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

Kathryn E. Spier

USC Center for Law, Economics & Organization
Research Paper No. C02-12

CENTER FOR LAW, ECONOMICS AND ORGANIZATION RESEARCH PAPER SERIES

Sponsored by the John M. Olin Foundation

University of Southern California Law School
Los Angeles, CA 90089-0071

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

Kathryn E. Spier*

October 5, 2001

Many settlement agreements in lawsuits involving either multiple plaintiffs or multiple defendants include so-called "most-favored-nation" clauses. If a defendant facing multiple claims, for example, settles with some plaintiffs early and settles with additional plaintiffs later for a greater amount, then the early settlers will receive the more favorable terms as well. These MFN provisions have been prominent in the recent MP3.com case, as well as tobacco litigation, class actions, and many antitrust lawsuits. This paper considers a defendant who is facing a large group of heterogeneous plaintiffs. Each plaintiff has private information about the (expected) award that he or she will receive should the case go to trial. MFN clauses are valuable because they commit the defendant not to raise his offer over time. This has two important effects. First, holding overall settlement rate fixed, MFNs encourage earlier settlement. Second, depending upon the distribution of plaintiff types, MFNs can either increase or decrease the overall settlement rate. Social welfare implications are discussed, and alternative theories, including the strategic use of MFNs to extract value from future plaintiffs, are explored.

---

* k-spier@kellogg.nwu.edu. Kellogg School of Management and National Bureau of Economic Research. Thanks to David Butz, Jim Dana, Andrew Daughety, Jennifer Reinganum, the Zodiac Room Group and Scott Stern for helpful discussions, and to Peter Dicola for the MP3.com example. All errors, of course, remain my own.

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.[1] 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.[2] Since the $50 million was a "judgment" rather than a "settlement" the MFN clause was not triggered.[3]

This paper argues that MFNs serve an important role in settlement with multiple plaintiffs and, in a parallel argument, with multiple defendants.[4] In a nutshell, MFNs commit the

---

[1] 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.

[2] Anna Wilde Mathews and Colleen DeBaise, "MP3.com Deal Ends Lawsuit on Copyrights," *The Wall Street Journal*, November 11, 2000. After liability was determined, Judge Rakoff ruled that the damages would be $25,000 for each Universal CD that was digitally copied by MP3.com. A second trial was scheduled to determine was 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.

[3] A salient feature of most MFN clauses is that they apply to settlement payments only, and not to awards made at trial. Not surprisingly, the four settling record companies challenged the $50 Million 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." See "Record Labels Fuming over Universal-MP3.com Ruling," *Daily News*, November 17, 2000.

[4] There is a literature on settlement of litigation with multiple litigants, although it does not discuss MFNs or the issues discussed here. Miller (1998) gives a survey of the economics of

defendant to be tough in future negotiations. Once the defendant has agreed to an MFN with one group of plaintiffs, it becomes very expensive for the defendant to settle on better terms with other plaintiffs. This paper highlights two reasons why the defendant may want to commit to be tough. First, MFNs mitigate the problem of asymmetric information when the plaintiffs are privately informed about the strengths of their cases. Second, MFNs are an effective mechanism for extracting value from future plaintiffs.

Formally, suppose that many plaintiffs are suing the same defendant. Each plaintiff has private information about what will happen if his or her case goes to trial. It is well known that settlement negotiations may break down under these circumstances: plaintiffs who know that they have very strong cases will reject the defendant's offer to settle and will seek compensation at trial instead.[5] There will also be delay in settlement: many plaintiffs with weaker cases will settle on the courthouse steps.[6] These plaintiffs reject the defendant's early offers to settle because they anticipate, correctly, that the defendant's offers will rise over time. MFNs induce these "11th hour" plaintiffs to accept early settlement offers instead.

This is an important effect. Since delay is inefficient -- protracted litigation involves time, energy, and legal costs -- MFNs can enhance both private and social welfare. Although MFNs unambiguously lead cases to settle earlier, the overall effect on the settlement rate is not

---

class action litigation, highlighting the conflict of interest between the (often self-appointed) attorney and the dispersed clients he represents. Che (1996, 1999) looks at the incentive for plaintiffs with private information to consolidate their claims. Che and Yi (1993) consider the role of precedent in litigation and Daughety and Reinganum (1999 and 2001) consider settlement negotiations when the defendant may want to keep information about the lawsuit a secret from future plaintiffs. Kornhauser and Revesz (1994a) and (1994b) look at multiple defendant lawsuits under joint and several liability, focusing on the externalities in settlement decisions, and Spier (2000) looks at similar externalities in lawsuits involving multiple plaintiffs and a potentially insolvent defendant.

[5] Hay and Spier (1998) and Daughety (2000) give recent surveys of the settlement literature.

[6] Spier (1992) shows why there are important deadline effects in pretrial bargaining.

as clear. When the plaintiffs' damages are uniformly distributed the overall settlement rate is unchanged by MFNs. More generally, however, we will see that the overall settlement rate falls if and only if the density function representing the plaintiffs' types is decreasing in the relevant range. A reduction in the overall settlement rate could potentially outweigh the cost savings from early settlement. In this case, MFNs could reduce social welfare.

The defendant may also want to adopt MFNs to capture a greater share of the future bargaining surplus. The mechanism is straightforward: an MFN commits the defendant to be tough in future negotiations, placing an upper bound on what a future plaintiff can extract in settlement. The defendant's tough commitment can backfire, however. The MFN destroys the settlement range when the 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 does not fully internalize the *ex post* cost of breakdowns since at least part of that cost will be borne by the future plaintiffs.

MFNs are quite common in settlement agreements. When the tobacco industry settled on favorable terms with Minnesota in 1998, for example, previous MFN clauses were triggered and payments to Florida,[7] Texas and Mississippi were increased $1.8 billion, $2.3 Billion, and $.6 billion, respectively.[8] They are also commonly used when a plaintiff is suing multiple

---

[7] Florida settlement, for example, reads: "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 on terms more favorable to such governmental plaintiff than the terms of this Settlement Agreement (after due consideration of relevant differences in population or other appropriate factors), the terms of this Settlement Agreement will be revised so that the State of Florida will obtain treatment at least as relatively favorable as any such non-federal governmental entity." The settlements with Texas and Mississippi were similar. See the full text at http://www.library.ucsf.edu/tobacco/litigation/fl/flsettle.html.

[8] See "States Suffer Tobacco Settlement Envy," *The National Law Journal*, May 25, 1998. The August 25, 1997. "Tobacco Companies Sweeten Florida's Settlement Pot," *Tobacco Industry Litigation Reporter*, September 25, 1998.

defendants, as in the 1996 price fixing case against steel pail manufacturers.[9] MFN clauses have also been used in class action litigation. In 1999, more than a dozen of the largest drug companies admitted to participating in a global cartel that fixed the prices of vitamins. The plaintiffs in this case included hundreds of small food and animal feed companies, in addition to large manufacturers. "Settlement negotiations were complicated by concerns that many of the largest companies in the case, including Tyson Foods Inc., Quaker Oats Co., Kellogg Co. And Cargill, Inc., would opt out of the settlement and file an individual lawsuit of their own."[10] The $1.2 billion settlement contract included a most-favored nation clause. While plaintiffs were free to subsequently opt-out of the class and file individual lawsuits, opt-outs could not receive more money unless the main class was paid the difference as well.[11]

A frequently voiced concern with MFNs is that they destroy settlement opportunities. "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."[12] In the vitamin price-fixing case, the court considered arguments both for and against MFN clauses before approving the class action settlement.[13] Some class members were

---

[9] See 85 F.3d 1198; 1996 U.S. App. LEXIS 13794; 1996 FED App. 0163P (6thCir.); 1996-1 Trade Cas. (CCH) P71,430. In this type of lawsuit, if a one defendant settles early and another defendant settles later at better terms, the early-settling defendant will receive a refund for the difference.

[10] Susan Warren and John R. Wilke, "Vitamin-Firm Accord to pay $1.17 Billion," *Wall Street Journal*, November 4, 1999.

[11] The MFN clause would expire after two years.

[12] In re: Chicken Antitrust Litigation, Civ. A. No. C74-2454A, 560 F. Supp. 943; 1979 U.S. Dist LEXIS 13430. I have reversed the identities of the plaintiffs and defendants in this quote to maintain consistency with the text.

[13] In re: Vitamins Antitrust Litigation, This Document Relates to: All Actions. Misc. No. 99-197 (TFH) United States Distric Court for the District of Columbia 1999 U.S. Dist. LEXIS 21963; 1999-2 Trade Cas. (CCH) P72, 726. "The court must protect the rights of the class to achieve resolution of their cases while at the same time preserving the rights of those who wish to opt-out of the class and negotiate their claims independently or pursue litigation."

unhappy with the terms of the settlement contract, believing it was too low. They argued that MFNs would deprive them of due process by forcing "them to either remain in the Class or to litigate this case to the end."[14] Class counsel argued that the clause was "essential to the settlement and that without it the parties would not have reached agreement." This debate is echoed in *The Manual for Complex Litigation* which states that "[MFN] clauses can provide an incentive for early settlement as well as an obstacle to later settlements."[15] The models presented here shed light on this important policy debate.

