in the event that the other states receive more.

Finally, some commentators have argued that MFNs may serve to drive competitors out of business. In *Chicken Antitrust* (a price-fixing case), plaintiffs entered into settlement agreements with several chicken processors. Several financially weaker processors refused to settle at the same terms.<sup>59</sup> The court was particularly concerned here that the defendants who were not part of the early settlement would be financially ruined by litigation, and that defendants who had settled early would come to dominate the market in the future. The court struck down the MFN in the *Chicken* case, arguing that the clause had “predatory intent.”<sup>60</sup>

Perhaps the most plausible alternative explanation for MFNs is that they can be an effective bargaining tool when plaintiffs have some bargaining power, too.<sup>61</sup> Intuitively, an MFN commits the defendant to be tough in future negotiations, placing an upper bound on what a future plaintiff can extract in settlement. This may allow the defendant and early plaintiffs to capture a greater share of the future bargaining surplus. This tough commitment can backfire, however. The MFN destroys the settlement range when a future plaintiff turns out to have a very valuable case. This may be bad from a social welfare perspective. When committing to the MFN *ex ante*, the defendant and the early plaintiffs do not fully internalize the *ex post* cost of breakdowns because at least part of that cost will be borne by the future plaintiffs.<sup>62</sup> These, and the previously discussed motivations for MFNs in settlement agreements, remain fruitful areas for further research.

## Appendix

- Proofs of Lemmas 2 and 3 and Propositions 2, 3, and 7 follow.

*Proof of Lemma 2.* To start, I will prove that  $S_2 \geq S_1 + c_p$  in any continuation equilibrium following  $S_1$ . Suppose not:

---

<sup>57</sup> Levy (2000) shows that best-price guarantees can be an effective signal of quality when a monopolist is selling a durable good, and that these guarantees may increase social welfare.

<sup>58</sup> See “New Mexico Cuts Deal with Microsoft,” *The Seattle Times*, July 13, 2001, p. A1.

<sup>59</sup> This rationale raises an important question: If a defendant would be financially ruined by going to trial, why doesn’t he settle on the same terms as the others?

<sup>60</sup> In re *Chicken Antitrust Litigation*, 560 F. Supp. 943 (Ga. 1979).

<sup>61</sup> These ideas are very related to the earlier work of Cooper and Fries (1991) and Neilsen and Winter (1994). They are also modelled in the earlier working paper version of this article, dated October 5, 2001, and available on SSRN and in Spier (2003). See also the signalling model of Daughety and Reinganum (2002b).

<sup>62</sup> The intuition is analogous to Aghion and Bolton (1987).



$S_2 < S_1 + c_p$ . Then all plaintiffs with  $x \leq x^*(S_1, S_2) = S_1 + c_p + k_p$  would accept  $S_1$  and all plaintiffs with  $x > x^*(S_1, S_2)$  would reject  $S_1$ . Facing a truncated distribution on the support  $(S_1 + c_p + k_p, \infty)$  in round 2, the defendant would make an offer that the lowest remaining types strictly prefer to accept:  $S_2 > S_1 + c_p$ . A contradiction.

*Case (i).*  $S_1 > \underline{x} - c_p - k_p$  implies  $M_1 > 0$ . If  $M_1 = 0$  instead, then  $S_2 = \underline{x} - k_p < S_1 + c_p$ , a contradiction (since I have shown  $S_2 \geq S_1 + c_p$ ). It must also be the case that  $S_2 = S_1 + c_p$ . If  $S_2 > S_1 + c_p$ , then no plaintiff would accept  $S_1$  ( $M_1 = 0$ , a contradiction). Notice that  $x^*(S_1, S_2) > \underline{x}$  and (suppressing the arguments)  $S_2 = x^* - k_p$ .

