Contract Penalties, Monopolizing Strategies, and Antitrust Policy

Joseph F. Brodley*
and Ching-to Albert Ma**

* Professor of Law and Economics and Kenison Distinguished Scholar of Law, Boston University.

** Assistant Professor of Economics, Boston University.

We received valuable comments from Phillipe Aghion, Patrick Bolton, Robert Bone, Paul Denis, Herbert Hovenkamp, Steve Marks, Robert Merges, Mark Pettit, Michael Salinger, and from workshop participants at Boston University, London School of Economics, Oxford University, and the University of Limburg (Maastricht). We thank Mark Chavey, Alison Fox, and Toni Leong for dedicated research assistance.

INTRODUCTION

Two conflicting interpretations of exclusionary conduct battle for ascendancy in modern antitrust law and industrial organization theory. The first takes a strategic view and maintains that, given sufficient incentives and inducing circumstances, firms with market power will exploit their rivals in ways that reduce economic welfare. The second approach, an efficiency explanation, holds that in the absence of cartelization, firms will behave efficiently and will rarely engage in strategic conduct. Contract penalty clauses, through which a monopolist may penalize customers for switching to rival firms, provide an illuminating context in which to explore these competing interpretations of exclusionary conduct.

In the famous *United Shoe* case,[1] the court ruled that a shoe machinery manufacturer's practice of leasing machines under long-term agreements that required its customers to pay a penalty for switching to a rival supplier violated the Sherman Act. The court found that the leases had been an important means by which United Shoe had monopolized the shoe machinery manufacturing market for over fifty years. Modern critics of *United Shoe* characterize the leasing agreements as an appropriate means of achieving efficient production. So the matter stood, until a recent article by Professors Aghion and Bolton illuminated the strategic potential of contract penalty

1. United States v. United Shoe Mach. Corp., 110 F. Supp. 295 (D. Mass. 1953), *aff'd per curiam*, 347 U.S. 521 (1954).

clauses when used by a monopolist.[2]

Breaking new economic ground, Aghion and Bolton demonstrate that a monopolist can use a contract penalty clause as a strategic mechanism to enhance its monopoly power. The penalty clause binds the monopolist and its customers in a coalition that wields monopoly buying power against potential entrants. The penalty clause forces an entrant to pay a penalty, through lower prices, in order to solicit customers from the monopolist. This deprives the entrant of its expected economic return, and in some cases, deters entry altogether. The economic consequences are reduced output, diminished return to innovation and new entry, and enhanced profit for the monopolist.

The Aghion-Bolton analysis powerfully illustrates how game theory enhances antitrust understanding; provides an example of an important family of strategic models based on credible precommitment; applies to numerous transactions; and has wide ranging implications for antitrust policy. This article attempts to distill and restate the new economic learning for legal readers; to formulate workable criteria for applying the theory to antitrust analysis; to illustrate several applications of the theory to specific cases; and to provide a framework for future policy analysis and investigation of contract penalty clauses.

I. A STRATEGIC APPROACH

Aghion and Bolton present a strategic explanation of contract penalties used by monopolists in long-term contracts. They demonstrate that a contract penalty clause provides a striking and effective means to exploit monopoly power. Under the contract law definition, a penalty allows the nonbreaching party to recover more than its actual or reasonably anticipated losses.[3] Under the definition we shall adopt, a penalty permits the nonbreaching party to recover more than its actual losses as determined at the time of breach.[4] Although the common law holds that penalties are not enforceable, judicial interpretations are highly permissive and often enforce contracts with penalty-type provisions. Moreover, in civil law jurisdictions, which include most industrialized nations except the United States, penalties are enforceable unless found to be unconscionable.[5]

Aghion and Bolton demonstrate the pernicious effects of a penalty on new entrants. The penalty acts much like an import tariff, raising the costs of market entry. The penalty deters entry by some potential entrants even though their costs are lower than those of the incumbent monopolist, and the penalty reduces the profitability of entry for even the most efficient potential competitors. The penalty thus decreases allocative efficiency and

2. Philippe Aghion & Patrick Bolton, *Contracts as a Barrier to Entry*, 77 AM. ECON. REV. 388 (1987); *see also* Peter A. Diamond & Erik Maskin, *An Equilibrium Analysis of Search and Breach of Contract, I: Steady States*, 10 BELL J. OF ECON. 282 (1979).

3. *See* RESTATEMENT (SECOND) OF CONTRACTS § 356(1) (1981); U.C.C. § 2-718(1).

4. *See* text accompanying note 141 *infra*.

5. *See* text accompanying notes 53–59 *infra*.

raises rent-seeking losses, enhancing the profit from monopoly. In addition, the penalty reduces long-term incentives for innovation and technological progress.

In *United Shoe*, an incumbent monopolist leased its most complex machines under long-term agreements that penalized customers who switched to a competitor's machines. Judge Wyzanski denounced these and other lease terms as creating "unnatural barriers" against actual and potential competition and ordered the penalty stricken.[6] Following *United Shoe*, debate focused on whether long-term leasing and other long-term supply contracts by a monopolist were inherently exclusionary. A naive strategic view maintained that a monopolist-supplier could preclude market entry through long-term contracts. Under this view, the monopolist closes off the market through long-term supply contracts before rivals have the opportunity to enter. Later, when potential rivals attempt entry, the monopolist's customers are already bound by the long-term contracts and cannot be a source of sales. Thus, long-term contracts enable an incumbent monopolist to perpetuate its monopoly.

Posner, Bork, and others challenge this logic, arguing that customers will not sign contracts restricting their future choices unless compensated for their lost economic freedom.[7] Posner and Bork's argument appears economically compelling on its surface. As a matter of economic logic, if a monopolist forecloses a customer's future source of supply through contract, the customer will have to be compensated, perhaps by a reduction in the monopoly price. However, Posner and Bork explicitly assume that the compensation must come from the incumbent monopolist; this is not necessarily true. Aghion and Bolton show that the compensation may come from the extraction of rents from future, more efficient suppliers. This rent extraction occurs through a contract penalty that forces future entrants to reduce prices to the monopolist's customers or prevents entry altogether, as discussed below.[8]

---

6. United States v. United Shoe Mach. Corp., 110 F. Supp. 295, 345, 347–50 (D. Mass. 1953), *aff'd per curiam*, 347 U.S. 521 (1954).

7. *See* RICHARD A. POSNER, ANTITRUST LAW: AN ECONOMIC PERSPECTIVE 203–04 (1976); *see also* ROBERT H. BORK, THE ANTITRUST PARADOX: A POLICY AT WAR WITH ITSELF 138–42, 304, 309 (1978).

8. The Posner-Bork argument can also be challenged for its unrealistic assumptions about consumer behavior. Posner and Bork implicitly assume that customers are fully informed, not influenced by technological switching costs, and capable of accurately calculating the lifecycle price of a durable good. However, some, if not all, of these assumptions may not be true. In a recent decision, the Supreme Court questioned the assumption of consumer rationality in the context of a tying arrangement for a durable good. Eastman Kodak Co. v. Image Technical Serv., Inc., 112 S. Ct. 2072 (1992). The Court would not accept without proof the assumption that consumers fully evaluated the costs and benefits of a durable good over the product lifecycle. *Id.* at 2085–87. However, for purposes of this analysis, we ignore the potential weaknesses of the Posner-Bork assumption of rational and fully informed consumer behavior because we believe that this context provides the most rigorous test of the Aghion-Bolton theory. If penalty contracts injure competition when consumers are fully aware of their costs and unhampered by technological switching costs, penalty contracts will surely injure competition when consumers are less informed and less able to switch. Moreover, even in a market with imperfect information, consumers' inability to evaluate economic

A. *The Aghion-Bolton Model*

Aghion and Bolton depict a strategic interaction or economic game among three actors: an incumbent monopolist, a customer, and a potential entrant who wishes to sell to the same customer. A self-enforcing coalition forms between the incumbent monopolist and its customer with the goal of exploiting the entrant. It is assumed that all actors seek only their individual profit and respond opportunistically to the actions of others. Nevertheless, the coalition will survive despite the entrant's effort to drive a wedge between the coalition partners. We illustrate the formation of the coalition with an informal explanation, written as a fable, followed by a more systematic description.

1. *An intuitive fable.*

Once there was a monopolist with many small customers. The monopolist, protected by patents, behaved as monopolists do, charging a high price and leaving its customers no choice but to pay. As the monopolist's patents gradually expired, there came a time when both the customers and the monopolist realized that a more efficient entrant might some day appear. If that happened, the market price would fall, the customers would pay less, society would save productive resources, and the monopolist would be disciplined to improve its performance. While it was not certain that an entrant would appear, the customers were hopeful. But the monopolist, who had never faced competition, was fearful. So the monopolist consulted a strategist versed in the theory of economic games, asking if there was some way to avoid this dreaded possibility.

The strategist had a proposal: If the monopolist would follow the strategist's advice, the monopolist would be able to increase its profits by capturing a share of a future entrant's earnings. The strategist advised the monopolist to offer its customers a long-term contract requiring payment of a large penalty for switching to a rival supplier; the monopolist followed this advice.

The customers were incredulous, asking why they should agree to such an onerous restriction. The monopolist told its unbelieving customers that they would benefit from signing: "It is true that if you sign this agreement, you must pay a penalty for switching to a new entrant, but the entrant will compensate you. The penalty will force the entrant to offer you a reduced price so you can pay the penalty and still gain from switching. This will be a much lower price than you could otherwise obtain from the entrant. To be sure, the entrant will not be happy about this because it stands to lose most of the expected profit from its lower cost technology, but there is no way to avoid this result. Renegotiation is impossible because you are bound by the contract, and of course, we will not release you from the contract."

effects is only partial. Thus, the Aghion-Bolton model would still explain the significant rational component of consumer behavior.

"Although the penalty will be paid to us, that detail need not concern you. We will divide the penalty with you! Not only that, we will pay you in advance by reducing your price immediately, and we will give you that price reduction regardless of actual entry. You could not do as well on your own because you have no way of negotiating with the entrant individually, and both the antitrust laws and your large numbers prevent you from negotiating jointly."

The customers were skeptical; they knew the monopolist would receive large gains. Indeed, if entry occurred, the monopolist would be likely to make more money through the penalty than through actual sales. Moreover, entrants with costs only slightly below the monopolist's would be deterred from entry because their cost advantage would be less than the penalty. While a highly efficient entrant could still enter, most of its economic profit would go to the monopolist through the penalty. But these facts did not prevent the customers from signing the penalty contract.

Each customer reasoned: "I am really facing a lottery. If I refuse to sign the contract, an entrant will either appear or not appear. If an entrant appears, I will get a cheaper price, but the entrant will confer no gifts. The entrant will offer me a price no lower than just necessary to displace the monopolist. If an entrant never appears, I will continue to pay the high monopoly price. On the other hand, if I agree to sign the penalty contract, I receive an immediate and unconditional price reduction."