There is a large theoretical literature on the use of most-favored-customer clauses in supplier relationships.[16] The paper most related to ours is Butz (1990), where best-price provisions mitigate the time inconsistency problem that a monopolist faces when selling a durable good.[17] 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 forego early

---

[14] The court observed that no case law exists supporting the argument that MFNs constitute legal prejudice or deny due process. Kenneth L. Adams, who represented many of these plaintiffs, commented: "It [the MFN clause] was custom-designed, apparently because the class lawyers believe their settlement is too low. … Why are they afraid somebody else is going to get more?" See "$1.17B Vitamin Settlement: Now the Fight Begins," *National Law Journal*, November 15, 1999.

[15] See *The Manual for Complex Litigation* (Third) @ 23.23 at 182 (1995). While this most recent edition of the manual gives a balanced view of MFNs, earlier editions were far more negative about their use.

[16] See the survey by Lyon (1998). Most of this literature asks whether these clauses can be used collusively to soften price competition. 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).

[17] 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.

purchases and wait for the lower price.[18] Butz (1990) observed that the best-price provision is the mechanism by which the monopolist can commit not to lower the price later, and allows the monopolist to capture the full monopoly rents.

This paper has both important similarities and important differences. The obvious similarity with Butz (1990) is that both papers 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 our 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.[19] Here, by encouraging early rather than late settlement, MFNs serve to raise social welfare and, for certain parameter values, MFNs may increase the settlement rate (further increasing social welfare). Therefore the policy implications from the well-known durable goods monopoly problem do not apply in the litigation context.[20]

The next section of the paper lays out the basic the model where the population of plaintiffs is drawn from a continuous distribution. The third section presents comparative statics and evaluates the social welfare consequences of MFNs. The fourth section discusses several important extensions: a finite number of plaintiffs, multiple groups of plaintiffs, and the emergence of new information during negotiations. The fifth section presents the second reason

---

[18] These ideas are formalized by Stokey (1981) and Bulow (1982).

[19] Butz does say that the reader should not draw "sweeping policy conclusions" from this result. MFNs may 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)).

[20] The underlying pretrial bargaining game is different from bilateral trade in two respects. First, the game is one of common values. A plaintiff's private information about damages will affect the payoffs of both the plaintiff and defendant should the case go to trial. Second, a settlement simply involves a cash transfer between the litigants. We will see that, as a consequence, time does not screen among the plaintiff types here as it does in models of bilateral trade. For both of these reasons, the more familiar results from the durable goods monopoly fail to hold.

for MFN clauses in settlement: to enhance the defendant's bargaining position with later plaintiffs.[21] The final section offers concluding remarks and discusses avenues for future research.

2. The Model

A continuum of plaintiffs has been injured by a defendant. Their damages, x, are drawn from a differentiable density function *f(x)* which is positive on the support $[0, \infty)$. We normalize the total volume or mass of lawsuits to be 1, or $\int_0^\infty f(s)ds = 1$, and assume that $F(x) = \int_0^x f(s)ds < 1$.[22] We also assume that *F(x)/f(x)* is increasing in x.[23] The defendant knows the density, *f(x)*, and each plaintiff privately observes his or her own damages. That is, all of the plaintiffs initially look the same to the defendant.

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

[21] The idea is that MFNs can be used strategically to extract value from disorganized plaintiffs is related to Aghion and Bolton's (1987) model of entry deterrence. There, damages for breach of contract extract value from a future trading partner. Here, the MFN extracts value from future litigants.
[22] This guarantees that a positive mass of plaintiff types will reject any given finite settlement offer.
[23] This monotone hazard rate assumption assures a unique solution in the settlement game.

reject S_2 go to trial. 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.[24] We assume that c_p, c_d, k_p, and k_d are strictly positive, so settlement negotiations and trials are costly.[25]

An equilibrium is characterized by settlement offers for the defendant, S_1 and S_2, and settlement strategies for the plaintiffs where everyone is behaving optimally at each stage of the game.[26] Let M_1 and M_2 represent the volume or mass of plaintiffs who settle in the first and second rounds, respectively, and let M_T represent the mass of plaintiffs who ultimately go to trial.

This basic framework could be easily modified to represent the situation where a single plaintiff is negotiating with multiple defendants. All defendants appear the same to the plaintiff *ex ante*, but each has private information, distributed *f(x)*. In price fixing cases such as the Steel Pail case mentioned in the introduction, each defendant may have had private information about prior participation in the formal collusive agreement. A defendant who knows that evidence proving his culpability will emerge at trial has higher expected damages than a defendant who knows that no such evidence exists. Legal cases along these lines, including the Corrugated Cardboard case, will be discussed in Section 4.

This simple model can easily be generalized in several other ways without changing the results. First, trials could be risky in this framework. For example, let p be the probability that a

---

[24] The careful reader may notice that I have not given the plaintiffs the option to drop their cases before trial, an option that a plaintiff would want to exercise if $S_2 < 0$ and $x < k_p$. Some 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) for the equilibria in games where the plaintiff has the option to drop the case.

[25] We are assuming the American rule where each party bears its own litigation costs. These costs would include both direct legal costs as well as opportunity costs such as loss of managerial focus.

[26] We will restrict attention to pure strategies for the defendant.

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 our results are maintained.[27] Second, we may interpret the trials for the plaintiffs as taking place either jointly or separately.[28] If all plaintiffs are originally part of the same class action lawsuit, as in the vitamins case, then plaintiffs who reject the first round settlement in our model have "opted out" of the class. Finally, the model assumes that the litigants do not discount the future. This assumption is made to simplify the exposition of the paper. If we assumed instead that the defendants and plaintiffs discount time at the same rate, then the results would all go through unchanged (except that the damages and costs would be in present discounted terms). Other generalizations of the model that would require additional analysis, including a finite number of plaintiffs and the revelation of information over time, are discussed in Section 4.

Before characterizing the defendant's optimal offers, we will first construct the plaintiffs' optimal settlement strategies given the (anticipated) sequence of offers, S_1 and S_2, assuming that the defendant is not using MFNs. Suppose that after observing the first round offer, S_1, the plaintiffs expect that the second round offer will be S_2. Each plaintiff has three choices -- to accept S_1, to wait and accept S_2, or to reject both offers and go to trial -- and makes this decision to maximize compensation net of costs. When measured from the beginning of round 1, a plaintiff's payoff from accepting S_1 is simply S_1; his payoff from accepting S_2 is $S_2 - c_p$; and

---

27 Risk neutrality is important here, however. With risk aversion, uncertainty introduces inefficient risk bearing at trial.

28 Importantly, though, we 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, for class members are able to take advantage of economies of scale, while individual plaintiffs who opt out would forego these economies.

his payoff from going to trial is $x - c_p - k_p$ where x is the plaintiff's damages or "type." A plaintiff with damages x would choose to reject both offers and go to trial when the payoff from going to trial exceeds the payoff from settling, or $x - c_p - k_p > \max\{S_1, S_2 - c_p\}$. This condition defines a cutoff,

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

A plaintiff with damages above this cutoff, $x > x^*$, will reject both offers and go to trial. A plaintiff with damages below this cutoff, $x \le x^*$, will settle out of court for the better of the two offers and may mix between the two offers when $S_1 = S_2 - c_p$.

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 settle and those with strong cases 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.[29]

<u>Lemma 1:</u> Suppose the plaintiffs anticipate a sequence of offers, S_1 (without an MFN) and S_2. Define $x^* = \max\{S_1, S_2 - c_p\} + c_p + k_p$. Plaintiffs with damages $x \le x^*$ settle out of court and plaintiffs with damages $x > x^*$ go to trial.[30]

---

[29] For more discussion of the relationship between these two models, see Spier (1992).

[30] We adopt the tie-breaking assumption that when a plaintiff is indifferent between accepting an offer and going to trial, he accepts an offer.

2.1 Settlement With Commitment

We will now establish an important benchmark. If the defendant can *commit* to the sequence of offers, S_1 and S_2, then a positive mass of settlement will take place in the first round and all remaining plaintiffs will go to trial. To see why this is true, suppose instead that there is a positive mass of settlement in each round: $M_1 > 0$ and $M_2 > 0$.[31] Since the plaintiffs are rational, they must be indifferent between the two settlement offers, S_1 and S_2. Using the previous lemma, there is a cutoff, x^*, where $S_1 = x^* - c_p - k_p$ and $S_2 = x^* - k_p$ and $M_1 + M_2 = F(x^*)$. The defendant's total expected payments are:

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

The defendant pays S_1 to the mass of plaintiffs who accept in the first round ($M_1 > 0$), he pays $S_2 + c_d$ for those cases that settle in the second round ($M_2 > 0$), and pays $x + c_d + k_d$ for those cases that go to trial. Since $S_1 = S_2 - c_p$, we know that $S_1 < S_2 + c_d$ -- *the defendant would like to shift the mass of settlement from the second round to the first round*. This is easily accomplished with an alternative sequence of offers, S_1' and S_2', where $S_1' = S_1$ and $S_2' < S_2$. Now *all* plaintiffs with $x \le x^*$ will accept the first offer, no plaintiff will accept the second offer, and the defendant's payments are lower than before:

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

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 the expression in (3).