Now I will show that  $M_2 = (k_p + k_d)f(x^*)$  is a necessary condition for  $S_2 = x^* - k_p$  to be the optimal choice for the defendant. Let  $G(x)$  represent the cumulative distribution function and  $g(x)$  (which I assume is integrable) represent the density of the plaintiffs remaining in round 2. Note that  $g(x) = f(x)$  for all  $x > x^*$  and  $g(x) \leq f(x)$  for all  $x \leq x^*$ .

Suppose that instead of  $S_2 = x^* - k_p$ , the defendant raised the offer to  $x^* + \Delta - k_p$ , where  $\Delta > 0$ . Since plaintiffs with damages below  $x^* + \Delta$  would accept this offer and those with damages above would reject it, the defendant's continuation payments are

$$(x^* + \Delta - k_p)[G(x^*) + F(x^* + \Delta) - F(x^*)] + \int_{x^* + \Delta}^{\infty} (x + k_d)f(x)dx. \tag{A1}$$

Since such a deviation should not be profitable for any  $\Delta > 0$ , upon differentiation we must have  $G(x^*) \geq (k_p + k_d)f(x^*)$ . If this were not true, the defendant would prefer a positive  $\Delta$  to raise the settlement offer.

Now suppose the defendant lowers the offer to  $x^* - \Delta - k_p$ , where  $\Delta > 0$ . Plaintiffs with damages below  $x^* - \Delta$  accept this offer and those with damages above reject it and go to trial. The defendant's continuation payments are

$$(x^* - \Delta - k_p)G(x^* - \Delta) + \int_{x^* - \Delta}^{x^*} (x + k_d)g(x)dx + \int_{x^*}^{\infty} (x + k_d)f(x)dx. \tag{A2}$$

Since such a deviation should not be profitable for any  $\Delta > 0$ , upon differentiation we must have  $G(x^*) \leq (k_p + k_d)g(x^*)$ . If not, the defendant would prefer a positive  $\Delta$  to lower the settlement offer. So we have

$$(k_p + k_d)f(x^*) \leq G(x^*) \leq (k_p + k_d)g(x^*). \tag{A3}$$

$g(x^*) \leq f(x^*)$  implies  $(k_p + k_d)f(x^*) = G(x^*)$ . Since  $G(x^*)$  is the mass of plaintiffs with  $x \leq x^*$  remaining in the second round and all of these plaintiffs accept  $S_2 = x^* - k_p$ , the result is proven.

*Case (ii).* Suppose  $S_1 \leq \underline{x} - c_p - k_p$ . It must be the case that  $M_1 = 0$ . Suppose not:  $M_1 > 0$ . It follows that  $S_2 = S_1 + c_p \leq \underline{x} - k_p$  and so  $x^*(S_1, S_2) \leq \underline{x}$ . As before, suppose the defendant is thinking about raising the offer to  $x^* + \Delta - k_p$ . The derivative of expression (A1) with respect to  $\Delta$  is  $G(x^*) - (k_p + k_d)f(x^*)$ . Using the fact that  $G(x^*) < F(x^*)$  when  $M_1 > 0$  and using the fact that  $x^*(S_1, S_2) \leq \underline{x}$  and the definition of  $\underline{x}$ , we have

$$G(x^*) - (k_p + k_d)f(x^*) < F(x^*) - (k_p + k_d)f(x^*) < F(\underline{x}) - (k_p + k_d)f(\underline{x}) = 0. \tag{A4}$$

We conclude that the defendant would want to raise  $S_2$  in order to reduce his total payments, a contradiction. Since  $M_1 = 0$ , we are in the benchmark case described above, with  $S_2 = \underline{x} - k_p$  and  $M_2 = F(\underline{x})$ .<sup>63</sup> *Q.E.D.*