After reflecting, the customer reasoned further: "I believe I am better off signing up with the monopolist. The immediate price reduction the monopolist offers is worth more than the lottery I will face if I refuse to sign. This difference is not an illusion, but arises because my commitment to pay a penalty for switching forces the low cost entrant to offer a price much below what the entrant would ever have considered otherwise. It is true that the immediate beneficiary of this economic strategy is the monopolist, but since it will share the gain with me, I am better off." Thus, the customer decided to sign the contract.

An additional reflection erased any doubt in the customer's mind about his decision. The customer thought to himself: "Suppose I were to refuse to sign the agreement with the monopolist. Unless a sizable number of others also refuse, even the most efficient entrant will not be able to achieve the scale economies necessary to enter the market. In that case, entry will not occur, and I will be forced to negotiate with a monopolist who knows I now have no other options, and has recently been making little asides on the importance of customer fidelity. On the other hand, if I sign the contract, I earn no disfavor, and I may even be rewarded for my loyalty if some other customers refuse to sign." The customer realizes that if a low cost supplier later enters the market, he may live to regret signing the agreement and instead wish to be free to switch to the entrant. But at the time he signs the contract, the customer knows that it is the best option, and his later regret

will be nothing more than the pangs he sometimes feels at the race track for not betting on a long shot who surprises everyone by winning.

Thus, the customer is induced to sign the penalty contract by both a carrot and a stick. The carrot is the lowering of the contract price by the monopolist. The stick, which comes into play when there are multiple customers, is a "reverse free rider" effect that discourages the customer from refusing to sign the penalty contract without a firm commitment that other customers will also refuse. The end result, whether achieved by the carrot alone or in combination with the stick, is that highly efficient entrants lose most of the profit from entry, and less efficient entrants are deterred altogether.

2. *A more systematic account.*

The key element of the Aghion-Bolton model is a penalty requiring customers who switch to a rival to pay an amount that exceeds the monopolist's lost profit. By way of the penalty contract, the monopolist forms a coalition with its customers to exert monopolistic buying power against more efficient entrants. In the absence of a penalty contract, a more efficient entrant could offer customers a lower price and keep the entire economic surplus resulting from its lower costs. However, if the incumbent monopolist and its customers enter into a long-term penalty contract *before* the entrant appears, they can force the entrant to price below the level at which entry would have been feasible without a penalty and thereby capture some or all of the entrant's economic surplus.

The leverage of the penalty contract is an application of Schelling's principle that credible commitment is a powerful bargaining strategy: "[T]he power to constrain an adversary may depend on the power to bind oneself . . . ."[9] The penalty contract binds customers in advance to pay a substantial switching penalty. When an entrant later appears and offers customers a lower price, customers can credibly respond that the penalty contract prevents them from accepting any offer that does not fully compensate them for the switching penalty. Accordingly, the entrant must lower its price to compensate the customers for the penalty; at the extreme, the entrant loses its entire economic profit because of the penalty.

This bargaining strategy will succeed, however, only if entry is uncertain at the time the penalty contract is signed. If entry is certain, as in the case of actual entry, a coalition between the incumbent and its customers will not be mutually beneficial. The lower cost entrant can always induce customers to reject the contract by undercutting the incumbent's price. Thus, the customers will not sign the penalty contract, preferring to play off the entrant against the incumbent, and the entrant will gain to the extent of its cost advantage over the incumbent. A penalty contract strategy will succeed only in cases of potential competition because the customers bind themselves

9. THOMAS C. SCHELLING, THE STRATEGY OF CONFLICT 22 (1960):

in advance to a negotiating strategy that forces the entrant to reduce its price significantly below the incumbent's costs.

The coalition between the monopolist and its customers has achieved what neither could do alone: the capture of the entrant's economic rent and the raising of entry costs. Depending on the entrant's costs, the penalty either deters a lower cost entrant from entering the market, or transfers most of the entrant's economic surplus to the monopolist. Deterrence occurs when the entrant does not have sufficient cost advantage to undercut the monopolist's price after absorbing the penalty; rent capture—collecting part of the profits the lower cost entrant would otherwise have earned—occurs when the entrant can absorb the penalty and offer consumers a lower net price. Since the penalty exceeds the monopolist's previous profit, the monopolist can readily compensate customers for signing the penalty contract by lowering the contract price.

Thus, the penalty is an effective means of maximizing the joint surplus of the coalition, but reduces overall efficiency and the incentive for efficient entry.[10] The economic effects of penalty contracts can be illustrated through applications to three different contract regimes: spot market sales, long-term contracts without penalties, and long-term contracts with penalties. We assume that an incumbent monopolist sells to a customer without bargaining power and that both parties have the same information on the probability of new entry. An initial analysis of nonpenalty regimes provides a benchmark to assess the effects of penalties.

Spot market sale. If an incumbent monopolist sells to customers on an individual or spot market basis, filling orders as they are received, the seller will charge a monopoly price, earn a monopoly return, and capture the customer's economic surplus. For example, if the customer values the product at $80 while the monopolist's cost is $50, the monopolist will charge $80, earn profit of $30, and capture all of the consumer surplus, defined to be the difference between a consumer's valuation of a good and the price he actually pays for it. An entrant with costs below $50 could displace the incumbent by undercutting its price. If the entrant's cost is $35 and there is no collusion, competition will drive the market price down to just under $50, the monopolist's cost. The entrant earns economic profit of $15, and con-

10. The detrimental economic impact of the transfer of surplus or rent between two producers may not be readily apparent. Rents are defined as returns above those necessary to induce the owner to devote an asset to a particular use. Entry and innovation are dynamic processes involving risky investment, and rent is a necessary inducement to such investment. Thus, although rent capture may not affect the *immediate* supply of the product, it reduces the incentive for *future* investment by potential entrants. Specifically, rent capture in contract penalty cases influences the supply of the good over a period of time. *Cf.* F.M. SCHERER & DAVID ROSS, INDUSTRIAL MARKET STRUCTURE AND ECONOMIC PERFORMANCE 624 (3d ed. 1990) (explaining that a policy of reneging on patent grant after invention is developed would discourage future patent investments). *See generally* Jennifer F. Reinganum, *The Timing of Innovation: Research, Development, and Diffusion, in* 1 HANDBOOK OF INDUSTRIAL ORGANIZATION 849, 904 (Richard Schmalensee & Robert D. Willig eds., 1989) (arguing that firms will underinvest in innovation if investors cannot reap appropriate rewards).

sumers enjoy a $30 surplus from the reduction in the market price.[11]

Long-term contract without penalty. In the second scenario, the monopolist offers the customer a long-term supply contract without a penalty for early termination. If both parties believe that entry by a more efficient supplier is probable during the term of the contract, the customer will be unwilling to sign a long-term contract at the same price he would have agreed to pay absent probable entry. The customer now can choose not to sign and continue purchasing on the spot market, with the prospect of dealing with a lower cost entrant in the future. Thus, the monopolist must reduce his price to induce the customer to sign the long-term contract.[12]

If a lower cost entrant appears during the course of the contract, the customer will breach its contract with the incumbent and pay contract damages equal to the incumbent's lost profit. The customer's breach is efficient because production is then carried out by the lower cost producer, while allowing the incumbent to receive its bargained-for profit.[13] Moreover, the

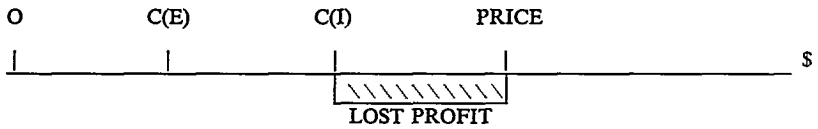
---

11. Aghion and Bolton assume Bertrand competition, highly competitive oligopoly pricing. *See* Aghion & Bolton, *supra* note 2, at 396. Under Bertrand competition, consumers buy from the firm offering the lowest price, and each firm is willing to reduce its price to marginal cost if necessary. Thus, price falls to just under the marginal cost of the higher cost producer, the monopolist. The lower cost producer, the entrant, captures the entire market and has no incentive to reduce price further. Other models of small numbers competition exist, but this oligopoly model shows the incentive of the lower cost producer to enter even under intense rivalry.

12. More specifically, in the absence of a long-term contract, the customer faces a lottery with two possible outcomes: Either an entrant appears or does not appear. If an entrant *does* appear, the market becomes more competitive and price is driven down to $50. If an entrant does *not* appear, the incumbent continues to charge the monopoly price of $80. Thus, the customer's alternative to signing the long-term contract is to accept the lottery and purchase on the spot market at a future price that will be either $80 or $50. Assuming further that each of these two alternatives is equally probable, the expected value of the (risk neutral) customer's spot market alternative is $65 *(.5 X 80 + .5 X 50 = 65).* Thus, the customer will sign a long-term contract only if price does not exceed $65.

13. This description of the standard case of efficient contract breach is further illustrated in the following diagram, which is similar to the model constructed by Masten and Snyder. Scott E. Masten & Edward A. Snyder, *The Design and Duration of Contracts: Strategic and Efficiency Considerations,* 52 LAW & CONTEMP. PROBS. 63, 66–69 (1989). PRICE represents the rental price under the incumbent's existing contract, C(I) represents the incumbent's production costs, and C(E) represents the entrant's production costs.

FIGURE 1



The breaching customer must pay the incumbent its lost profit (PRICE - C(I)); however, because C(E) is less than C(I), the entrant can offer a price lower than C(I), which allows the customer to gain even after paying the lost profit.

Some analysts challenge the conclusion that the standard contract damage rule leads to an efficient result because legal remedies may not fully compensate the nonbreaching party in light of the uncertainty of recovery, time lag, and other inefficiencies. *See* Lisa Bernstein, *Opting Out of the Legal System: Extralegal Contractual Relations in the Diamond Industry,* 21 J. LEGAL STUD. 115, 135 (1992); Daniel Friedmann, *The Efficient Breach Fallacy,* 18 J. LEGAL STUD. 1, 23–24 (1989) (increased negotiation and litigation costs). However, these insights do not challenge the basic concept that breach is efficient if the gain exceeds the breaching party's losses. In our analysis, a penalty

efficient entrant's superior productivity is rewarded by economic profit earned from its cost advantage over the incumbent, providing the necessary incentive to induce risky investment.

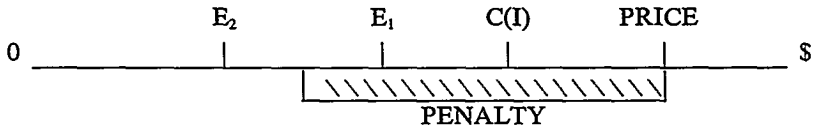
Thus, under both a spot market sale and a long-term contract without penalty, a lower cost entrant will displace a less efficient monopolist and earn economic rent because of its lower costs. An examination of long-term contracts *with* a penalty reveals adverse effects on efficiency and rent capture.

Long-term contract with penalty. If the incumbent monopolist offers customers a long-term contract requiring customers who switch to a rival supplier to pay stipulated damages *greater* than the incumbent's lost profit, a more efficient entrant is now disadvantaged despite its lower costs. Depending upon its costs, the entrant either loses much, if not all, of its expected profit or is excluded from the market altogether.