Proposition 1: If the defendant can commit to a sequence of settlement offers (and MFNs are not

[31] The case of $M_1 = 0$ and $M_2 > 0$ is very similar and is not presented here.

used), then $S_1 = \hat{x} - c_p - k_p$ and $S_2 < \hat{x} - k_p$ where $\hat{x}$ solves:

$$F(\hat{x}) - (c_p + c_d + k_p + k_d) f(\hat{x}) = 0 .^{32} \tag{4}$$

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

Intuitively, the defendant does not want settlement to take place in the second round because *delay is inefficient*. By committing to a sufficiently unattractive second round offer, the defendant encourages all plaintiffs with low damages, $x \le \hat{x}$, to settle in the first round. In this way, the defendant saves both his own delay cost c_d and extracts c_p from these plaintiffs in settlement. The first-order condition defining the cutoff $\hat{x}$, equation (4), may be understood intuitively. When the defendant increases the settlement offer S_1 by a small amount, Δ, there are both costs and benefits. The cost is that he pays an additional Δ 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 $\Delta f(\hat{x})$ and the cost savings on these plaintiffs is $(c_p + c_d + k_p + k_d)$. The optimal cutoff, $\hat{x}$, equates the marginal cost and marginal benefit.

2.2 Settlement Without Commitment

In the previous section, a truncated distribution of plaintiffs, *F(x)* on $(\hat{x}, \infty)$, remain in the second round. The defendant's second round offer, $S_2 < \hat{x} - k_p$, was not *sequentially rational*. If not fully committed to this sequence, the defendant would make an offer that *at least some* of the remaining plaintiffs would accept: $S_2 > \hat{x} - k_p$. (In this way, the defendant would save his

[32] This is the first-order necessary condition for $\hat{x}$ to minimize (2). The monotone hazard rate condition assures that the solution is unique and that the second-order condition is satisfied.

litigation costs, k_d, and extract the plaintiffs' litigation costs, k_p.) Of course, anticipating that the settlement offer would rise in this way, the plaintiffs with damages $x \leq \hat{x}$ would never be willing to accept $S_1 = \hat{x} - c_p - k_p$ to begin with. The next proposition characterizes the defendant's optimal *sequentially rational* settlement strategy.[33]

Proposition 2: If the defendant cannot commit to the sequence of settlement offers and MFNs are not used, then he will offer $S_1 = \tilde{x} - c_p - k_p$ and $S_2 = \tilde{x} - k_p$ where $\tilde{x}$ solves:

$$F(\tilde{x}) - (c_p + c_d + k_p + k_d) f(\tilde{x}) + (c_p + c_d)(k_p + k_d) f'(\tilde{x}) = 0. \tag{5}$$

In equilibrium, $M_1 = F(\tilde{x}) - (k_p + k_d) f(\tilde{x}) > 0$, $M_2 = (k_p + k_d) f(\tilde{x}) > 0$, and $M_T = 1 - F(\tilde{x})$.

In contrast to the benchmark case in Proposition 1, both M_1 and M_2 are positive here -- there is a positive mass of settlement in each round. This is a necessary feature of the equilibrium: for $S_2 = \tilde{x} - k_p$ to be sequentially rational, it is necessary that mass $M_2 = (k_p + k_d) f(\tilde{x})$ of plaintiffs with damages $x \leq \tilde{x}$ remain in the second round. If *fewer* than $(k_p + k_d) f(\tilde{x})$ remained, then the defendant would want to *raise* his offer above $\tilde{x} - k_p$. If *more* than $(k_p + k_d) f(\tilde{x})$ remained in the second round then the defendant would want to make an offer *below* $\tilde{x} - k_p$. (The formal argument is given in the proof.)

The first-order necessary condition stated in the Proposition may be understood intuitively. In addition to the marginal cost and the marginal benefit described for the

---

33 Spier (1992) assumed either a binary or a uniform distribution of damages, and assumed a skimming condition where the lowest plaintiff types accept the early offers. The analysis presented here places no exogenous restrictions on the plaintiff's settlement strategies and only very general restrictions on the distribution of types. Spier (1992) also did not consider MFN clauses or the welfare consequences of commitment.

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, Δ, 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.

Proof: Let S_1 and S_2 be the equilibrium sequence of offers. Following Lemma 1 we define $\tilde{x} = \max\{S_1, S_2 - c_p\} + c_p + k_p$.

First we will prove that $\tilde{x} = S_2 + k_p$, so S_2 is at least as attractive to the plaintiffs as S_1. Let *g(x)* be the population density of plaintiff types that remain in round 2, and let *G(x)* be the cumulative distribution.[34] (To simplify the proofs we assume that these functions are well-behaved.) By Lemma 1, *g(x)* = *f(x)* for $x > \tilde{x}$ and $g(x) \le f(x)$ for $x \le \tilde{x}$. By standard arguments, the defendant would settle with a positive mass of plaintiffs in the second round (See Bebchuk, 1984). It follows that $M_2 > 0$. Therefore the second round offer must be at least as attractive as the first round offer: $S_2 - c_p \ge S_1$. Therefore $\tilde{x} = S_2 + k_p$.

Now we will show that $M_2 = (k_p + k_d) f(\tilde{x})$ is necessary for $S_2 = \tilde{x} - k_p$ to be sequentially rational. Suppose that the defendant instead offers $S_2 = \tilde{x} + \Delta - k_p$ where $\Delta > 0$. Since plaintiffs with damages below $\tilde{x} + \Delta$ accept this offer and those with damages above reject the offer, the defendant's continuation payments are:

[34] Recall that the total population, represented by *f(x),* was normalized to 1. The mass of plaintiffs remaining in the second round, *g(x),* will in general be less than 1.

$$V(\tilde{x}+\Delta)=(\tilde{x}+\Delta-k_p)[G(\tilde{x})+F(\tilde{x}+\Delta)-F(\tilde{x})]+\int_{\tilde{x}+\Delta}^{\infty}(x+k_d)f(x)dx. \tag{6}$$

Taking the derivative with respect to Δ tells us that $G(\tilde{x})\geq(k_p+k_d)f(\tilde{x})$. If this were not true, the defendant would choose a positive Δ. Now suppose instead that the defendant offers $S_2=\tilde{x}-\Delta-k_p$ where $\Delta>0$. Plaintiffs with damages below $\tilde{x}-\Delta$ accept this offer and those with damages above reject the offer and go to trial. The defendant's continuation payments are

$$V(\tilde{x}-\Delta)=(\tilde{x}-\Delta-k_p)G(\tilde{x}-\Delta)+\int_{\tilde{x}-\Delta}^{\tilde{x}}(x+k_d)g(x)dx+\int_{\tilde{x}}^{\infty}(x+k_d)f(x)dx. \tag{7}$$

Differentiation establishes that $G(\tilde{x})\leq(k_p+k_d)g(\tilde{x})$, for otherwise the defendant would lower the settlement offer. Combing these two conditions we have

$$(k_p+k_d)f(\tilde{x})\leq G(\tilde{x})\leq(k_p+k_d)g(\tilde{x}). \tag{8}$$

Since $g(\tilde{x})\leq f(\tilde{x})$ and $M_2=G(\tilde{x})$ the claim is proven.

Finally, the fact that all $x\leq\tilde{x}$ settle out of court (Lemma 1) and $M_2=(k_p+k_d)f(\tilde{x})$ tells us that $M_1=F(\tilde{x})-(k_p+k_d)f(\tilde{x})\geq 0$. Combining these claims, we may write the defendant's total payment as a function of $\tilde{x}$ alone:

$$(\tilde{x}-c_p-k_p)F(\tilde{x})+(c_p+c_d)(k_p+k_d)f(\tilde{x})+\int_{\tilde{x}}^{\infty}(x+c_d+k_d)f(x)dx. \tag{9}$$

The defendant could do no better than offer $S_1=\tilde{x}-c_p-k_p$ and $S_2=\tilde{x}-k_p$ where $\tilde{x}$ minimizes this expression, and the first-order necessary condition is in the proposition. ■

Sequential rationality dictates that the defendant makes a second round offer that a positive mass of plaintiffs will accept. It follows that a positive mass of plaintiffs must "wait" for the second round offer, plaintiffs who ideally should have settled in the first round. The defendant is of course hurt by his inability to commit to the sequence of settlement offers.

Holding the cutoff $\tilde{x}$ fixed in Proposition 2, the defendant could strictly reduce his total payments by committing to a second period 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.[35]

2.3 Settlement With Most-Favored-Nations Clauses

In this section we will show that the defendant can achieve full commitment power through a most-favored-nation clause in the first round. Specifically, suppose a mass of plaintiffs accepts S_1 with an MFN in the first round and that the defendant subsequently offers to settle for $S_2 > S_1$ in the second round. A most-favored-nation provision obligates the defendant to pay the early settling plaintiffs the difference between the offers: $S_2 - S_1$. We will argue that this obligation makes raising the offer in the second round prohibitively expensive for the defendant. Although this commitment not to raise the offer 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.

<u>Proposition 3:</u> 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 a most-favored-nation (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.