*Proof of Proposition 2.* The derivative of expression (6) is  $F(x^*) - (c_p + c_d + k_p + k_d)f(x^*) + (c_p + c_d)(k_p + k_d)f'(x^*)$ . When  $x^* = \underline{x}$ , then, using (5), the derivative simplifies to  $-(c_p + c_d)f(\underline{x}) + (c_p + c_d)(k_p + k_d)f'(\underline{x})$ . The following two facts establish that this expression is negative: (i) the monotone hazard rate assumption implies that  $f'(x)/f(x) < f(x)/F(x)$ ; (ii) using (5),  $f(\underline{x})/F(\underline{x}) = 1/(k_p + k_d)$ . We can conclude that the defendant's total payments are falling in  $x^*$  when  $x^* = \underline{x}$ , and therefore  $\bar{x} > \underline{x}$ . Note that  $\bar{x}$  may not be uniquely defined. The results stated later do not hinge on uniqueness. *Q.E.D.*

*Proof of Proposition 3.* Suppose that the defendant offers  $S_1 = \hat{x} - c_p - k_p$  together with an MFN provision. If the plaintiffs expect that  $S_2 < S_1$ , then plaintiffs with  $x \leq \hat{x}$  accept  $S_1$  and plaintiffs with  $x \in (\hat{x}, \infty)$  reject  $S_1$  and remain in round 2. The defendant will rationally choose *not to settle* in round 2 with any of these remaining plaintiffs. Suppose  $S_2 = x^* - k_p > \hat{x} - k_p > S_1$ . Plaintiffs with damages  $x \in (\hat{x}, x^*]$  accept  $S_2$ , and plaintiffs with damages above  $x^*$  reject  $S_2$  and go to trial. We may write the defendant's total continuation payments as

$$[F(x^*) - F(\hat{x})][S_2] + \int_{x^*}^{\infty} (x + k_d)f(x)dx + (S_2 - S_1)F(\hat{x}). \tag{A5}$$

<sup>63</sup> Intuitively, if a positive mass were removed from the lower part of the distribution of plaintiff types, then it would be the case that  $S_2 > \underline{x} - k_p$ .

The first term represents the settlement payments made to  $x \in (\hat{x}, x^*]$ ; the second term represents the total payments made at trial with the higher types; the third term represents the additional payments made to the early settlers under the MFN. Replacing  $S_2$  with  $x^* - k_p$  and rearranging terms gives the defendant's continuation payments:

$$F(x^*)(x^* - k_p) + \int_{x^*}^{\infty} (x + k_d)f(x)dx - S_1 F(\hat{x}). \quad (A6)$$

The derivative with respect to  $x^*$  is

$$F(x^*) - (k_p + k_d)f(x^*). \quad (A7)$$

From Proposition 1,  $\hat{x}$  was defined by  $F(\hat{x}) - (c_p + c_d + k_p + k_d)f(\hat{x}) = 0$ , and so  $F(\hat{x}) - (k_p + k_d)f(\hat{x}) > 0$ . Since  $F(x)/f(x)$  is increasing in  $x$  by assumption and  $x^* > \hat{x}$ , we have  $F(x^*) - (k_p + k_d)f(x^*) > 0$ . Therefore the defendant offers  $S_2 < \hat{x} - k_p$  and  $M_2 = 0$ .<sup>64</sup> This confirms the plaintiffs' expectations, and the full-commitment outcome is achieved. (This is the only outcome of the continuation game. If the plaintiffs expected  $S_2 > S_1$ , then  $x \leq x^{**} = S_2 - k_p$  would accept  $S_1$  in anticipation of a future MFN payout. An argument similar to the one above shows that the defendant would never offer to settle for  $S_2 > S_1$ .) *Q.E.D.*