If the entrant's costs are low enough to offer customers a significantly reduced price, the customer will be able to pay the penalty and still gain from switching to the entrant. However, the entrant will lose some or perhaps all of its economic surplus because price will be reduced below the monopolist's costs by the amount of the penalty. If the entrant's costs are not low enough to provide switching customers with a gain after payment of the penalty, the entrant will be excluded from the market entirely.

The two effects of penalty contracts—entry deterrence and surplus extraction—are illustrated in the following diagram. PRICE represents the incumbent's contract price; C(I) represents the incumbent's cost; E_1 represents the cost of Entrant 1; E_2 is the cost of Entrant 2; and the penalty for switching to another supplier during the term of the contract is shown by the shaded area. As shown in the diagram, the penalty not only exceeds the

FIGURE 2



monopolist's lost profit from forgone sales (PRICE - C(I)), but is so large that Entrant 1 must reduce its price below its own cost to induce the customer to switch. The penalty produces an inefficient result because Entrant 1 will not enter even though its costs are below those of the incumbent. Although the penalty does not deter Entrant 2, whose costs are even lower than those of Entrant 1, Entrant 2 sacrifices much of its economic reward from entry. In the absence of a penalty, Entrant 2 would be able to enter the market by offering a price slightly below the incumbent's cost (C(I)). The penalty contract forces the entrant to reduce price substantially below the incumbent's cost in order to provide the customer with the means to pay the

contract exists when the stipulated damages exceed the promisee's losses, *realistically assessed. See* text accompanying note 141 *infra.*

penalty. Consequently, the monopolist receives the penalty amount in addition to its lost profit; the customer receives a price reduction at least equal to the penalty; but the more efficient entrant loses most of the anticipated profit from its lower costs savings.

The transfer of economic surplus from highly efficient entrants to an incumbent monopolist provides a striking insight from the Aghion and Bolton model: Under a penalty contract regime, the profit maximizing monopolist may actually welcome market entry by the highly efficient firm. Indeed, as shown by the shaded area in Figure 2, the incumbent gains more from the penalty paid when the customer switches to highly efficient Entrant 2 than from its profit under the original contract.[14]

Although it is initially unclear why a customer would be willing to sign a contract with a penalty, the previous analysis shows that the customer strictly benefits from signing. In the absence of a penalty contract, the entrant, however efficient, has no incentive to reduce its price below the incumbent's cost; the penalty forces the efficient entrant to reduce its price below this amount. This further reduction provides additional surplus, which the incumbent can share with the customer by reducing the customer's price.[15]

Customers may also be motivated to sign penalty contracts when the market has multiple customers and the entrant's product has fixed costs or economies of scale. Under these conditions, typical in modern industry, the individual customer confronts a dilemma in refusing to sign the contract. Unless enough other customers also refuse to sign, the entrant will not have access to enough customers to operate at efficient scale. If entry does not occur, there will be no supplier other than the monopolist. Thus, each customer finds it prudent to sign the agreement, while hoping that others will refuse and provide the necessary customer base for a new entrant. Each customer who signs the agreement increases the cost for the remaining customers by reducing the future entrant's base of available customers. Salop describes this effect as the "free rider in reverse."[16] The cumulative effect is

14. These results can be illustrated numerically. Assume under the long-term contract that the incumbent's price is $65, its cost is $50, and the switching penalty is $25. Entrant 1's cost is $45 and Entrant 2's cost is $35. Entrant 1 is excluded from the market because its cost plus the penalty ($70) exceeds the $65 contract price. Due to its lower costs, Entrant 2 may enter the market because it can undercut the $65 contract price, even with the penalty *($35 + $25 = $60).* However, Entrant 2's surplus is reduced from $15 without the penalty *($65 - $35 - $15 (damages) = $15)* to less than $5 with the penalty. The monopolist could exclude Entrant 2 by reducing price, but has no incentive to do so because the $25 penalty exceeds the $15 profit under the contract. Thus, the highly efficient entrant is able to enter the market, but loses most of its economic surplus through the penalty.

15. An objection to this analysis is that it neglects later time periods. Although both the customer and the monopolist achieve gains within the time frame of the model, circumstances may change in the long term. For example, in a longer time frame, the penalty payment the monopolist receives from an entrant might be offset by loss of market share when present contracts expire. In that scenario, the monopolist may gain more by preventing entry altogether, rather than seeking profit through the penalty contract. Nevertheless, consideration of distant time periods does not negate the previously demonstrated implications of the model. It merely indicates that a fully specified model must include some assessment of expected future losses from entry, discounted to present value. This alteration would complicate, but not modify, the basic insight that a penalty contract can increase the private welfare of the monopolist and its customers.

16. Steven C. Salop, *Practices that (Credibly) Facilitate Oligopoly Coordination, in* NEW DE-

that all customers tend to sign the agreement, thereby blocking entry. The customer's dilemma intensifies if the monopolist offers special inducements to customers who sign the agreement when others refuse.[17]

In sum, a penalty contract provides a strategically effective way for a coalition between a monopolist and its customers to exploit their bargaining power and extract rents from a more efficient producer. Although the penalty has private advantages for the monopolist and its customers, such clauses result in social inefficiency by deterring lower cost entrants and perpetuating monopoly.[18] Furthermore, the penalty allows the monopolist to capture the economic surplus of the most efficient entrants, who are not deterred despite the penalty, and thus reduces the profitability of entry in the industry. This reduced profitability discourages the substantial investment in fixed costs required for research and development (R&D). Consequently, over time firms will invest less in R&D projects, thereby reducing innovation and forestalling the possibility of new entry into monopolized markets.[19]

B. *Assumptions and Possible Objections to the Aghion-Bolton Model*

The Aghion-Bolton model necessarily rests on simplifying assumptions. This raises the issue of whether these assumptions limit the model as a policy guide. The most questionable assumptions are the presumed ability of a new entrant to earn economic rent, symmetric information, and lack of negotiation alternatives that might reduce the harmful effects of penalty contracts.

VELOPMENTS IN THE ANALYSIS OF MARKET STRUCTURE 265, 272–73, 278, 284 (Joseph E. Stiglitz & G. Frank Mathewson eds., 1986); *see* Aghion & Bolton, *supra* note 2, at 396–98; Eric B. Rasmusen, J. Mark Ramseyer & John S. Wiley, Jr., *Naked Exclusion,* 81 AM. ECON. REV. 1137 (1991).

The free-rider-in-reverse effect asserts that self-interested parties, unable to make binding agreements, will fail to achieve an efficient outcome. Even though an outcome preferred by all parties can be made possible through collective action, each individual privately prefers to avoid the costly action necessary to implement the collective optimum, thus making the outcome unsustainable. In our context the entrant may need a substantial customer base to obtain scale economies, and if all, or a substantial number of, customers can agree to switch, significant cost savings can be achieved. However, uncertain about other customers' choices, an individual customer may prefer to wait to see if the entrant can build up a sufficient customer base. Waiting avoids the risk that were the entrant unsuccessful, a switching customer would be at the mercy of the incumbent monopolist. But when most of the customers are withholding their participation, the entrant does not have enough customers. So although collectively customers would like to exploit the entrant's technological advantage, each of them lacks the individual incentive to enter the entrant's new customer base.

17. In this case, the agreement may be characterized as a full prisoner's dilemma in which signing the agreement is the dominant strategy, although each buyer would prefer not to sign if others would act similarly. *See* Eric Rasmusen, *Recent Developments in the Economics of Exclusionary Contracts, in* CANADIAN COMPETITION LAW AND POLICY AT THE CENTENARY 371, 380–83 (R.S. Khemani & W.T. Stanbury eds., 1991).

18. *See* Tai-Yeong Chung, *On the Social Optimality of Liquidated Damage Clauses: An Economic Analysis,* 8 J.L. ECON. & ORGANIZATION 280, 283 (1992).

19. We are not assuming that there is too little investment in R&D. We simply assert that the incentives for R&D should not be regulated through Aghion-Bolton penalties against new entrants. Rather, the proper level of R&D should be adjusted by general incentives, such as patents, taxes, or other subsidies. Accordingly, Congress has recognized a general need to stimulate R&D investment through cooperative research. *See* National Cooperative Research Act of 1984, Pub. L. No. 98-462, 98 Stat. 1815 (codified as amended at 15 U.S.C. §§ 4301–4305 (1988)).

However, none of the objections undermines the force of the model or its utility as a policy guide.

1. *The entrant's profitability.*

The key assumption of the Aghion-Bolton model is that, absent a penalty contract, an efficient entrant can earn economic profit in competition with the higher cost incumbent. This assumption is critical because a penalty contract could not deprive an entrant of economic profit if the entrant could not earn any profit initially. Clearly, the assumption is satisfied if there is only a single entrant. When there is only one entrant, price will not fall below the monopolist's cost even under the highly competitive pricing of Bertrand competition, where all customers immediately switch to the lower cost supplier.[20] Thus, even under the most competitive regime, the entrant earns economic profit equal to its cost advantage over the incumbent.

If two or more equally efficient entrants enter the market simultaneously, it appears that the entrants will earn no economic profit because competition between the entrants will lower price to the entrants' marginal cost. However, this is an exceedingly unlikely scenario. The preclusion of economic profit by multiple entry holds only under a set of very special conditions: Entrants are equally efficient; entry is simultaneous; the products are homogeneous; no capacity constraints limit an entrant's ability to supply the whole market; and the firms engage in no express or tacit collusion. In the absence of any one of these conditions, prices in oligopoly markets are likely to range somewhere between full monopoly pricing and rigorously competitive pricing.[21] While these conditions may be satisfied over the long term, multiple entry is unlikely to occur without significant time lag.[22] Indeed, an entrant would not undertake risky investment in new entry unless it anticipates a steady stream of economic profit over a significant time period following successful entry. Moreover, even simultaneous entry will not eliminate economic profit if entry involves significant product differentiation, which is typically the case for technological advances.[23] Thus, in most industrial markets, a potential entrant with lower costs will earn economic profit over an appreciable period; hence, the key assumption of the Aghion-Bolton model is satisfied.

20. More precisely, price will fall to the cost of the next to lowest cost producer, but no lower. The lowest cost producer, here the entrant, does not have any competition below that price, and thus has no economic incentive to reduce price below that level. *See* note 11 *supra.*

21. *See* JEAN TIROLE, THE THEORY OF INDUSTRIAL ORGANIZATION 218–21 (1988).

22. Even if firms are equally efficient, their incentives to enter a particular market may differ due to financial constraints or diminishing returns to the diffusion of managerial effort. *See* Aghion & Bolton, *supra* note 2, at 390; Joseph E. Stiglitz & Andrew Weiss, *Credit Rationing in Markets with Imperfect Information,* 71 AM. ECON. REV. 393 (1981); *cf.* Masten & Snyder, *supra* note 13, at 73 (explaining that differential barriers to entry may prevent simultaneous entry).

23. Sequential entry by firms of *differing* efficiency will not eliminate profitability because the most efficient firm has no incentive to reduce price below the level necessary to undersell its rivals, with the lowest level being slightly less than the production cost of the *second* most efficient firm. In that event, the most efficient entrant sets a price above its own marginal cost and earns economic profit.