<u>Proof:</u> Suppose that the defendant offers $S_1 = \hat{x} - c_p - k_p$ with an MFN provision, and suppose that all plaintiffs with $x \leq \hat{x}$ accept this offer. So the mass of plaintiff who settle in the first

[35] With more equal sharing of bargaining power, early settlement would be in the joint interest of the plaintiff and defendant. (Here, the defendant pockets the entire savings.)

round is $M_1 = F(\hat{x})$.

First, we will show that the defendant will rationally choose *not to settle* in round 2 with any of the remaining plaintiffs. Suppose not. The defendant's offer in the second round, S_2, makes some plaintiff with damages $x^* > \hat{x}$ indifferent between accepting S_2 and going to trial. Plaintiffs with $x \in (\hat{x}, x^*]$ strictly prefer to accept the second round offer, and those above x^* will reject it and go to trial. Notice that $x^* > \hat{x}$ implies that $S_2 > S_1$:

$$S_2 = x^* - k_p > \hat{x} - k_p > \hat{x} - c_p - k_p = S_1. \tag{10}$$

The MFN from the first round will obligate the defendant to pay the early settlers the difference, $S_2 - S_1$. We may write the defendant's total continuation payments as

$$[F(x^*) - F(\hat{x})][S_2] + \int_{x^*}^{\infty}(x + k_d)f(x)dx + (S_2 - S_1)F(\hat{x}). \tag{11}$$

The first term represents the settlement payments made to all types $x \in (\hat{x}, x^*]$; 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:

$$V(x^*) = F(x^*)(x^* - k_p) + \int_{x^*}^{\infty}(x + k_d)f(x)dx - S_1 F(\hat{x}). \tag{12}$$

Taking the derivative, we see that the slope of this function is

$$F(x^*) - (k_p + k_d)f(x^*). \tag{13}$$

From Proposition 1, the first-order necessary condition for $\hat{x}$ to be the optimal commitment cutoff is $F(\hat{x}) - (c_p + c_d + k_p + k_d)f(\hat{x}) = 0$ and it follows that $F(\hat{x}) - (k_p + k_d)f(\hat{x}) > 0$. Since *F(x)/f(x)* is increasing in *x* by assumption, we see that $F(x^*) - (k_p + k_d)f(x^*) > 0$ for all

$x^* > \hat{x}$. Therefore the defendant offers $S_2 < S_1 = \hat{x} - c_p - k_p < \hat{x} - k_p$. [36]

It is clearly in the interest of plaintiffs with $x \le \hat{x}$ to accept the first offer. The plaintiffs do not expect the offers to improve over time, and so they interpret $S_1 = \hat{x} - c_p - k_p$, together with the MFN clause, as a take-it-or-leave-it offer. ■

Finally, it is worth noting that this outcome is not renegotiation-proof for the following reason. 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 rewrite their settlement agreement to allow for more settlement, they could be made jointly better off *ex post*. (They may be worse off *ex ante*, however.) 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. And even if they could coordinate their actions ex post, it would in many legal situations (a class action, for example) require the approval of a judge.

3. Comparative Statics and Social Welfare Implications of MFNs

In the previous section, we argued that MFNs commit the defendant not to raise his settlement offers over time. Plaintiffs who would have otherwise waited to settle on the courthouse steps are induced to settle early. This has an obvious private and social benefit: MFNs economize on the costs of delay. Indeed, holding the settlement offers fixed at $S_1 = \tilde{x} - c_p - k_p$ and $S_2 = \tilde{x} - k_p$, a social planner would want the parties to include an MFN in

[36] Notice that this strategy is not unique. In fact, the defendant could offer to settle for $\hat{x} - k_p > S_2 > S_1 = \hat{x} - c_p - k_p$. None of the remaining plaintiffs would accept this offer, so the MFN clause would not obligate the defendant to pay additional money to the early settlers.

their contract. But the terms of settlement are chosen opportunistically by the defendant, and will generally change when MFNs are adopted. This section compares the two equilibria from the previous section to see how the terms of settlement, the rate of settlement, and social welfare change when MFNs are permitted.

The next lemma characterizes the relationship between $\hat{x}$, the cutoff when MFNs are used, and $\tilde{x}$, the cutoff when MFNs are not permitted.

<u>Lemma 2:</u> $\hat{x} > \tilde{x}$ (resp. $\hat{x} < \tilde{x}$, $\hat{x} = \tilde{x}$) if and only if $f'(\tilde{x}) > 0$ (resp. $f'(\tilde{x}) < 0$, $f'(\tilde{x}) = 0$).

<u>Proof</u>: Restating equation (4), the cutoff without MFNs, $\hat{x}$, is the solution to:

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

Restating equation (5), the cutoff with MFNs, $\tilde{x}$, is the solution to:

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

These two expressions differ in that equation (5) has an additional term.

If $f'(\tilde{x}) = 0$ (as for the uniform distribution) then clearly $\hat{x} = \tilde{x}$. If $\hat{x} = \tilde{x}$, then combining the two equations shows $(c_p + c_d)(k_p + k_d) f'(\tilde{x}) = 0$. Taken together, $f'(\tilde{x}) = 0$ if and only if $\hat{x} = \tilde{x}$. If $f'(\tilde{x}) > 0$ then $(c_p + c_d)(k_p + k_d) f'(\tilde{x}) > 0$ in equation (5), and so it must be the case that $F(\tilde{x}) - (c_p + c_d + k_p + k_d) f(\tilde{x}) < 0$. Using the assumption that *F(x)/f(x)* is increasing in *x*, we see that $\hat{x} > \tilde{x}$. Now suppose that $\hat{x} > \tilde{x}$. Since *F(x)/f(x)* is increasing in *x* we have $F(\hat{x}) - (c_p + c_d + k_p + k_d) f(\hat{x}) = 0 > F(\tilde{x}) - (c_p + c_d + k_p + k_d) f(\tilde{x})$. It follows from equation (5) that $(c_p + c_d)(k_p + k_d) f'(\tilde{x}) > 0$. Taken together, $f'(\tilde{x}) > 0$ if and only if $\hat{x} > \tilde{x}$. The proof that $f'(\tilde{x}) < 0$ if and only if $\hat{x} < \tilde{x}$ is analogous and we are done. ■

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 used, all settlement takes place in the first round, $M_1 = F(\hat{x})$ and $M_2 = 0$, and so the overall rate of settlement is simply $F(\hat{x})$. (See Proposition 3.) 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 $F(\tilde{x})$. (See Proposition 2.) 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 from Lemma 2.

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

It is also straightforward to characterize the effect of MFNs on the plaintiffs' 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 $S_1 = \tilde{x} - c_p - k_p$. If $\hat{x} > \tilde{x}$ then the first-round offer is clearly higher under the MFN regime. When the first round offer is higher, plaintiffs who accept the offer are better off and those who reject the offer and go to trial instead are no worse off than before.

Proposition 5: MFNs weakly increase (resp. weakly decrease, leave unchanged) the plaintiffs' welfare if and only if $f'(\tilde{x}) > 0$ (resp. $f'(\tilde{x}) < 0$, $f'(\tilde{x}) = 0$). [37]

What can we say about social welfare more generally? We end this section with a

[37] "Weakly" because plaintiffs who go to trial under both regimes receive exactly the same payoff in each.

comparison of the total costs of litigation and delay under the two regimes. Since settlement payments are simply monetary transfers between the plaintiff and defendant, we start by taking these payments from a social welfare perspective.

When MFNs are prohibited or commitment is not possible, the total costs are:

$$\tilde{TC} = [1 - F(\tilde{x})](c_p + c_d + k_p + k_d) + (k_p + k_d) f(\tilde{x})(c_p + c_d), \tag{14}$$

where $\tilde{x}$ is defined in Proposition 2. This expression is very intuitive. Consider 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 costs. From Proposition 3, the total wasted resources when the defendant uses MFNs, thereby committing to settle in the first round only, are:

$$\hat{TC} = [1 - F(\hat{x})](c_p + c_d + k_p + k_d). \tag{15}$$

Using these two expressions we see that the cost savings that come from the use of MFNs, which is the change in social welfare, is:

$$\tilde{TC} - \hat{TC} = (c_p + c_d)(k_p + k_d) f(\tilde{x}) + [F(\hat{x}) - F(\tilde{x})](c_p + c_d + k_p + k_d), \tag{16}$$

where $\tilde{x}$ is the cutoff without MFNs and $\hat{x}$ is the cutoff with MFNs. The first term captures the benefit of early (rather than late) settlement. Formally, it is the volume of cases that would have settled in the second round, $(k_p + k_d) f(\tilde{x})$, multiplied by the savings from having these cases settle in the first round, $(c_p + c_d)$. This first term is always positive: the commitment to settle early increases social welfare. The second term represents the change in the settlement rate. Formally, it is the increase in volume of settlement, $[F(\hat{x}) - F(\tilde{x})]$, multiplied by the cost savings from settlement, $(c_p + c_d + k_p + k_d)$. Since $\hat{x}$ may be either smaller than or greater than $\tilde{x}$ depending on the sign of $f'(\tilde{x})$ (Lemma 2), MFN clauses may either raise or lower the overall volume of settlement. If $f'(\tilde{x}) < 0$ then the second term in expression (16) is negative and, in theory, this effect could dominate the former.