*Proof of Lemma 3.* Comparing equation (4), which defines the cutoff with MFNs,  $\hat{x}$ , to equation (8), which defines the cutoff without MFNs,  $\bar{x}$ , we see that equation (8) has an additional term:  $(c_p + c_d)(k_p + k_d)f'(\bar{x})$ . If  $f'(\bar{x}) = 0$  (as for the uniform distribution), then clearly  $\hat{x} = \bar{x}$ . If  $\hat{x} < \bar{x}$ , then combining (4) and (8) shows  $(c_p + c_d)(k_p + k_d)f'(\bar{x}) = 0$ . Taken together,  $f'(\bar{x}) = 0$  if and only if  $\hat{x} = \bar{x}$ . If  $f'(\bar{x}) > 0$ , then  $(c_p + c_d)(k_p + k_d)f'(\bar{x}) > 0$  in equation (8), and so it must be the case that  $F(\bar{x}) - (c_p + c_d + k_p + k_d)f(\bar{x}) < 0$ . Using the assumption that  $F(x)/f(x)$  is increasing in  $x$ , we see that  $\hat{x} > \bar{x}$ . Now suppose that  $\hat{x} > \bar{x}$ . Since  $F(x)/f(x)$  is increasing in  $x$ , we have  $F(\hat{x}) - (c_p + c_d + k_p + k_d)f(\hat{x}) = 0 > F(\bar{x}) - (c_p + c_d + k_p + k_d)f(\bar{x})$ . It follows from equation (12) that  $(c_p + c_d)(k_p + k_d)f'(\bar{x}) > 0$ . Taken together,  $f'(\bar{x}) > 0$  if and only if  $\hat{x} > \bar{x}$ . The proof that  $f'(\bar{x}) < 0$  if and only if  $\hat{x} < \bar{x}$  is analogous, and we are done. *Q.E.D.*

*Proof of Proposition 7.* Let the damage multiplier be  $m$ . Let  $\hat{V}(m)$  represent all of the payments made by the defendant when MFNs are used.  $\hat{V}(m)$  equals the defendant's direct payments to the plaintiffs plus the defendant's delay and litigation costs. Let  $\hat{\Pi}(m)$  be the aggregate payoffs received by the plaintiffs.  $\hat{\Pi}(m)$  equals the settlements and judgments received from the defendant minus the plaintiffs' relevant delay and litigation costs. We may write  $\hat{V}(m) = \hat{\Pi}(m) + \hat{\Omega}(m)$ , where  $\hat{\Omega}(m)$  are the total costs of litigation and delay (as defined earlier). Similarly, we let  $\tilde{V} = \tilde{\Pi} + \tilde{\Omega}$  be the defendant's total payments when MFNs are prohibited. Since  $f'(\bar{x}) \geq 0$  by assumption, we know from Proposition 5 that  $\hat{\Pi}(1) \geq \tilde{\Pi}$  and from Proposition 6 that  $\hat{\Omega}(1) < \tilde{\Omega}$ . Since  $\hat{V}(1) < \tilde{V}$  and  $\hat{V}'(m) > 0$  (easily shown with the envelope theorem), we know that there exists a damage multiplier where  $\hat{V}(m^*) = \tilde{V}$ , or  $\hat{\Pi}(m^*) + \hat{\Omega}(m^*) = \tilde{\Pi} + \tilde{\Omega}$ . Since  $\hat{\Pi}'(m) > 0$  (the plaintiffs are better off with higher damage multipliers) we know that  $\hat{\Pi}(m^*) > \hat{\Pi}(1) > \tilde{\Pi}$ . Therefore  $\hat{\Omega}(m^*) < \tilde{\Omega}$ , and we are done. *Q.E.D.*