A more realistic alternative—sequential entry—does not significantly affect the Aghion-Bolton argument. If customers expect that several firms will enter the market sequentially, then the customers will anticipate that the spot market price will fall over time. As we have seen under Bertrand competition, the most efficient firm will undercut the price of all less efficient firms, but will not lower price to its own marginal cost. The firm has no incentive to reduce price below that necessary to drive out the other firms, the lower limit being slightly less than the production cost of the second most efficient firm. The surviving entrant earns economic profit because its price is above its marginal cost. In this scenario, the incumbent and its customers can still form a coalition to extract some of the low cost entrant's rent. Thus, the alternative of sequential entry by firms of varying efficiency does not change the conclusions of the Aghion-Bolton model.[24]

2. *Symmetric information.*

The Aghion-Bolton model makes two informational assumptions concerning a future entrant's costs. First, the model assumes that the incumbent and its customers have imperfect knowledge of the entrant's costs. Second, the model assumes that such information as the incumbent and its customers do possess is identical. The assumption of imperfect knowledge of the entrant's costs is realistic, because at the time the penalty contract is signed the entrant is typically unknown and its plant unbuilt. The assumption of identical information is less secure. An incumbent engaged in manufacturing, for example, might have better information than its customers on the costs of entry into its own market.

Thus, Aghion and Bolton extend their analysis to the asymmetric information case, where a monopolist is better informed than its customers about an entrant's costs. Under these circumstances, an incumbent facing a low cost entrant will have an incentive to write a long-term penalty contract, while an incumbent facing a high cost entrant will not have similar incentives.[25] However, when the customers are aware that the monopolist has superior information, they may interpret the incumbent's actions as an indi-

24. Another possible objection to the Aghion-Bolton analysis might be grounded in its implicit assumption of imperfect competition at the customer-reseller level. If perfect competition and free entry dominated the customer-reseller market, then the efficient entrant into the supply market could bypass any contract regardless of penalty by selling to customers not bound by contracts. Indeed, under a free entry assumption, the entrant could itself enter the customer-reseller market to avoid customers bound by penalty contracts. Both of these cases may apply to an extreme long-run situation where relatively free entry and vertical integration can occur; however, entry in the short run is never likely to be simple and easy. Thus, a long period may elapse after an entrant has appeared before a significant number of new customers consider entry. Similar impediments also confront integration. Beginning production in both upstream and downstream markets requires more preparation and investment than entry at only one level. Furthermore, integration is preferable to single market entry only if integration achieves significant scale economies.

25. If the incumbent knows that the entrant's costs are likely to be low, more can be gained from extracting rent or deterring entry through an anticompetitive penalty contract. However, if the potential entrant's costs are likely to be high, then the threat of entry diminishes; the incumbent will prefer not to sign a penalty contract and instead will pursue short-term profit maximization.

cation of the entrant's cost. If the incumbent anticipates such a response by customers, complex signaling strategies may emerge.[26] Asymmetric information merely makes the analysis more complex, but does not alter the essential result of the Aghion-Bolton model. To the extent such contracts are used, they retain their anticompetitive potential, even if assessment of the magnitude of the effect may be less straightforward.

3. *Negotiation alternatives.*

The Aghion-Bolton analysis does not consider whether the entrant could negotiate around the adverse effects of the penalty contract. If the entrant could negotiate directly with customers, a low cost entrant could prevent customers from signing a penalty contract by offering them a price below that of the incumbent. The model justifiably excludes this possibility, however, because the customers sign the penalty contract before the entrant has appeared. After the entrant has appeared, the customers are already bound by the penalty contract and negotiation to prevent signing is not possible.

Negotiations by the entrant with the incumbent offer little hope of overcoming the inefficiencies of the penalty contract. The entrant cannot negotiate with the incumbent before the penalty contract is signed for the same reason as above, namely, because the entrant has not yet appeared. Ex post negotiation between the entrant and the incumbent might appear more hopeful for bypassing the penalty contract.[27] There are three possibilities: The incumbent might (1) acquire the entrant, (2) "delegate" production to the entrant through purchase of its output or by forming a joint venture, or (3) reduce the penalty by an amount sufficient to allow entry. But none of these alternatives appears likely to prevent loss of social efficiency.

The first two alternatives, acquisition of entrant or delegation of production, raise serious antitrust issues since they foreclose competition in a monopolized market. In addition, both acquisition and delegation of production are likely to involve transactional inefficiencies. Acquisition effects a change of ownership, causing loss of high powered ownership incentives within the acquired firm and possible loss of economies of scope since an entrant will frequently make other products.[28] Delegation of production to entrant through purchase of its output is not feasible if sale of the product requires special services or know-how that the incumbent lacks and the entrant is unable or unwilling to transfer.[29] Joint ventures are less vulnerable

26. For example, the customer may interpret the incumbent's offer of a long-term penalty contract as a signal that there is a low cost potential entrant. Realizing the likelihood of such a response, the incumbent may shorten the duration of the contract or reduce the penalty to mute the signal. The result reduces, but does not remove, the Aghion-Bolton effect.

27. *See* Masten & Snyder, *supra* note 13, at 71–72.

28. In addition, a merger may not even be feasible if the entrant produces products unrelated to the incumbent's product line or its strategic plan.

29. For the incumbent to successfully market the entrant's output, the products must be homogeneous and technologically similar. However, most new and superior products involve innovative technology. If selling a product includes a transfer of technological know-how to the buyer, the incumbent cannot simply "substitute" the entrant's product for its own. The incumbent may lack

to these problems, but confront difficulties in management and valuation of the inputs of the participants.[30]

Moreover, acquisition or delegation of production will not prevent an Aghion-Bolton effect. The existence of the penalty contract reduces the alternative value of independent entry, and this necessarily lowers the buyout price. Thus, the buyout of the entrant in the shadow of the penalty contract simply becomes an alternative means of extracting the penalty. Rather than pay the penalty by sales to customers at low prices, the entrant sells out to the incumbent at a penalty-reduced price. The effect on investment incentives remains adverse and competition is suppressed in a monopolized market.

The third alternative, reduction of the penalty to allow entry, will not eliminate the adverse effect of penalty contracts. The penalty contract sets a status quo that favors the incumbent: Highly efficient entrants enter the market after paying the penalty, while relatively less efficient entrants are deterred. Since the penalty exceeds the incumbent's previous monopoly profit, it might appear that the incumbent would always be willing to renegotiate the penalty to allow entry. But if the incumbent allows entry by firms with only a moderate cost advantage, it would trade a small rent capture gain in the current contract period for a loss of monopoly profit in all future periods. Thus, unless the penalty substantially exceeds the current monopoly profit, the incumbent is likely to do best by deterring entry and retaining its monopoly profit in future periods. As a result, the incumbent will be unwilling to reduce the penalty, and negotiation will not occur.

Even if the incumbent reduces the penalty to allow entry, incentive effects remain adverse. In allowing entry, the incumbent will attempt to extract as much as possible of the entrant's surplus. While the reduced penalty shifts production to the more efficient entrant, the entrant loses much, if not all, of the economic rent that motivates risky new entry, reducing its incentive to make future investment in the market; and the incumbent earns additional monopoly profit, increasing unproductive rent-seeking.

In sum, a consideration of negotiation alternatives does not modify the conclusions and policy implications of the Aghion-Bolton analysis.

II. PENALTIES UNDER CONTRACT AND ANTITRUST LAW

Courts deciding antitrust cases have not considered the Aghion-Bolton model in their analysis, nor have they developed an effective alternative theory for evaluating contract penalties. In *United Shoe*, the penalty provisions in the long-term leases were part of a broader array of practices, including price discrimination, tying arrangements, mergers, and deliberate prevention

---

the knowledge to service or maintain the entrant's product, and the costs associated with mastering the new technology could be prohibitive.

30. *See generally* Joseph F. Brodley, *Joint Ventures and Antitrust Policy*, 95 HARV. L. REV. 1521 (1982).

of a second-hand market.[31] Thus, *United Shoe* may be read as holding only that the penalties in the long-term leases were part of a web of practices, which together unlawfully excluded competition on a basis other than superior economic performance.[32] As a result, the decision could not serve as a clear precedent for evaluating specific practices.

In recent years, however, penalty contracts have been challenged in several cases in the lower federal courts. The courts have consistently rejected such challenges, relying primarily on the contract law rule against penalties as the appropriate antitrust standard. Thus, courts have held that a stipulated damages provision enforceable under contract law is also valid under antitrust law.[33]

In contrast to the reluctance of U.S. antitrust courts to accept challenges to penalty contracts, foreign courts have held in two recent decisions that penalty contracts can be anticompetitive.[34] Understandably, none of the decisions refers to the Aghion-Bolton analysis, which was neither available when these cases were decided, nor presented in a form accessible to most legal readers. We believe that the Aghion-Bolton analysis provides an effective method for analyzing antitrust cases. In this Part we reexamine some of these cases using the Aghion-Bolton model as a guide.

We conclude first that contract law is highly permissive in its approach to contract penalties, both in common law and civil law regimes, and that contract law would become even more permissive if the penalty rule were abandoned, as many law and economics scholars recommend. Second, such permissiveness, while perhaps appropriate for contract law, does not address the concerns of antitrust law regarding contract penalties in monopolistic markets. Third, courts have developed no coherent legal theory for evaluating penalty contracts under antitrust law, and the Aghion-Bolton analysis demonstrates that many of the criteria courts have used are mistaken. Fourth, recent antitrust decisions provide instructive examples for the proper application of the Aghion-Bolton analysis and the selection of better criteria for future antitrust decisionmaking.

A. *Inadequacy of the Contract Law Standard*

Contract law has historically refused to enforce a stipulated damage provision that constitutes a "penalty."[35] Thus, contract law might appear to preclude the penalty contract that drives the Aghion-Bolton model. Indeed, respected courts and commentators have stated that the contract law rule

31. The legal issue was further clouded by an earlier Supreme Court decision upholding the validity of United Shoe's leases under the Sherman Act. United States v. United Shoe Mach. Co. of New Jersey, 247 U.S. 32 (1918).

32. *See* Williamsburg Wax Museum, Inc. v. Historic Figures, Inc., 810 F.2d 243, 253 (D.C. Cir. 1987) (stating that the *United Shoe* decision rested on "a panoply of practices").

33. *See, e.g.*, Barry Wright Corp. v. ITT Grinnell Corp., 724 F.2d 227, 239 (1st Cir. 1983).

34. *See* Elopak Italia Srl v. Tetra Pak, 4 C.M.L.R. Antitrust Rep. 551 (E.C. Comm'n 1991); *In re* Laidlaw Waste Sys. Ltd., Canadian Competition Tribunal, Jan. 20, 1992.

35. *See* 3 E. ALLAN FARNSWORTH, FARNSWORTH ON CONTRACTS § 12.18 (1990).

sets the appropriate standard for antitrust law.[36] However, these cursory assertions have not fully examined the limitations of contract law's treatment of penalty contracts in the antitrust context. Contract law provides a wholly inadequate standard for antitrust evaluation of penalty contracts, whether one adheres to: (1) the prevailing judicial approach in the United States that claims to deny enforcement of penalty contracts; (2) the law and economics approach, which would generally enforce penalty contracts as economically efficient; or (3) the civil law rule, applicable in most EEC countries, as well as in Japan, Russia, and Switzerland, which enforces penalty contracts unless they are unconscionable.