Proposition 6: If $f'(\tilde{x}) \geq 0$ then total costs fall when MFNs are used. If $f'(\tilde{x}) < 0$ then total costs may either rise or fall when MFNs are used.

Example #1: The Uniform Distribution

When $f(x)$ is uniformly distributed then $f'(x) = 0$. Comparing (4) and (5) we see that the plaintiff type who is indifferent between settling and litigating in the two regimes is exactly the same: $\hat{x} = \tilde{x}$. This has several important implications. The overall rate of settlement is unchanged when MFNs are used and the plaintiffs are equally well off under the two regimes.[38] The defendant, however, is strictly better off because MFNs shift the timing of settlement. Cases that would have settled in the second round settle in the first round instead. It follows that the total social costs unambiguously fall when the plaintiffs' damages are uniformly distributed.

Example #2: The Exponential Distribution

We will now show that MFNs can have very positive welfare effect even when the distribution of plaintiff types is everywhere decreasing. Consider, for example, the exponential distribution, $f(x) = (1/\lambda)e^{-x/\lambda}$. A straightforward evaluation of equations (4) and (5) give us closed-form solutions for $\hat{x}$ and $\tilde{x}$:

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

Comparing these two expressions we see that $\hat{x} < \tilde{x}$ -- the overall settlement rate falls when MFNs are adopted. This cost is represented by the second term in equation (16). The benefit is

[38] The first round settlement offers are the same: $S_1 = \hat{x} - c_p - k_p = \tilde{x} - c_p - k_p$.

that cases that would have settled in the second round now settle in the first round instead. This benefit is represented by the first term in equation (16). It is easy to verify that the benefit always outweighs the cost for the exponential distribution.

Example #3: The Normal Distribution

When the plaintiff's damages are normally distributed, social welfare 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 increase social welfare. MFNs reduce social welfare when the parameter values are changed. 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 general results are not obtained.

In closing, there are other reasons to think that the social welfare consequences of MFNs may be either higher of lower than those discussed here. First, the analysis was limited to the costs of delay and litigation. One could go further and argue that MFNs are also bad for social welfare because they allow the defendant to reduce his total expected payments, leading the defendant to take fewer precautions. It is important to note, however, that the dilution of *ex ante* incentives could be addressed by increasing the defendant's liability through, for example, a damage multiplier. Second, one might be concerned about the public good aspect of litigation (e.g. trials are important so the law can evolve efficiently). Since the main effect of MFNs is to encourage early settlement rather than create a barrier to litigation, the public interest in litigation is preserved. A full treatment of these issues is beyond the scope of this paper.[39]

4. Discussion

This section changes several assumptions of the basic analysis with two goals in mind. The first is to indicate how the basic results would be modified in more realistic situations. The second is to shed light on recent legal cases and the practical use of MFNs in settlement.

4.1 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 section argues that although MFNs are ineffective when there is exactly one plaintiff, they may be useful in cases involving small numbers of similar plaintiffs, although they will not typically 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.[40]

---

[39] Shavell (1997) and Spier (1997) discuss the social desirability of private settlement.

[40] The defendant, however, might want to commit to pay damages to some third party if the settlement offer is raised second round. This would require the third party to be involved in these negotiations and to have full information about the offers. The author is not aware of side

Suppose instead that there are N plaintiffs, each with damages that are independently drawn from a common probability density function *f(x)*. The defendant would like to commit to a first period offer, $S_1 = \hat{x} - c_p - k_p$, and a sufficiently unattractive second round offer that any remaining plaintiff would reject. In the full commitment outcome, plaintiffs with damages $x \leq \hat{x}$ would accept the first round offer and plaintiffs with damages $x > \hat{x}$ would reject both offers and go to trial. MFNs cannot implement the full commitment outcome here for the following simple reason: If all plaintiffs happen to have $x > \hat{x}$, then nobody will accept the first round offer. In this case, nothing prevents the defendant from raising his offer to the remaining plaintiffs, skimming off plaintiffs with damages near the bottom of the truncated distribution. The same would be true if an unexpectedly small number of plaintiffs accepted in the first round. In this case, MFN adjustments would be paid on the equilibrium path.

Although the defendant cannot implement the full commitment outcome, MFNs are useful because they will prevent settlement in the second round with positive probability (albeit smaller than unity). The partial commitment power afforded by MFNs should be preferred to no commitment at all. A full formal analysis of this case is left for future research.

4.2 Multiple Groups of Plaintiffs

The main analysis assumed that although individual plaintiffs were heterogeneous, they all looked the same to the (uninformed) defendant before trial. 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. We argued that MFNs are valuable *within* these plaintiff groups because they commit the defendant to not raise his offer in the future. But what

---

contracts of this form, perhaps because they would be very difficult to write and even harder to enforce.

if plaintiffs did not all look alike to the defendant?

Formally, suppose that plaintiffs fall into two groups, A and B. The damages of the plaintiffs in Group A are distributed $f_A(x)$ and the damages of the Group B plaintiffs are distributed $f_B(x)$.[41] Furthermore, suppose that the defendant can observe the individual case characteristics that identify each plaintiff as belonging to one group or another. Each plaintiff's actual damages remain private information, however. For example, the first group of plaintiffs may include smokers from Florida, while the second group would include smokers from Minnesota. Under what circumstances would the defendant want the MFN from Florida to apply to the plaintiffs from Minnesota?

At first glance, it would appear that, in general, 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.[42] 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."*

Many settlement agreements with MFN clauses do, in fact, include exceptions.[43] The

[41] More generally, these two groups may have different litigation costs or other distinguishing features.

[42] Proposition 1 shows how the optimal settlement offer depends on the distribution of damages as well as the litigation and delay costs.

[43] The desire to distinguish between groups appears in many other contexts as well. In natural gas markets, for example, MFNs typically apply withing the boundaries of each field and not across fields. See Butz (1986).

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.[44] More often, however, the exceptions are broadly specified as litigants who are not "similarly situated" with the early settlers. In Corrugated Container litigation, a price fixing case involving many defendants, the MFN clause was interpreted to not apply to defendants who received better terms later because of their inability to pay ("hardship settlements") and was held not to apply to defendants who were not indicted in related criminal cases.[45] Presumably defendants who were not the targets of criminal investigations were less culpable than their indicted counterparts.

The strategy of 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. Consequently, the litigants may try to take advantage of the "similarly situated" clause *ex post*. This may manifest itself in a variety of ways.

First, consider the Group A plaintiffs who reject the settlement offer in the first round. They may argue that they, in fact, are not "similarly situated" to the others in Group A and should therefore be exempt. This situation frequently arises in practice. In the Vitamins case, for example, plaintiff Nutra-Blend argued that it was unique position as a blender and suffered greater damages than the other class members and should therefore be permitted to receive more

[44] See 22 (c) of the Vitamins Antitrust Litigation Settlement Agreement. (United States District Court for the District of Columbia).
[45] See In Re Corrugated Container Litigation, M.D.L. No. 310, 1983 U.S. Dist LEXIS 15969; 1983-1 Trade Cas. (CCH) P65, 451.

in settlement.[46]

Second, suppose the defendant wants to give members of Group B a better deal than Group A. The members of Group A who have already accepted settlement offers now have every incentive to try to prove that, in fact, the members of Group B were "similarly situated" to themselves. If they can convince a court that the Group B plaintiffs are similar to them, then they would be entitled to the MFN payout! 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.[47]

Third, the defendant may have an incentive to misrepresent the truth *ex post* as well. At 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. Of course, if these types of activities were anticipated by the Group A plaintiffs before the first round, then the value of MFNs would be greatly diminished.

In summary, the use of MFNs across groups will lead to opportunistic behavior. Opportunism will tend to undermine the positive effects of MFNs that were identified in previous sections. In addition, these rent-seeking activities may be costly in and of themselves, as is evidenced by the follow-on litigation described in the preceding paragraphs. 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. Therefore the defendant may

---

[46] In Re: Vitamins Antitrust Litigation, Misc. No. 99-197 (TFH), 2000 U.S. Dist. LEXIS 8931; 2000-1 Trade Cas. (CCH) P72,862.

[47] In re: Corrugated Container Litigation, M.D.L. No. 310, 1983 U.S. Dist LEXIS 15969; 1983-1 Trade Cas. (CCH) P65, 451.

forego price discrimination and use MFNs across the two groups, or not use MFNs at all.

Another reason why MFNs may be useful across groups (in addition to within them) arises when groups have small numbers of plaintiffs.[48] As discussed in Section 4.1, MFNs are completely ineffective when a group has exactly one plaintiff, and are less than fully effective 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.

4.3 The Defendant's Liability is Established After the MFN is Signed

Many settlement agreements with MFNs, including the Corrugated Container case, stipulate that the provision is voided if "present circumstances" change. The chairman of the Plaintiff's Steering Committee in the Corrugated Container case writes to defendant's counsel: "You, however, fully understand that we must retain the freedom to enter into settlement agreements on non-comparable terms with defendants under different circumstances presented as the events of this litigation unfold. For example, as the Court issues rulings in this case and the criminal case, as evidence unfolds in either case, or as the law governing liability or class action changes, it might become obvious that we have a weaker or stronger case than we now think

[48] This argument would also apply if there are many plaintiffs represented by a small number of decision makers such as lawyers.