## References

- AGHION, P. AND BOLTON, P. "Contracts as a Barrier to Entry." *American Economic Review*, Vol. 77 (1987), pp. 388–401.
- BEBCHUK, L.A. "Litigation and Settlement Under Imperfect Information." *RAND Journal of Economics*, Vol. 15 (1984), pp. 404–415.
- BESANKO, D. AND LYON, T.P. "Equilibrium Incentives for Most-Favored-Customer Clauses in an Oligopolistic Industry." *International Journal of Industrial Organization*, Vol. 11 (1993), pp. 347–367.
- BROADMAN, H.G. AND MONTGOMERY, W.D. *Natural Gas Markets After Deregulation*. Washington, D.C.: Resources for the Future, 1983.
- BULOW, J.I. "Durable-Good Monopolists." *Journal of Political Economy*, Vol. 90 (1982), pp. 314–332.
- BUTZ, D.A. "Long-Term Contracting in Field Markets for Natural Gas: A 'New' Perspective on the Most-Favored Nation Provision." Ph.D. dissertation, Department of Economics, Northwestern University, 1986.
- . "Durable-Good Monopoly and Best-Price Provisions." *American Economic Review*, Vol. 80 (1990), pp. 1062–1076.
- . "Most-Favored Treatment Provisions as Nondiscrimination Guarantees." *International Journal of the Economics of Business*, Vol. 2 (1995), pp. 65–86.
- CHE, Y.-K. "Equilibrium Formation of Class Action Suits." *Journal of Public Economics*, Vol. 62 (1996), pp. 339–361.
- . "The Economics of Collective Negotiations in Pretrial Bargaining." *International Economic Review*, Vol. 43 (2002), pp. 549–576.
- AND YI, J.G. "The Role of Precedents in Repeated Litigation." *Journal of Law, Economics and Organization*, Vol. 9 (1993), pp. 399–424.

<sup>64</sup> Notice that this strategy is not unique (although the properties in the lemma always hold) because the defendant could offer to settle for  $S_2 \in (S_1, \hat{x} - k_p)$ . Since none of the remaining plaintiffs would accept this offer, the MFN clause would not obligate the defendant to pay additional money to the early settlers. But offering  $S_2 > S_1$  is weakly dominated by  $S_2 \leq S_1$ .