1. *The prevailing U.S. approach.*

Most U.S. courts assume that although contracting parties have incentives to write penalty contracts, such conduct undermines the strong public policy favoring purely compensatory contract remedies.[37] The courts regulate these undesirable motives by refusing to enforce excessive stipulated damage provisions, declaring such provisions to be void as penalties.[38] Despite the rhetoric of the penalty rule, courts enforce many contracts where rigorous judicial scrutiny would reveal that stipulated damages exceed actual damages.[39] Several reasons might explain the courts' permissive attitude.

First, the current standard used to assess stipulated damages, which combines in a single test both prior anticipations of loss and after-the-fact experience, is both vague and contradictory.[40] Courts using the standard generally enforce stipulated damages in commercial cases, rather than deter-

36. *See, e.g., Barry Wright,* 724 F.2d at 238–39 (asserting that a stipulated damages provision valid under contract laws is reasonable under antitrust law); Masten & Snyder, *supra* note 13, at 71–73 (concluding that illegality of penalties in contract law undermines their effective strategic use); Oliver E. Williamson, *Delimiting Antitrust,* 76 GEO. L.J. 271, 283 (1987) (unreasonably large stipulated damages would be void as a penalty under contract law).

37. *See* Lake River Corp. v. Carborundum Co., 769 F.2d 1284, 1289 (7th Cir. 1985) (noting that penalty clauses increase risks to other creditors, increase bankruptcies, and can amplify the business cycle by increasing bankruptcies in economic downturns when contracts are more likely to be broken).

38. *See, e.g., id.* at 1290–92; Brecher v. Laikin, 430 F. Supp. 103 (S.D.N.Y. 1977).

39. *See* 3 FARNSWORTH, *supra* note 35, § 12.18. Farnsworth notes a trend towards enforcement of stipulated damages provisions and suggests that the development of the doctrine of unconscionability has made it "increasingly difficult to justify the peculiar historical distinction between liquidated damages and penalties." *Id.*

40. The liquidated damages provisions are enforced only if "reasonable in the light of the anticipated or actual loss caused by the breach." RESTATEMENT, *supra* note 3, § 356(1). Reasonableness is measured by the "difficulty of proof of loss." *Id.* cmt. b. Reasonableness is examined both at the time of contract formation and the time of loss. *See* 3 FARNSWORTH, *supra* note 35, § 12.18 (noting courts' difficulty distinguishing liquidated damages from penalties). The greater the difficulty of proof, the greater the latitude to be given to the stipulated damage provision. *See* RESTATEMENT, *supra* note 3, § 356(1) cmt. b. Thus, the reasonableness standard requires that the parties have information sufficient to show that the predicted damages were reasonable, while also requiring that the parties not have information sufficient to make those damages easy to prove. *See id.*; *see also* Samuel A. Rea, Jr., *Efficiency Implications of Penalties and Liquidated Damages,* 13 J. LEGAL STUD. 147, 150 (1984) (describing courts' inconsistency in distinguishing liquidated damages from penalties); 2 ROY RYDEN ANDERSON, DAMAGES UNDER THE UNIFORM COMMERCIAL CODE §§ 13:01–13:05 (1991) (U.C.C. § 2-718 confusing and difficult to apply).

mine damages under such a convoluted approach. Courts readily justify enforcement of stipulated damages when the agreement was "made by knowledgeable, sophisticated businessmen in the industry on both sides acting at arm's length without any coercion or inability to act freely."[41]

Second, the trend toward increased enforcement of stipulated damages is also encouraged by a shifting of the burden of proof to the party who asserts the existence of an unlawful penalty. The shifted burden of proof, enacted by statute in some states,[42] has probably now become the majority rule,[43] replacing the earlier rule requiring the enforcer of a contract to prove the *absence* of an unlawful penalty. In commercial litigation, typically involving complex issues of assessing lost profit expectancies, the party who bears the burden of proof clearly faces a significant disadvantage. Thus, the change in the burden of proof increases the difficulty of invalidating stipulated damage provisions.

Third, contracting parties can adopt many alternative covenants to mask penalty provisions and circumvent the penalty rule.[44] While courts will strike down "disguised penalties," the continued use and variety of alternative covenants suggest that courts do not uncover all hidden penalties. For example, some courts hold that alternative performance agreements are not considered penalties, provided it is conceivably possible that either alternative might be chosen at the time of performance.[45] Similarly, agreements on payment of attorneys' fees, avoidable damages, valuation of performance conditions, security deposits, and a variety of other terms may escape penalty classification while having economic effects similar to those of contract penalties.[46]

Finally, many commentators advocate increased enforcement of stipulated damage contracts with limitation or even outright abrogation of the contract law penalty rule. Advocates of such reform include not only law and economics scholars,[47] but also those writing from more general

41. United Air Lines, Inc. v. Austin Travel Corp., 681 F. Supp. 176, 188 (S.D.N.Y. 1988), *aff'd*, 867 F.2d 737 (2d Cir. 1989).

42. *See, e.g.*, CAL. CIV. CODE § 1671(b) (West 1985) (stipulated damage provision valid unless party seeking to invalidate shows provision to be "unreasonable under the circumstances existing at the time the contract was made"); *cf.* Wassenar v. Panos, 111 Wis. 2d 518, 539, 331 N.W.2d 357, 367 (1983) (burden of proof on party challenging contract to prove that stipulated damage is "grossly disproportionate to the actual harm and thus unreasonable").

43. 25A C.J.S. *Damages* § 144(f) (1966) (majority rule places burden of proof on party seeking to show "that the intention of the parties was other than appears from the face of the contract").

44. *See* Frank C. Dunbar, Jr., *Drafting the Liquidated Damage Clause—When and How*, 20 OHIO ST. L.J. 221, 230 (1959) ("Contracting parties' lawyers have been ingenious . . . in cloaking penalty provisions in other forms, sometimes with success in the courts."); Justin Sweet, *Liquidated Damages in California*, 60 CAL. L. REV. 84, 90–94 (1972) (listing contractual clauses controlling damages, such as security deposits, alternative performances, and limitations on liability).

45. Chandler v. Doran Co., 44 Wash. 2d 396, 401–02, 267 P.2d 907, 910–11 (1954) (quoting 3 SAMUEL WILLISTON, A TREATISE ON THE LAW OF CONTRACTS § 781 (Rev. ed. 1936)).

46. *See, e.g.*, 3 FARNSWORTH, *supra* note 35, § 12.18 (asserting that attorneys' fees and "other legal expenses" are usually not classified as disguised penalties); Sweet, *supra* note 44, at 90–91 (cataloging variety of forms liquidated damage clauses may take).

47. *See, e.g.*, Charles J. Goetz & Robert E. Scott, *Liquidated Damages, Penalties and the Just Compensation Principle: Some Notes on an Enforcement Model and a Theory of Efficient Breach*, 77

frameworks of analysis.[48] Indeed, these recent developments suggest that U.S. law may be moving closer to the civil law rule, under which penalty contracts are enforceable unless they are held to be unconscionable.[49]

2. *The law and economics perspective.*

Under the law and economics approach to penalties, contracting parties in competitive markets do not intentionally write penalty contracts. Such contracts would reduce the transactional surplus from trade, binding the parties when the costs of performance exceed their expected gains. Parties have no incentive to enter into such contracts, which would result in their exclusion from competitive markets as high cost producers. Hence, what appears to a court as a penalty is most likely a hidden loss suffered by the nonbreaching party that escapes judicial recognition. Since contracting parties do not purposely enter into penalty contracts in competitive markets, law and economics scholars believe that courts should enforce stipulated damage provisions, unless such provisions are unconscionable.[50] Full acceptance by courts of this view, however, would not solve the problem exposed by Aghion and Bolton, but in fact would make it worse!

The law and economics view of penalty contracts magnifies the antitrust problem by failing to consider the full range of strategic interactions in monopolized markets. While law and economics writers emphasize the gains from bilateral or multilateral transactions among contract participants, they have until recently ignored strategic interactions between contract participants and outsiders.

Under the Aghion-Bolton analysis, parties who enter into penalty contracts engage in a predatory, exploitive bargaining strategy against a third party, generally a subsequent market entrant, to increase their joint welfare.[51] Although the contract enhances the contracting parties' joint surplus, it reduces net social surplus, including the welfare of third parties.[52]

COLUM. L. REV. 554, 594 (1977) (concluding that the unconscionability doctrine is a less costly alternative to the broad invalidation powers of the penalty rule); Alan Schwartz, *The Myth that Promisees Prefer Supracompensatory Remedies: An Analysis of Contracting for Damage Measures,* 100 YALE L.J. 369, 370 (1990) (asserting that both the ex ante and ex post branches of liquidated damages rules should be abandoned). *But see* Masten & Snyder, *supra* note 13, at 83 (noting that "invalidation of contract penalties is a sensible legal precept").

48. *See, e.g.,* MARVIN A. CHIRELSTEIN, CONCEPTS AND CASE ANALYSIS IN THE LAW OF CONTRACTS 172 (1990) (arguing that the best approach is the unconscionability doctrine); DAN B. DOBBS, HANDBOOK ON THE LAW OF REMEDIES § 12.5 (1973) (claiming that, except for adhesion contracts, enforcement of liquidated damages clauses should not be restricted); 3 FARNSWORTH, *supra* note 35, § 12.18 (noting that it is "increasingly difficult to justify the peculiar historical distinction between liquidated damages and penalties").

49. Unconscionability is not well-defined, but the basic concept, involving both procedural and substantive unfairness, appears to focus on uninformed or powerless consumers, not commercial actors. *See* CAL. CIV. CODE §§ 1671(c), 1671(d) (West 1985); CHIRELSTEIN, *supra* note 48, at 74, 77; Schwartz, *supra* note 47, at 383–84.

50. *See generally* Schwartz, *supra* note 47.

51. *See generally* Chung, *supra* note 18 (extending the analysis to contracts having externality effects on third parties).

52. Thus, as we saw in the Intuitive Fable, the monopolist and its customers both gained from

Hence, a permissive penalty regime, although economically desirable in competitive markets, would provide fertile soil for anticompetitive Aghion-Bolton contracts to flourish in monopoly markets.

3. *The civil law regime.*

Under the civil law applicable in most EEC countries and in commercially vital jurisdictions such as Japan, Russia, and Switzerland, contract penalties exceeding actual losses are lawful unless held to be unconscionable. These regimes recognize that penalties may serve a valid *in terrorem* function and may sometimes be enforced even if no actual damages are suffered.[53] For instance, Swiss law explicitly provides that payment is due even if the nonbreaching party has suffered no damages.[54] Moreover, in Germany and Japan, there are no statutory or other explicit limitations on the size of the penalty in commercial cases.[55]

Furthermore, statutory formulations of the unconscionability limitation seem to contemplate payments exceeding actual damages. Thus, these statutes require manifest or startling unconscionability, indicating that only grossly excessive penalties are proscribed. The French statute allows modification of the penalty only if it is "manifestly excessive" or "manifestly derisory";[56] the Italian law contains similar language;[57] and the Russian statute permits reduction of the penalty only when it is "extraordinarily large in comparison with the creditor's actual losses."[58] Moreover, even if a European court finds unconscionability, reduction of the penalty generally remains subject to judicial discretion.[59]

Evidently, non-common law commercial jurisdictions have followed a permissive approach toward penalty contracts, approximating the views of

---

the penalty contract, but society was denied the full benefits of lower cost production, and the entrant was either excluded or lost much of its return from superior economic performance.