…"[49] Other settlement agreements have allowed the MFN clause to apply even as new information is revealed, as in the MP3.com case. This section extends the previous analysis to consider "changes in present circumstances."[50]

The previous sections assumed that the underlying distribution of expected damages did not change over time. That is, the litigants did not learn any additional information relevant to their cases before the ultimate trial. This section relaxes this assumption by assuming that the defendant's liability is established when the costs c_p and c_d are sunk before the round 2 offers. One could think of this second round as taking place after discovery, or alternatively after the first stage of a bifurcated trial. At this point, the plaintiff and defendant will learn whether the case has sufficient merit to proceed. If it is determined at this point that the defendant is not liable, then the plaintiffs cases will be dismissed or dropped. Would the defendant still want to use an MFN provision under these circumstances? We will see that, perhaps surprisingly, the answer is yes.

Suppose that $y = px$ where p is the probability that defendant will be found liable, and x is a plaintiff's damages. As before, plaintiffs are heterogeneous and their damages x are drawn from the density function $f(x)$. The parameter p is common to all cases, however, and the defendant and plaintiffs are symmetrically uninformed about its value in round 1. Assume that p takes on the value 1 with probability θ and 0 with probability $1-\theta$. If it is discovered that $p = 0$ then the game ends. There is no more basis for the lawsuit and the cases are either dismissed or dropped with no further costs incurred. If it is discovered that $p = 1$, however, then the stakes

[49] In re: Corrugated Container Litigation, M.D.L. No. 310, 1983 U.S. Dist LEXIS 15969; 1983-1 Trade Cas. (CCH) P65, 451.

[50] MP3.com settled for $20 Million each with the four record labels before the first stage where the liability for copyright infringement was determined. The fifth record label, Universal, settled (in the judge's chambers) for $50 Million after this first stage but just before the next trial to determine Universal's damages.

are higher than before. When it is discovered that the defendant is liable for sure, the settlement offer would rise to reflect the stakes. We will show that all of the earlier results continue to hold. In particular, the defendant would like to commit not to deal with plaintiffs who delay settlement, even when the damning information comes to light.

Suppose that the plaintiffs *expect* a sequence of offers S_1 and S_2. A plaintiff will prefer the early offer to the later offer when $S_1 > \theta\, S_2 - c_p$, and will prefer to settle for S_2 rather than go to trial when $S_2 \geq x - k_p$. It is then easy to establish the following result.

<u>Proposition 7:</u> Suppose that the defendant's liability is established before the round 2 settlement offers, while damages remain private information of the plaintiffs. The defendant would still like to commit not to settle in the second round, and can use an MFN provision to assure this.

<u>Proof:</u> The proof is analogous to the one for Proposition 2. Suppose that settlement takes place in each round, so $M_1 > 0$ and $M_2 > 0$. Then $S_1 = \theta\, S_2 - c_p$. The defendant's expected payments are:

$$S_1 M_1 + (\theta S_2 + c_d) M_2 + \int_{\hat{x}}^{\infty} [\theta(x + k_d) + c_d] f(x) dx\,. \tag{17}$$

Since $S_1 = \theta\, S_2 - c_p$, the defendant would like to shift settlement from the second round to the first round. This is easily accomplished with an alternative sequence of offers: let $S_1' = S_1$ and $S_2' < S_2$. All plaintiffs who would have accepted in the second round would now accept in the first round instead, and the defendant's payments are lower than before:

$$S_1 [M_1 + M_2] + \int_{\hat{x}}^{\infty} [\theta(x + k_d) + c_d] f(x) dx\ . \tag{18}$$

■

There are examples of lawsuits where defendants have tried to renege on their MFN obligations as a consequence of this type of information. The message from the theory is clear. First, it is in the defendant's interest to commit to MFNs under these circumstances. Second, although it might be in a defendant's *ex post* interest to renege on the deal (it was in Section 3 as well), it would not be in their *ex ante* interest. Plaintiffs, expecting that the obligation would be voided, would not be willing to settle early, eliminating the positive benefits of MFNs. Therefore it is in the interest of defendants more generally for these agreements to be enforced as written.

It is important to note, however, that MFNs may also have a downside when information is arriving over time. Better information allows the defendant to better tailor his settlement offer to the situation at hand. Under more general conditions than those presented here, new information would affect the cutoff in plaintiff type space in the second round. In other words, the arrival of new information may create option value for the defendant, and therefore he may not want to settle in the first round only.[51] A full treatment of the value of information in settlement is beyond the scope of this paper and is left for future research.

5. MFNs as a Bargaining Tool

The previous sections argued that MFNs are valuable to the defendant (in cases with multiple plaintiffs) -- and often to society as a whole -- because they mitigate the problem of asymmetric information and lead to earlier settlement. For tractability, this point was illustrated in a model where the defendant could make take-it-or-leave-it offers to the plaintiffs. When this

[51] In the simple example presented above, all cases were dropped if the defendant was found not liable. The cutoff that the defendant wanted to implement before the arrival of information was the same as the cutoff he wanted to implement before the arrival of information. If we had instead assumed that the probability of winning at trial was revealed to be either 1/4 or 1, the optimal cutoff would change.

assumption is relaxed, so the plaintiffs have bargaining power as well, then MFNs may be adopted for a second reason as well. Simply put, MFNs commit the defendant to be a tough negotiator and limit the surplus that future plaintiffs can capture in settlement negotiations.[52] Although MFN clauses may be privately desirable for this reason, they tend to be *socially wasteful*. MFNs can lead future negotiations to break down, and the resulting costs of breakdown are not fully internalized by the original contracting parties.[53]

To illustrate this new set of issues, this section assumes that the defendant and the plaintiffs are *symmetrically informed* about the value of the cases. First, we will characterize the effect of a pre-existing MFN has on settlement negotiations between the defendant and a single plaintiff. Second, we will present two stylized examples to illustrate the incentives of the defendant to include an MFN clause in his settlement contracts to begin with.

5.1 The Effect of a Pre-Existing MFN on Settlement.

Suppose that the defendant has settled an earlier case for S_1 with an MFN clause. The defendant is now facing a second plaintiff whose expected damage award at trial is x. This value is observed by the defendant as well as the plaintiff, so there is symmetric information during settlement negotiations. Will the defendant and the second plaintiff agree to settle the case? If so, what will the terms of settlement be?

To answer these questions, we will construct the lower and upper bounds of the bargaining range. The least that the second plaintiff is willing to accept, $\underline{S}$, is simply:

$$\underline{S} = x - k_p, \tag{19}$$

the plaintiff's expected award at trial minus the litigation costs. This lower bound is unaffected

[52] Cooper and Fries (1991) and Neilson and Winter (1994) make arguments along these lines in supplier relationships.

[53] This kind of effect was also featured in Aghion and Bolton (1987).

by the terms of settlement from the earlier case, S_1 .

The upper bound on the bargaining range is more subtle.[54] If the defendant settles with the second plaintiff for $S_2 > S_1$, then he will be obligated contractually to pay $S_2 - S_1$ to the first plaintiff. The defendant's continuation payout from accepting an offer S_2 may be written $S_2 + \max\{S_2 - S_1, 0\}$.[55] The most that the defendant is willing to pay the second plaintiff, $\overline{S}$, makes the defendant indifferent between accepting the offer and going to trial, $\overline{S} + \max\{\overline{S} - S_1, 0\} = x + k_d$, or:

$$\overline{S} = \min\{x + k_d, \tfrac{1}{2}(S_1 + x + k_d)\}. \tag{20}$$

Using the expressions for $\overline{S}$ and $\underline{S}$, it is easy to prove the following:

<u>Proposition 8:</u> Suppose that the defendant has settled an earlier case for S_1 with an MFN clause, and is currently facing a new plaintiff with expected damages, *x*.

1. If $x < S_1 - k_d$ the bargaining range is $[x - k_p, x + k_d]$ and the case settles.
2. If $x \in [S_1 - k_d, S_1 + k_d + 2k_p]$ the bargaining range is $[x - k_p, \tfrac{1}{2}(S_1 + x + k_d)]$ and the case settles.
3. If $x > S_1 + k_d + 2k_p$ the bargaining range is empty and the case goes to trial.

When the second case is much less valuable (*x* is in the lower range) then the parties

[54] If the earlier settlement *did not include an MFN*, then the most that the defendant is willing to accept is $x + k_d$, the expected liability at trial plus the defendant's litigation costs. Notice that settlement will certainly occur in this case, since $\underline{S} < x + k_d$. The division of the bargaining surplus, which is simply the sum of the litigation costs, will depend upon their relative bargaining strengths.