- COASE, R.H. "Durability and Monopoly." *Journal of Law and Economics*, Vol. 15 (1972), pp. 143–149.
- COOPER, T.E. "Most-Favored-Customer Pricing and Tacit Collusion." *RAND Journal of Economics*, Vol. 17 (1986), pp. 377–388.
- AND FRIES, T.L. "The Most-Favored-Nation Pricing Policy and Negotiated Prices." *International Journal of Industrial Organization*, Vol. 9 (1991), pp. 209–223.
- CROCKER, K.J. AND LYON, T.P. "What Do Facilitating Practices Facilitate? An Empirical Investigation of Most-Favored-Nation Clauses in Natural Gas Contracts." *Journal of Law and Economics*, Vol. 37 (1994), pp. 297–322.
- DAUGHETY, A.F. "Settlement." In B. Bouckaert and G. de Geest, eds., *Encyclopedia of Law and Economics*, Vol. 5. Northampton, Mass. Edward Elgar, 2000.
- AND REINGANUM, J.F. "Hush Money." *RAND Journal of Economics*, Vol. 30 (1999), pp. 661–678.
- AND ———. "Information Externalities in Settlement Bargaining: Confidentiality and Correlated Culpability." *RAND Journal of Economics*, Vol. 33 (2002a), pp. 587–604.
- AND ———. "Exploiting Future Settlements: A Signaling Model of Most-Favored-Nation Clauses in Settlement Bargaining." Mimeo, Department of Economics, Vanderbilt University, 2002b.
- DEGRABA, P. AND POSTELWAITE, A. "Exclusivity Clauses and Best Price Policies in Input Markets." *Journal of Economics and Management Strategy*, Vol. 1 (1992), pp. 423–454.
- HAY, B. AND SPIER, K.E. "Settlement of Litigation." In P. Newman, ed., *The New Palgrave Dictionary of Economics and the Law*. New York: Macmillan Reference, 1998.
- KORNHAUSER, L.A. AND REVESZ, R.L. "Multidefendant Settlements: The Impact of Joint and Several Liability." *Journal of Legal Studies*, Vol. 23 (1994a), pp. 41–76.
- AND ———. "Multidefendant Settlements Under Joint and Several Liability: The Problem of Insolvency." *Journal of Legal Studies*, Vol. 23 (1994b), pp. 517–542.
- LEVY, S.M. "Best Price Guarantees as a Quality Signal." Unpublished mimeo, 2002. Available from author at Slevy@nathaninc.com.
- LYON, T.P. "Most Favored Customer Clauses." In Peter Newman, ed., *The New Palgrave Dictionary of Economics and the Law*. New York: Macmillan Reference, 1998.
- MACAVOY, P.W. *Price Formation in Natural Gas Fields: A Study of Competition, Monopsony, and Regulation*. New Haven, Conn.: Yale University Press, 1962.
- MARX, L.M. AND SHAFER, G. "Opportunism and Nondiscrimination Clauses." Mimeo, University of Rochester, 2000.
- MCAFEE, R.P. AND SCHWARTZ, M. "Opportunism in Multilateral Vertical Contracting: Nondiscrimination, Exclusivity, and Uniformity." *American Economic Review*, Vol. 84 (1994), pp. 210–230.
- MILLER, G. "Class Actions." In P. Newman, ed., *The New Palgrave Dictionary of Economics and the Law*. New York: Macmillan Reference, 1998.
- NALEBUFF, B. "Credible Pretrial Negotiation." *RAND Journal of Economics*, Vol. 18 (1987), pp. 198–210.
- NEILON, W.S. AND WINTER, H. "Enhancing Bargaining Power with Most-Favored-Customer Pricing." *Economics Letters*, Vol. 44 (1994), pp. 61–66.
- P'NG, I.P.L. "Most-Favored-Customer Protection Versus Price Discrimination over Time." *Journal of Political Economy*, Vol. 99 (1991), pp. 1010–1028.
- REINGANUM, J.F. AND WILDE, L.L. "Settlement, Litigation, and the Allocation of Litigation Costs." *RAND Journal of Economics*, Vol. 17 (1986), pp. 557–566.
- SALOP, S.C. "Practices That (Credibly) Facilitate Oligopoly Coordination." In J.E. Stiglitz and G.F. Matthews, eds., *New Developments in the Analysis of Market Structure*. Cambridge, Mass.: MIT Press, 1986.
- SCHNITZER, M. "Dynamic Duopoly with Best-Price Clauses." *RAND Journal of Economics*, Vol. 25 (1994), pp. 186–196.
- SCOTT MORTON, F. "The Strategic Response by Pharmaceutical Firms to the Medicaid Most-Favored-Customer Rules." *RAND Journal of Economics*, Vol. 28 (1997), pp. 269–290.
- SHAVELL, S. "The Fundamental Divergence Between the Private and the Social Motive to Use the Legal System." *Journal of Legal Studies*, Vol. 26 (1997), pp. 575–612.
- SPIER, K.E. "The Dynamics of Pretrial Negotiation." *Review of Economic Studies*, Vol. 59 (1992), pp. 93–108.
- . "Pretrial Bargaining and the Design of Fee-Shifting Rules." *RAND Journal of Economics*, Vol. 25 (1994), pp. 197–214.
- . "A Note on the Divergence between the Private and the Social Motive to Settle Under a Negligence Rule." *Journal of Legal Studies*, Vol. 26 (1997), pp. 613–621.
- . "Settlement with Multiple Plaintiffs: The Role of Insolvency." *Journal of Law, Economics and Organization*, Vol. 18 (2002), pp. 295–323.
- . "'Tied to the Mast': Most-Favored-Nation Clauses in Settlement Contracts." *Journal of Legal Studies*, forthcoming.
- AND SYKES, A.O. "Capital Structure, Priority Rules, and the Settlement of Civil Claims." *International Review of Law and Economics*, Vol. 18 (1998), pp. 187–200.
- STOKEY, N.L. "Rational Expectations and Durable Goods Pricing." *Bell Journal of Economics*, Vol. 12 (1981), pp. 112–128.