53. *See generally* INTERNATIONAL CHAMBER OF COMMERCE, GUIDE TO PENALTY AND LIQUIDATED DAMAGES CLAUSES (1990).

54. G.H. TREITEL, REMEDIES FOR BREACH OF CONTRACT: A COMPARATIVE ACCOUNT 218 (1988) (citing translation of Swiss Code of Obligations art. 161, ¶ 1).

55. *See* 3 ZENTARO KITAGAWA, DOING BUSINESS IN JAPAN § 1.15[5] (1990); BERND RUSTER, BUSINESS TRANSACTIONS IN GERMANY (FRG) § 10.05[1][d] (1992). In addition, Article 420 of the Japanese Civil Code states that courts cannot increase or decrease the stipulated damage amount. 3 KITAGAWA, *supra,* § 1.15[5]. In Germany, penalty clauses are in general valid and enforceable. While German courts have discretion to modify penalties if the penalty is disproportionately high and has not been paid, this provision is inapplicable to commercial transactions entered into by merchants. *See* TREITEL, *supra* note 54, at 223, 227–28. The French law contained no explicit constraint on penalty clauses until 1975, when an unconscionability condition was adopted. *Id.* at 220.

56. TREITEL, *supra* note 54, at 220, 224 (citing translation of French Civil Code art. 1152, ¶ 2). The French court's power to reduce the stipulated sum, conferred by French Civil Code art. 1152, ¶ 2, is an exception to the general rule of art. 1152, ¶ 1, which requires literal enforcement of the prior damage agreement.

57. *See* 1 LOUIS F. DEL DUCA & PATRICK DEL DUCA, COMMERCIAL, BUSINESS AND TRADE LAWS: ITALY § 5, ¶ 1384 (1983).

58. 2 CHRISTOPHER OSAKWE, SOVIET BUSINESS LAW: INSTITUTIONS, PRINCIPLES & PROCESSES app. at 64 (1991) (citing translation of GK RSFSR art. 190).

59. *Id.*; *see also* DEL DUCA & DEL DUCA, *supra* note 57, § 5, ¶ 1384; TREITEL, *supra* note 54, at 223–24.

the law and economics scholars. Thus, even if the common law rule effectively prevented penalties, the Aghion-Bolton analysis would remain relevant to the development of competition law in most of the industrial world, as well as to international trade subject to the laws of these foreign jurisdictions.

B. *Recent Antitrust Decisions*

In several recent antitrust cases, U.S. courts have reviewed alleged penalty provisions. The courts have rejected the antitrust claims, without considering the Aghion-Bolton analysis. Absent that analysis, the opinions in these cases lack consistency, as evidenced by the variety of grounds on which the courts rejected antitrust claims. The courts dismissed antitrust claims for the following reasons: (1) The stipulated damages provision was not a penalty;[60] (2) stipulated damages provisions that were considered penalties would not be enforceable under contract law;[61] (3) entry was occurring despite the penalty, thereby reducing the dominant firm's market share;[62] (4) the defendant offered a short-term lease or sale alternative;[63] (5) the penalty was a reasonable response to terms offered by rivals;[64] (6) the duration of the contract did not exceed the 5-year period allowed under the *United Shoe* decree;[65] (7) an industry-wide prohibition of penalties or a shortening of the lease term would impede new entry;[66] (8) a lease was no more exclusionary than a sale;[67] or (9) the defendant lacked monopoly power.[68] In addition, defendants have argued for legality based on customer preference for such contracts.[69]

The courts have not suggested how these diverse criteria are to be combined, or how any one element is to be balanced against another. The range of the factors considered indicates absence of a coherent theory of how penalty contracts injure competition. Moreover, many of these specific justifications are mistaken in light of the Aghion-Bolton analysis and the realities of modern contract law.

Thus, the assertion that contract law prohibits penalty contracts neglects the permissiveness of modern contract law in allowing penalties, both in the

60. *See* United Air Lines, Inc. v. Austin Travel Corp., 867 F.2d 737, 740 (2d Cir. 1989).

61. *See* Barry Wright Corp. v. ITT Grinnell Corp., 724 F.2d 227, 239 (1st Cir. 1983).

62. *See* Telex Corp. v. IBM, 510 F.2d 894, 905 (10th Cir.), *cert. dismissed,* 423 U.S. 802 (1975).

63. United States v. United Shoe Mach. Corp., 110 F. Supp. 295, 314, 350–51 (D. Mass. 1953), *aff'd per curiam,* 347 U.S. 521 (1954).

64. *Austin Travel,* 867 F.2d at 740–41 (stipulated damages not excessive when competitors required higher payments); *Telex,* 510 F.2d at 920 (competitors used longer leases).

65. *Telex,* 510 F.2d at 920.

66. *In re* "Apollo" Air Passenger Computer Reservation Sys. (CRS), 720 F. Supp. 1068, 1076 (S.D.N.Y. 1989).

67. Northeastern Tel. Co. v. AT&T, 651 F.2d 76 (2d Cir. 1981), *cert. denied,* 455 U.S. 943 (1982).

68. *Austin Travel,* 867 F.2d at 742 (finding 31% control of national market insufficient to constitute a national monopoly).

69. *See* CARL KAYSEN, *UNITED STATES V. UNITED SHOE MACHINERY CORPORATION*: AN ECONOMIC ANALYSIS OF AN ANTI-TRUST CASE 202 (1956).

United States and in foreign jurisdictions. The fact that entry is occurring despite the penalty neglects the teaching of Aghion and Bolton that a monopolist may allow entry by highly efficient producers in order to capture economic rent through the penalty. The preference of customers for long-term penalty contracts is not exculpatory, but rather is predicted by the Aghion-Bolton analysis, because customers may willingly sign penalty contracts to obtain a share of the benefits. A 5-year limit on duration of the contract does not validate the penalty because the significance of contract length depends upon industry characteristics, such as product lead time and technological life, as discussed in Part III. The assertion that a lease is less exclusionary than a sale is correct but irrelevant, because a lease *with a penalty* can be much more exclusionary than a sale.[70]

In contrast to the U.S. decisions, foreign courts have recognized in two recent decisions the anticompetitive effects of penalty contracts within the context of larger antitrust violations. These and other recent decisions from both domestic and foreign jurisdictions provide highly instructive applications for the Aghion-Bolton analysis, with examples of deterrence and rent capture effects, as well as the deficiencies of judicial reasoning. In the discussion of these cases our goal is to explain how contract penalty analysis applies to antitrust cases. In Part III, we develop detailed criteria for evaluating the antitrust implications of penalty contracts.

The key for proper application of the Aghion-Bolton analysis is to recall that a mutually beneficial bargain between the incumbent monopolist and its customers can be made only if the prospects for entry are uncertain at the time of bargaining. If entry is certain, or has already occurred, the customer will not join the coalition, preferring instead to bargain directly with the entrant. Only in cases of entry uncertainty or *potential entry* does a customer benefit by signing a penalty contract. In examining the recent cases, we divide our discussion into cases involving actual entry, potential entry, and cases involving both actual and potential entry.

1. *Actual entry*: Telex v. IBM.

In *Telex*,[71] and several similar cases, manufacturers of peripheral devices for computers challenged stipulated damage provisions in IBM leases.[72] IBM had previously used 30-day leases, but in response to the success of

70. Other reasons offered for upholding penalties have greater validity, but frequently require qualification. Justifying the penalty as a response to competitive terms offered by rivals is reasonable in a competitive market, but is less reasonable when the "competitive" interaction is between duopolists. Similarly, the assertion that industry-wide prohibition of stipulated damages could impede new entry may be valid if an entrant is attempting to set stipulated damages *below* contract damages. In that case, a ceiling on damages may assure nervous customers the option of shifting to another supplier if dissatisfied. However, that justification does not apply where a penalty contract sets stipulated damages at more than contract damages.

71. Telex Corp. v. IBM, 510 F.2d 894 (10th Cir.), *cert. dismissed,* 423 U.S. 802 (1975).

72. Greyhound Computer Corp. v. IBM, 559 F.2d 488, 498–99 (9th Cir. 1977), *cert. denied,* 434 U.S. 1040 (1978); ILC Peripherals Leasing Corp. v. IBM, 458 F. Supp. 423 (N.D. Cal. 1978), *aff'd sub nom.* Memorex Corp. v. IBM, 636 F.2d 1188 (9th Cir. 1980), *cert. denied,* 452 U.S. 972 (1981).

rival peripheral equipment manufacturers, adopted one- and 2-year leases. These longer leases were subject to a cancellation penalty of either five times the monthly rental (for a 2-year lease canceled during its first year), or two and a half times the monthly rental (for any lease canceled in its final year). IBM reduced its monthly rental charges to induce customers to sign the new leases. The 30-day lease and a sale alternative remained available, but without any price reduction. The court upheld the IBM leases on several grounds, but primarily because IBM's new leasing terms, including the cancellation penalties, were in response to similar lease provisions offered by IBM's peripheral equipment rivals.

The Aghion-Bolton model does not apply to *Telex* because the case involves actual competition, rather than potential competition; entry by other peripheral equipment manufacturers had made the market competitive. Thus, IBM could not have induced customers to sign penalty contracts limiting their freedom to deal with lower cost rivals. Customers would gain more by exploiting the open competition between IBM and its rivals. Furthermore, IBM had little if any monopoly return in the peripheral equipment market to use in compensating customers for signing penalty contracts due to the vigorous price competition between IBM and its rivals; in fact, the main issue in the case was predatory pricing. Even if a penalty contract had been a viable business strategy, the IBM lease was short and thus unlikely to impose significant delay cost on new entrants.

2. *Potential entry with rent capture: the* Hazeltine *cases.*

The *Hazeltine* cases,[73] involving patent royalty agreements, illustrate the rent capture effects of a penalty contract. In *Automatic Radio*, the first *Hazeltine* decision, Hazeltine licensed a package of 200 of its radio patents under a 10-year agreement. The license required the licensee-manufacturer to pay both a fixed sum annually and a royalty based on its radio sales, whether or not the patents were used.[74] Thus, a licensee who used the technology of a rival firm still had to pay royalties to Hazeltine; such payments amounted to a penalty for switching to a rival firm if the licensee did not continue to use the Hazeltine patents. Moreover, the penalty was discriminatory: Because the royalty was based on a percentage of end product sales, the more successful the rival patent, the larger the royalty that would have to be paid to Hazeltine. The case involved potential competition because the licensing agreement encompassed future technology. The licensee had made limited use of the Hazeltine patents and would have been unlikely to have signed the agreement if it had known of the existence of a superior technology.