[55] Notice that this function is continuous and everywhere increasing in S_2. It is not differentiable at $S_2 = S_1$.

would settle for $S_2 < S_1$ even without the MFN, so the MFN has no effect on negotiations. When the second case is roughly of the same size as the first (x is in the intermediate range) then the MFN reduces the upper bound of the settlement range. It is in this range that the MFN will serve to extract value from the second plaintiff. Finally, when the second case is very valuable (x is in the high range) the pre-existing MFN causes negotiations to fail.[56]

Would a defendant facing multiple plaintiffs want to include an MFN clause in early settlement contracts? If the defendant anticipates having all of the bargaining power in the future the answer is "no." An MFN clause would not improve the terms of settlement and may lead negotiations to break down in the future (in which case the defendant will bear litigation costs). The defendant may want to adopt an MFN clause if he anticipates not having all of the bargaining power in the future, however. Suppose the defendant does not know the value of x when the MFN is adopted. Proposition 8 suggests an important tradeoff. A *private benefit* accrues in the intermediate range, where the MFN tends to reduce the surplus captured by the second plaintiff.[57] But a *private cost* is borne in the high range, since negotiations break down and cases proceed to costly litigation. Intuitively, the defendant and the early plaintiffs will adopt MFNs if and only if the private benefits of including the clause exceed the private costs.

5.2 The Private Incentive to Use MFNs

This section presents two stylized examples to show why MFNs may be adopted in

[56] If the defendant could renegotiate his MFN contract with the first plaintiff, then efficiency would be restored. The plaintiff would concede on the claim, and each player could capture some of the bargaining surplus. While negotiations would be efficient in this case, MFNs may still be used. In renegotiation, the MFN brings three parties to the bargaining table instead of two, probably reducing the value captured by the second plaintiff. See the related issue in Perotti and Spier (1992) and Spier and Sykes (1998).

[57] This would be true with the Nash bargaining solution, for example.

equilibrium. In the first example, two plaintiffs arrive *simultaneously* in the first round but negotiate independently with the defendant (they do not consolidate their claims). We will see that the defendant can exploit the two plaintiffs through MFN clauses, settling the two cases for less than they would settle for otherwise. In this first example, MFNs simply redistribute value among the litigants. In the second example, the two plaintiffs arrive *sequentially*, one in each round. The defendant and the first plaintiff adopt an MFN clause to "grab value" from the second plaintiff. In addition to redistributing value among the litigants, we will see that value will be destroyed in equilibrium. When the second plaintiff's damages are high, second round negotiations will fail and litigation costs will be wasted.

5.2.1 Simultaneous Plaintiffs

A defendant is facing two plaintiffs *simultaneously*. Each plaintiff has damages x drawn from distribution $f(x)$. The realization of the parameter x is not known in the first round of negotiations, so the litigants are symmetrically uninformed at that point. The parameter x will be revealed to the all litigants before the second round, however. Consider the following timing. In the first round, *the defendant offers to settle with each plaintiff for S_1,* with or without an MFN clause. Each plaintiff individually and non-cooperatively decides whether or not to accept the offer. If a plaintiff rejects, then *the plaintiff can make a single offer to settle in the second round, S_2.* If the plaintiff's second round offer is rejected the case goes to trial and costs k_p and k_d are borne.[58] For simplicity, there is no discounting and there are no additional costs of delay ($c_p = c_d = 0$).

[58] The assumption that the defendant has all of the bargaining power in the first round and the plaintiff has all of the bargaining power in the second round was made to streamline the analysis. A more general analysis would allow a coin flip to determine who makes the offer in each round, but the same underlying effects would be present.

First, suppose that the defendant offers S_1 *without an MFN clause* in the first round. The plaintiffs are no dummies. They know that they will have all of the bargaining power in the second round and would settle for $S_2 = x + k_d$ if the case reaches that stage. Therefore the best that the defendant can do in the first round is to offer to settle for $S_1 = E(x) + k_d$. Without MFNs, all of the bargaining surplus is conceded to the plaintiffs. This is a standard result: the player who makes the last offer captures the surplus.

Now suppose instead that the defendant offers S_1 *with an MFN clause*. If one plaintiff expects that the other plaintiff will accept the offer, then using Proposition 8 we can write that plaintiff's payoff from rejecting the first round offer as:

$$\int_0^{S_1-k_d}(x+k_d)f(x)dx+\int_{S_1-k_d}^{S_1+k_d+2k_p}\tfrac{1}{2}(S_1+x+k_d)f(x)dx+\int_{S_1+k_d+2k_p}^{\infty}(x-k_p)f(x)dx\,. \tag{21}$$

In the low range, the plaintiff demands the full $x + k_d$, the defendant's expected payments at trial. In the intermediate range, the plaintiff demands only $\tfrac{1}{2}(S_1 + x + k_d)$. This reflects the additional payments that the defendant will have to make to the other plaintiff who settled earlier. In the high range the case goes to trial and the plaintiff receives $x - k_p$. The plaintiff's expected payoff is, of course, strictly smaller than $E(x) + k_d$, what the plaintiff would get if no MFN were in effect. It is straightforward to show that if the defendant offers to settle for $S_1 \in [E(x) - k_p, E(x) + k_d]$ with an MFN in round 1, then there is an equilibrium where both plaintiffs accept.[59] In a nutshell, plaintiffs are concerned about rejecting early offers, for fear of

[59] The careful reader will notice that there is a second equilibrium where both plaintiffs reject the settlement offers. The plaintiffs do better in this second equilibrium, for they will receive $E(x) + k_d$ in expectation. Although this second equilibrium certainly Pareto dominates the first, it is not the "risk dominant" equilibrium. Therefore we would expect that the defendant could

being the victim of an MFN in the future. In this way, the defendant gets the plaintiffs to accept less than $E(x)+k_d$.

There is *no direct social waste* from the use of MFNs in this simple example. Both plaintiffs settle in the first round and no litigation costs are spent on the equilibrium path. Note, however, that there may be *indirect social costs*. Since the defendant uses MFNs to evade liability, he may not take as much care in his primary activities as he would if MFNs were banned. While interesting and important, the effects on incentives to take *ex ante* precautions is beyond the scope of this paper.

MFNs are used here precisely because they can take advantage of a lack of coordination among the plaintiffs. If the plaintiffs are well-organized, as they probably were in class action lawsuits like the Vitamins case, we would not expect to see MFNs adopted for the reason identified here. Indeed, the ability of the defendant to take advantage of plaintiffs in this way gives the plaintiffs the incentive to coordinate their lawsuits and hire a common lawyer to represent them.

5.2.2 Sequential Plaintiffs

Now suppose instead that the two plaintiffs arrive sequentially. The defendant negotiates with the first plaintiff *before the second plaintiff appears*. At the time of the first round negotiations, it is known that the second plaintiff's damages, x, are drawn from distribution $f(x)$, but the realization of x is not known until the second plaintiff actually arrives. We will see that the decision to include an MFN in the first round hinges upon the defendant's expectations of his future bargaining power in the second round.

Suppose that the defendant expects to have all of the bargaining power when negotiating

---

take advantage of the plaintiffs as described in the text so long as S_1 is not too small. See the related discussion in Spier (2000).

with the second plaintiff or, equivalently, the ability to make a take-or-or-leave-it offer. Without an MFN clause, the defendant would offer to settle with the second plaintiff for $S_2 = \underline{S} = x - k_p$, and would capture all of the bargaining surplus, $k_p + k_d$. So there is no private benefit of including the MFN. There is a private cost, however. When x is in the high range the case would go to court and defendant would pay $x + k_d$ in total. It is clear that MFNs will not be used when the defendant expects to have all of the bargaining power in the second round.

Now suppose instead that the defendant expects to have no bargaining power in the second round. The second plaintiff will make a take-it-or-leave-it offer to the defendant. Without the MFN, the defendant would be forced to settle for $S_2 = x + k_d$ in the second round. With an MFN in place, the defendant will still pay exactly $x + k_d$. (He will either settle for S_2 that makes him indifferent between going to trial and not going to trial, or he will actually go to trial.) Once the second round is reached, the defendant neither gains nor loses from having an MFN.[60] The first plaintiff, on the other hand, receives an additional payment of $S_2 - S_1 = \frac{1}{2}(x + k_d - S_1) > 0$ when x is in the intermediate range, surplus that would otherwise have gone into the second plaintiff's pocket.[61] Therefore the defendant and the first plaintiff jointly gain from including an MFN clause in their settlement contract.

Although MFNs may be in the mutual interest of the first plaintiff and the defendant in

---

[60] The defendant does not lose anything when negotiations break down. The losses are borne by the second plaintiff, who can no longer extract the bargaining surplus.

[61] The magnitude of the first round offer will be determined by negotiations. The defendant would like S_1 to be as small as possible (he pays S_1 to the first plaintiff and is driven down to his outside option in the second round), while the first plaintiff would like S_1 to be as large as possible. To see why this latter statement is true, write the first plaintiff's continuation payoff as $S_1 + \int_{S_1 - k_d}^{S_1 + k_d + 2k_p} \frac{1}{2}(x + k_d - S_1) f(x) dx$. Differentiation shows that this is everywhere increasing in S_1.

this example, they are *socially wasteful*. The defendant and the first plaintiff adopted the MFN for purely rent-seeking reasons: to extract surplus from the second plaintiff in the intermediate range. When the second plaintiff has a very valuable case, however, the case goes to trial and the litigation costs, $k_p + k_d$, are wasted. (This did not happen in the first example because both plaintiffs were present in the first round.) The expected social cost is $(k_p + k_d)[1 - F(S_1 + k_d + 2k_p)]$, the litigation costs multiplied by the probability that negotiations ultimately break down.