In *Automatic Radio*, the Supreme Court upheld the agreement as a legiti-

73. Zenith Radio Corp. v. Hazeltine Research, Inc., 395 U.S. 100 (1969); Automatic Radio Mfg. Co. v. Hazeltine Research, Inc., 339 U.S. 827 (1950).

74. *Automatic Radio*, 339 U.S. at 829.

mate and convenient exploitation of the patent monopoly.[75] In *Zenith*, the second *Hazeltine* decision, the Supreme Court ruled that the licensing agreement was a patent misuse because the licensee was required to pay royalties on radios not using the patented technology. The Court also found the license illegal because Hazeltine had coerced the licensee to sign the agreement.[76] Presumably, if the agreement had been voluntary, the Court would have upheld the license as a mutually convenient way to arrange royalty payments. Aside from Justice Harlan's brief reference in a separate opinion,[77] the Court failed to consider the royalty provision's effect on innovation rivalry, the relevant issue under the Aghion-Bolton analysis.[78]

The *Hazeltine* license exemplifies a penalty agreement designed to capture an entrant's rent. While the fixed yearly rental would deter entry of a rival technology if the expected return was less than the yearly rental, such an effect was unlikely because the annual lump sum payment was minimal compared with the royalties the licensee had to pay, which amounted to 1 percent of its total radio sales.[79]

The rent capture effect, however, was acute. The licensee paid a royalty of 1 percent of end product sales. The more an entrant's patent might increase the licensee's sales, the higher the licensee's payment to Hazeltine would be in the event of breach.[80] Thus, the patent owner would gain more as a result of the switch than if its monopoly had not been challenged by a superior technology. Consequently, the patent owner could afford to compensate the licensee for agreeing to the take-or-pay provision by reducing the royalty or fixed yearly rental; the reduced royalty received before entry would be outweighed by the higher royalty the owner would receive in the event entry occurred.

Although licensing a package of patents may result in transactional cost savings, it is highly doubtful that such savings could outweigh the social losses sustained from collecting royalty payments when the end products do not use any of the licensed patents. Moreover, where patents can be licensed individually, a royalty based on sales of end products not embodying the patent raises the penalty contract issue without the mitigating convenience of block licensing.

---

75. *Id.* at 833.

76. *Zenith*, 395 U.S. at 135–36, 139–40.

77. *Id.* at 141 (Harlan, J., dissenting).

78. Justice Harlan noted that the royalty provision might reduce the incentive for the licensee to substitute cheaper inputs for the patented product. The licensee's obligation to pay royalties on manufactured products not using the patents reduces a new entrant's gain from improvements in the technology. *Id.* at 145 (Harlan, J., dissenting) (citing William F. Baxter, *Legal Restrictions on Exploitation of the Patent Monopoly: An Economic Analysis*, 76 YALE L.J. 267 (1966)).

79. *See Automatic Radio*, 339 U.S. at 829 (annual lump sum payments of $10,000).

80. A licensee might switch to a rival patent because the patent reduced the licensee's costs or enhanced its product quality. In either case, the licensee's sales would likely increase, either by allowing a higher selling price or by allowing a price reduction and increasing the number of units sold; royalties payable to the patent owner will then exceed those it would have earned if no switch occurred.

3. *Potential entry and buyout of rivals: the* Tetra Pak *case.*

The *Tetra Pak* case,[81] which involved a penalty contract within a pervasive monopolizing system, illustrates the use of penalty contracts to deter entry and to capture rents through the buyout of potential rivals at penalty-depressed prices. Arising under EEC law, *Tetra Pak* also demonstrates how penalty contracts may fortify a monopoly within a legal system where such contracts need not be disguised.

Tetra Pak, the dominant producer of packaging machines and cartons for milk and other liquid food products, used long-term contracts that required customers to pay large penalties for switching to rival packaging machines. As in *United Shoe*, the contracts were part of an elaborate network of restrictive agreements and practices, including tying arrangements, price discrimination, predatory pricing, buyout of rivals, and close surveillance of customers. The penalty contracts were central to this monopolizing plan. Tetra Pak held a monopoly position in the market for aseptic containers, which are filled and sealed under sterilized conditions to enable long shelf life; Tetra Pak's market share ranged from 90 to 95 percent for both packaging machines and containers, and the company faced only one rival.

The penalty contracts involved packaging machines, the vehicle for Tetra Pak's market dominance. In many EEC countries Tetra Pak supplied machines on a lease-only basis, while in others sale was an option. However, both the leasing and sale contracts required customers to pay penalties for switching to a rival firm.

The leasing contracts, with durations of up to nine years, imposed two penalties on customers who sought early termination. First, customers forfeited all prepaid rent. The contracts required customers to pay nearly all rent at the beginning of the lease. Early cancellation, or even entry into a sublease, resulted in forfeiture of all prepaid rent. Second, any breach of the lease made the customer liable for an explicit penalty of up to 10 percent of the total rent, or approximately one year's rental, determined solely at Tetra Pak's discretion.[82]

The sale contracts indirectly limited the customer's ability to switch vendors by severely restricting resale; no resale was permitted without Tetra Pak's authorization, and Tetra Pak could opt to repurchase at a nominal price. Any breach of the sale agreement subjected the customer to a penalty of up to 80 percent of the cost of the machine, in addition to contract damages and interest. This provision clearly was a penalty for switching, because Tetra Pak offered nonswitching customers much higher repurchase prices if they replaced their old machines with new Tetra Pak machines.[83]

Thus, under both leasing and sale contracts, customers who switched to a rival supplier lost substantially all of the residual value of the machine and also paid an explicit penalty. Customers who switched early in the contract

81. Elopak Italia Srl v. Tetra Pak, 4 C.M.L.R. Antitrust Rep. 551 (E.C. Comm'n 1991).
82. *Id.* at 567, 603–05.
83. *Id.* at 579, 599–601, 604–05.

period faced a potential penalty exceeding their lease or purchase price.[84]

Entry barriers were very high in the industry, especially for aseptic packaging machines. Because there was little growth in demand, a rival or new entrant could attract customers only by replacing Tetra Pak's machines.[85] When potential entry occasionally threatened, Tetra Pak bought out the potential entrants or their technology.[86]

The EEC Commission held that Tetra Pak had abused its dominant position in violation of Article 86. The unlawful practices included contractual provisions that had the effect of "unduly binding" customers to Tetra Pak by "artificially eliminating potential competition,"[87] and the buying out of potential rivals in order to exclude competing technologies.[88]

In the *Tetra Pak* case, an undoubted monopolist used penalty contracts to deter new entry and to capture rents by acquiring potential entrants at reduced prices, thereby perpetuating its monopoly. The facts strongly suggest that the penalty exceeded the monopolist's actual losses. First, a customer seeking to terminate the lease or to sell a purchased machine in the early years of the contract faced penalty charges potentially exceeding the total value of the machine. The evidence is most striking in the case of the sale contracts, where the stipulated penalty was expressly additional to contract damages and interest.[89] Second, customers who sought to sell a previously purchased Tetra Pak machine in order to switch to a rival machine paid penalties that were not imposed on customers who replaced an old machine with a new Tetra Pak machine. Because the higher payments did not stem from additional costs incurred by Tetra Pak, the payments exacted from switching customers constituted penalties. Third, the lease term exceeded the product's technological life, indicating an intent to impede entry by raising an entrant's waiting cost in a stagnant market.[90]

Tetra Pak's restrictive agreements resulted in both entry deterrence and rent capture. Because of the stagnant demand and Tetra Pak's extremely high market share, successful entry required sales to Tetra Pak's existing customers. The penalties imposed on switching customers necessarily raised entry costs by requiring entrants to reduce price to gain customers, thereby diminishing the number of entrants. In fact, over a 20-year period, only one actual entrant and three potential entrants appeared in this highly profitable industry.[91]

---

84. *Id.* at 607–08.
85. *Id.* at 561.
86. *Id.* at 616–17.
87. *Id.* at 617.
88. *Id.* at 618.
89. Most of Tetra Pak's profit came from the supply of cartons, which were subject to a tying arrangement requiring Tetra Pak cartons to be used with Tetra Pak machines. But Tetra Pak could have offered its cartons to customers separately from its machines, continuing to profit on the sale of cartons, provided that customers would have bought cartons without the tying contract. If customers would have refused to buy cartons without the tying agreement, this is not a loss that antitrust law should recognize as legitimizing a penalty.
90. *See Tetra Pak*, 4 C.M.L.R. Antitrust Rep. at 604–05.
91. *See id.* at 558–59, 582–83.

Rent capture occurred primarily through the buyout of potential entrants at reduced prices. In each instance of attempted entry, Tetra Pak bought out the rival firm, either by direct acquisition or by purchase of its technology. The value of an entrant and its technology would necessarily have been reduced by the costs the entrant would face in compensating customers for switching penalties. This reduction in the buyout price constituted the rent that Tetra Pak was able to extract from entrants.[92]

Finally, Tetra Pak's agreement also probably caused a reverse free rider effect. Making the reasonable assumption of significant fixed costs and economies of scale, customers who considered switching to an entrant faced the risk that too few other customers would sign up with the entrant to make entry viable. If entry was not viable, the customers who dealt with the entrant faced contract penalties under the Tetra Pak contract, giving Tetra Pak the right to impose a penalty for any breach, however immaterial. Facing such repercussions, customers would refuse to deal with the entrant, thereby denying access to a customer base sufficient to achieve economies of scale.[93]

4. *Potential entry, toehold expansion, and rent capture from consumers and rivals: the* Airline Reservation *cases.*

The *Airline Reservation* cases involved potential entry by firms outside

92. *See id.* at 582–83, 616–18. The conclusion of rent capture assumes that the potential entrants had sufficiently low costs that, in the absence of a penalty contract, they could have earned economic profits. Such an assumption appears reasonable here since it is doubtful that the entrants would have made risky investments in an industry with a dominant and aggressive monopoly firm unless their expected costs were sufficiently low to allow them to earn economic profits. *See* text accompanying notes 21–24 *supra.* Moreover, Tetra Pak would have had little incentive to buy high cost entrants, who could not overcome the penalty and presented no competitive threat. Thus, entrants must have anticipated earning rent, which Tetra Pak sought to capture.

93. A recent Canadian decision also involved the use of penalty contracts to deter entry and capture rent through buyout of rivals. In *In re* Laidlaw Waste Systems Ltd., Canadian Competition Tribunal, Jan. 20, 1992, a monopoly supplier of solid waste collection services entered into long-term penalty contracts which were potentially of unlimited duration since the 3-year contract period was subject to automatic renewal unless the customer served notice of termination within a narrow time period. The contracts contained large stipulated damage provisions and other oppressive conditions, which suggested a penalty since above a minimum scale of operations, most of the costs involved in solid waste collection are avoidable. In addition, the stipulated damages provision applied only to switching customers, not to customers terminating for other reasons. Moreover, the defendant's conduct revealed an intent to impede entry by denying an entrant access to a viable customer base, as evidenced by threatened litigation against switching customers even when defendant had no intention of bringing suit. *Id.* at 28-41, 94-95. Thus, both objective and intent evidence pointed to a penalty.