6. Concluding Remarks

This paper has focused on two particular reasons for the existence of MFNs in settlement contracts. In the main analysis, MFNs are valuable because they commit the defendant not to raise his settlement offers over time and therefore encourage early settlement. We also presented a second model where MFNs may be an effective bargaining too. MFNs commit the defendant to be tougher in future negotiations, and this allows the defendant to extract more of the bargaining surplus.

A defendant facing a group of plaintiffs may enjoy other advantages from using MFNs as well. 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 have otherwise 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.[62] Another rationale is that MFNs may serve to discourage

[62] Notice these inefficiencies would be greater if MFNs applied to judgments at trial. See McAvoy (1962) and Broadman and Montgomery (1983) for discussions of risk sharing and most favored customer clauses in natural gas contracts.

future suits. Since plaintiffs would be forced to bring their cases all the way 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 their case is too weak to pursue.

It is also plausible that MFNs may allow an informed defendant to signal his type to the plaintiffs.[63] 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 signaling story may have some appeal, it is not consistent with an important feature of most (but not all) 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 is the antitrust settlement between Microsoft and the state of New Mexico where 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." [64] 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.

---

[63] 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.

[64] See "New Mexico Cuts Deal with Microsoft," *The Seattle Times*, July 13, 2001.

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.[65] The court was particularly concerned here that 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."[66] These, and the previously discussed motivations for MFNs in settlement agreements, remain fruitful areas for further research.

[65] 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?
[66] In re: Chicken Antitrust Litigation, Civ. A. No. C74-2454A, 560 F. Supp. 943; 1979 U.S. Dist LEXIS 13430.

7. Bibliography

Aghion, Philippe, and Patrick Bolton, 1987, "Contracts as a Barrier to Entry," *American Economic Review*, 77(3), pp. 388-401.

Bebchuk, Lucian, 1984, "Litigation and Settlement Under Imperfect Information," *RAND Journal of Economics*, 15(3), pp. 404-415.

Besanko, David and Lyon, Thomas P., 1993, "Equilibrium Incentives for Most-Favored Customer Clauses in an Oligopolistic Industry," *International Journal of Industrial Org*anization; 11(3), pages 347-67.

Broadman, Harry G. and Montgomery, W. David, *Natural Gas Markets After Deregula*tion, Washington D.C., Resources for the Future.

Bulow, Jeremy, 1982, "Durable-Goods Monopolists," *Journal of Political Eco*nomy, 90(2), 314-32.

Butz, David A., 1986, "Long-term Contracting in Field Markets for Natural Gas: A 'New' Perspective on the Most-Favored Nation Provision," Northwestern University Ph.D. Thesis.

Butz, David A., 1990, "Durable Goods Monopoly and Best-Price Provisions," *American Economic Review*, 80(5), pp. 1062-76.

Butz, David A., 1995, "Most-Favored Treatment Provisions as Nondiscrimination Guarantees," *International Journal of the Economics of Business*, 2(1), pp. 65-86.

Coase, Ronald H., 1972, "Durability and Monopoly," *Journal of Law and Economics*, 15(1), pp. 143-149.

Che, Y.-K. 1999. "The Economics of Collective Negotiations in Pretrial Bargaining," mimeo.

Che, Y.-K. 1996. "Equilibrium Formation of Class Action Suits," *Journal of Public Economics,* 62(3), pp. 339-361.

Che, Y.-K. and J. G. Yi. 1993. "The Role of Precedents in Repeated Litigation," *Journal of Law, Economics, and Organization* , 9(2), pp. 399-424..

Chang, H. F. and H. Sigman. 2000. "Incentives to Settle Under Joint and Several Liability: An Empirical Analysis of Superfund Litigation," *Journal of Legal Studies*, 29(1), pp. 205-36

Coffee, J. 1999, "Class Wars: The Dilemma of Mass Tort Action," 95 *Columbia Law Review* 1343.

Cooper, Thomas E., 1986, "Most-Favored-Customer Pricing and Tacit Collusion," *RAND Journal of Economics*; 17(3), pp. 377-88.

Cooper, Thomas E., and Fries, Timothy L., 1991, "The Most Favored Nation Pricing Policy and Negotiated Prices," *International Journal of Industrial Organiz*ation, 9(2), pp. 209-23.

Crocker, Keith J., and Lyon, Thomas P., 1994, "What do Facilitating Practices Facilitate? An Empirical Investigation of Most-Favored-Nation Clauses in Natural Gas Contracts," *Journal of Law and Economics*; 37(2), pp. 297-322.

Daughety, A. F. 2000. "Settlement," in *Encyclopedia of Law and Economics Volume 5*, Bouckaert, B. and De Geest, G. (eds.). Edward Elgar Publishing Co., pp. 95-158.

Daughety, A. F. and J. F. Reinganum. 1999. "Hush Money," *RAND Journal of Economics,* 30(4), 661-78.

Daughety, A. F. and J. F. Reinganum. 2001. "Informational Externalities in Settlement Bargaining: Confidentiality and Correlated Culpability," Vanderbilt University mimeo.

DeGraba, Patrick, and Postelwaite, Andrew, 1992, "Exclusivity Clauses and Best Price Policies in Input Markets," *Journal of Economics and Management Strate*gy, 1(3), pp. 423-54.

Hay, Bruce, and Kathryn E. Spier, "Settlement of Litigation," in Peter Newman, ed., *The New Palgrave Dictionary of Economics and the Law*, Macmillan Reference Limited, 1998. pp. 442-451.

Kornhauser, L. and Revesz, R. 1994a. "Multidefendant Settlements: The Impact of Joint and Several Liability," *Journal of Legal Studies*, 23(1), pp. 41-76.

Kornhauser, L. and Revesz, R. 1994b. "Multidefendant Settlements Under Joint and Several Liability: The Problem of Insolvency," *Journal of Legal Studies*, 23(1), pp. 517-42.

Levy, Stephan M., "Best Price Guarantees as a Quality Signal," mimeo, June 2000.

Lyon, Thomas P., "Most Favored Customer Clauses," Peter Newman, ed., *The New Palgrave Dictionary of Economics and the Law*, Macmillan Reference Limited, 1998.

Marx, Leslie M., and Shaffer, Greg, "Opportunism and Nondiscrimination Clauses," University of Rochester mimeo, October 2000.

McAfee, R. Preston and Schwartz, Marius, 1994, "Opportunism in Multilateral Vertical Contracting: Nondiscrimination, Exclusivity, and Uniformity," *American Economic Review*; 84(1), pp. 210-30.

MacAvoy, Paul, 1962, *Price Formation in Natural Gas Fie*lds, New Haven, CT, Yale University Press.

Fiona Scott Morton, 1997, "The Strategic Response by Pharmaceutical Firms to the Medicaid

Most-Favored-Customer Rules," *RAND Journal of Economics*; 28(2), pp. 269-90.

Nalebuff, Barry, "Credible Pretrial Negotiation" *RAND Journal of Economics*; 18(2), Summer 1987, pages 198-210.

Neilson, William S., and Harold Winter, 1994, "Enhancing Bargaining Power with Most-Favored-Customer Pricing," *Economics Letters*, 44(1-2), pp. 61-66.

Perotti, Enrico, and Kathryn E. Spier, 1993, “Capital Structure as a Bargaining Tool: The Role of Leverage in Contract Renegotiation,” *The American Economic Review,* 83(5), pp. 1131-1141.

P'ng, I.P.L. 1991, "Most-Favored-Customer Protection versus Price Discrimination over Time," *Journal of Political Economy*; 99(5), pp. 1010-28.

Salop, Stephen C., "Practices that (Credibly) Facilitate Oligopoly Coordination," in J.E. Stiglitz and G. F. Matthewson, eds., *New Developments in the Analysis of Market Structure*, Cambridge, MA: MIT Press, 1986.

Schnitzer, Monika, 1994, *"Dynamic Duopoly with Best-Price Clauses,"* RAND Journal of Economics; 25(1), pp. 186-96.

Shavell, Steven, 1997, "The Fundamental Divergence between the Private and the Social Motive to Use the Legal System," *The Journal of Legal Studies*, 26(2), pp. 575-612.

Spier, Kathryn E., 1992, “The Dynamics of Pretrial Negotiation,” *The Review of Economic Studies*, 59(1), pp. 93-108.

Spier, Kathryn E., "Settlement with Multiple Plaintiffs: The Role of Insolvency," mimeo, October 2000.

Spier, Kathryn E., 1997, “A Note on the Private Versus Social Incentive to Settle Under a Negligence Rule.” *The Journal of Legal Studies*, 26(2), pp. 613 - 621.

Spier, Kathryn E., and Alan Sykes, 1998, “Capital Structure, Priority Rules, and the Settlement of Civil Claims," *The International Review of Law and Economics*, 18(2), pp.187-200.

Stokey, Nancy, 1981, "Rational Expectations and Durable Goods Pricing," *Bell Journal of Economics*, 12(1), pp. 112-128.