As a result of these practices the defendant completely monopolized the relevant markets, holding a market share approaching 100% and having a continuing right to increase the agreed contract price through the immediate pass through of all cost increases. *Id.* at 35-36, 96. Thus, the defendant was able to charge a fully monopolistic price at all times. The contracts resulted in both entry deterrence and also rent capture when entrants sold out to defendant at prices necessarily depressed by existence of the penalty contracts. *Id.* at 14-15, 19-21. Customers were induced to sign such contracts by a reverse free rider effect that made it risky for a customer to switch to a new entrant. An entrant could not achieve scale-efficient production unless it converted a sufficient number of defendant's customers to reach a viable operating scale, the demand being inelastic. *Id.* at 97. Laidlaw's aggressive policy of threatening each switching customer with immediate suit necessarily impeded the entrant's effort to recruit a critical mass of customers, *id.* at 105-09, and thus powerfully reinforced customers' tendency to let others take the risk of angering Laidlaw by switching to a rival.

the market, as well as potential expansion by an inside firm with a small toehold position. The cases illustrate that rent capture from final consumers and existing rivals can provide an alternative source of rents to induce customers, here travel agents, to sign penalty contracts.

In several civil actions,[94] as well as in administrative proceedings,[95] private litigants and the government challenged leasing practices for computerized reservation systems ("CRSs"). The CRSs, all of which were owned by airlines, were leased to travel agents for a period of five years.[96] In *Austin Travel* and *Apollo*, the principal cases for discussion here, private litigants challenged the validity of United Air Lines' leases for its Apollo CRS under the antitrust laws. The leases contained a stipulated damage clause that, in the event of early termination, required the travel agent to pay 80 percent of its remaining rental payments and also to compensate United for lost booking charges otherwise earned from other airlines for use of United's CRS service.[97]

United and American Airlines dominated the CRS market with a combined share of 77 percent of national CRS rentals.[98] The two airlines also dominated most regional CRS markets, either individually or jointly, as well as the related local air carrier markets.[99] While United's leases were written for five years, almost none of the leases actually expired at the end of five years, due to United's constant solicitations for early renewal of its leases.[100] Thus, a travel agent wishing to terminate a lease from United usually faced stipulated damages covering several years of lease rentals.

Litigation arose when System One, then owned by Texas Air, induced several travel agents to replace United's Apollo CRS with its own system. System One agreed to defend travel agents in suits brought by United and to

94. *In re* "Apollo" Air Passenger Computer Reservation Sys. (CRS), 720 F. Supp. 1068 (S.D.N.Y. 1989); *In re* Air Passenger Computer Reservations Sys. Antitrust Litig., 694 F. Supp. 1443 (C.D. Cal. 1988), *aff'd sub nom.* Alaska Airlines, Inc. v. United Airlines, Inc., 948 F.2d 536 (9th Cir. 1991), *cert. denied,* 112 S. Ct. 1603 (1992); United Air Lines, Inc. v. Austin Travel Corp., 681 F. Supp. 176 (S.D.N.Y. 1988), *aff'd,* 867 F.2d 737 (2d Cir. 1989).

95. *See* Reply Comments of U.S. Dep't of Justice Before the Dep't of Transp. Regarding Notice of Proposed Rulemaking, Computer Reservation System Regulations (Aug. 8, 1991) [hereinafter DOJ 1991 Reply Comments]; Comments of U.S. Dep't of Justice Before the Dep't of Transp. Regarding Notice of Proposed Rulemaking, Computer Reservation System Regulations (July 9, 1991) [hereinafter DOJ 1991 Comments]; Comments of U.S. Dep't of Justice Before the Dep't of Transp. Regarding Advanced Notice of Proposed Rulemaking, Computer Reservation System Regulations (Nov. 22, 1989) [hereinafter DOJ 1989 Comments]; Comments and Proposed Rules of U.S. Dep't of Justice before the Civil Aeronautics Board regarding Advance Notice of Proposed Rulemaking, Airline Computer Reservation Systems (Nov. 17, 1983) [hereinafter DOJ 1983 Comments].

96. *See Apollo,* 720 F. Supp. at 1077.

97. *Austin Travel,* 867 F.2d at 739. American Airlines used similar leasing arrangements. *See Computer Reservations Sys. Litig.,* 694 F. Supp. at 1452.

98. *Apollo,* 720 F. Supp. at 1073 (noting that in 1986 American and United accounted for 45% and 32% of industry sales, respectively).

99. *See* DOJ 1989 Comments, *supra* note 95, at 24 (citing individual CRS market shares as high as 89.1% for American in El Paso, and 60.4% for United in Honolulu).

100. *Apollo,* 720 F. Supp. at 1077; DOJ 1989 Comments, *supra* note 95, at 27–28. American followed a similar lease renewal policy, and also required CRS customers to sign a new five-year contract whenever installing or replacing any piece of equipment during the term of the contract. *Computer Reservations Sys. Litig.,* 694 F. Supp. at 1452.

indemnify agents for any damages from breaching their leases. System One held only approximately 9 percent of the national CRS market and much smaller shares in many regional markets.[101]

System One and several travel agents challenged the United leases as agreements restraining trade and monopolizing the CRS market in violation of sections 1 and 2 of the Sherman Act.[102] The courts rejected the challenges on several grounds, including lack of monopoly power,[103] absence of a penalty,[104] and a failure to prove either that the substantive contract provisions were anticompetitive or that United acted in bad faith.[105]

Although the courts found that United lacked monopoly power,[106] United nonetheless held dominant market shares of both CRS and air carrier service in several regional markets,[107] setting the basic condition for an Aghion-Bolton effect. Air carrier dominance reinforced CRS dominance because CRS affiliation with a regionally dominant air carrier provided travel agents both real advantages, such as instantaneous confirmation, more accurate booking, ticketing, and better flight information, and perceived advantages, including goodwill of the airline on which most flights would be booked.[108] For these reasons travel agents, who generally dealt with only one CRS supplier, had strong incentives to use the system of the regionally dominant carrier.[109]

Under these conditions the leasing contracts restricted potential competition not only from firms outside the market, but also from existing CRS vendors with small toehold shares. A rival CRS supplier could not become an acceptable substitute for the CRS service of the regionally dominant airline unless it gained a major share of the local air carrier market. Not only was this a highly uncertain prospect for outside firms, such as small airlines without a CRS affiliate, but it was also uncertain for an inside CRS supplier such as System One, which held only a toehold position.[110] Thus, uncertainty regarding entry, a prerequisite for Aghion-Bolton analysis, existed for

101. *See Apollo,* 720 F. Supp. at 1073–74; DOJ 1989 Comments, *supra* note 95, at 24.

102. *See* United Air Lines, Inc. v. Austin Travel Corp., 867 F.2d 737, 739 (2d Cir. 1989); *Apollo,* 720 F. Supp. at 1071.

103. The *Austin Travel* court held that United lacked market power in the national market and in the Long Island regional market that was at issue in the case. *Austin Travel,* 867 F.2d at 742.

104. *Id.* at 740–41. No penalty existed because the stipulated damages were reasonable in light of the anticipated or actual loss caused by the breach. *See id.* at 740.

105. *Apollo,* 720 F. Supp. at 1076.

106. The *Austin Travel* court's finding of absence of market power seems questionable, particularly in the context of regional markets where United, or United in conjunction with American, was dominant. It is outside the scope of this article to resolve the issue of market power. Instead, we assume for the purposes of the following analysis that United had market power in regional CRS markets where it was both the dominant CRS vendor and the dominant air carrier.

107. *See* DOJ 1989 Comments, *supra* note 95, at 24–26.

108. *See id.* at 25; DOJ 1991 Reply Comments, *supra* note 95, at 4–5.

109. *See* DOJ 1991 Reply Comments, *supra* note 95, at 6; *see also* United Air Lines, Inc. v. Civil Aeronautics Bd., 766 F.2d 1107, 1115 (7th Cir. 1985) (noting that United was able to persuade 72% of Denver travel agents to subscribe to its CRS).

110. In Dallas, for example, System One had a market share of 1.4% against American's 87.3%. *See* DOJ 1989 Comments, *supra* note 95, at 24.

both classes of potential entrants.[111]

The key issue in the *Austin Travel* case was therefore whether the stipulated damage provisions constituted penalties, and the court's analysis here was unsatisfactory. Although System One urged that the stipulated damages substantially exceeded United's losses, the court rejected that assertion because United had high sunk costs, which could not be recaptured if the machines were returned.[112] The court upheld stipulated damages requiring switching travel agents to pay 80 percent of future leasing rentals based on a finding that United's avoidable costs, which could be recaptured, did not exceed 20 percent.[113]

However, the court did not indicate any discounting for present value of the future lease rentals. Further, the avoidable cost finding does not justify using an 80 percent cancellation penalty for leases extending beyond five years. Finally, most of the stipulated damages consisted not of unrecoverable sunk costs, but of lost booking charges that United expected to receive from other airlines over the remaining life of the contract. The calculation of these charges assumed that United would have collected fees on 50 percent of the average bookings through the CRS. However, because booking fees are only paid for flights on other airlines, United was unlikely to lose such a large percentage of bookings where it was the dominant carrier.[114]

Assuming that the cancellation charges were penalties, the leases would have acted as a strong deterrent to new entry and toehold expansion. Furthermore, System One's inducement of travel agents by indemnification against liquidated damages claims suggests that rent capture was probably also present.[115] Although it is uncertain whether System One ever paid such indemnifications, after the leases were upheld System One ceased indemnifying agents. Conversions of customers both to System One and other smaller systems "virtually ground to a halt," and entry of new CRS systems appeared highly improbable.[116] Thus, any rent capture from potential entrants previously occurring disappeared. Accordingly, there must have been another reason why travel agents were induced to sign penalty contracts.

Two mechanisms would help persuade travel agents to sign the penalty contracts with a regionally dominant airline. First, the large sunk costs and economies of scale required for CRS systems triggered a reverse free rider

111. An objection might be made that because a firm with a toehold position had actually entered the market, there was actual competition, rather than simply potential competition. However, there is a significant difference between placing a CRS service with a few agents in a dominated market and challenging the dominant CRS supplier across the market. A small CRS service such as System One might have value for agents servicing customers flying predominantly on System One routes. For most airline agents, however, the small CRS service is nothing more than a potentially viable alternative to a dominant CRS.

112. United Air Lines, Inc. v. Austin Travel Corp., 867 F.2d 737, 740 (2d Cir. 1989).

113. *Id.* System One argued that United's avoidable costs approximated 40 to 50%, but the testimony was dismissed as that of a rival. *Id.*

114. For example, if 80% of the bookings were on United, United's CRS would collect booking charges on only 20% of the flights and a 50% charge would be clearly excessive.

115. *See* DOJ 1989 Comments, *supra* note 95, at 29.

116. *Id.* at 29, 32.

effect.[117] Second, dominance of both the regional air carrier market and the local CRS market enabled United to earn additional rents from passengers and other airlines, thereby creating a surplus to compensate travel agents for signing penalty contracts. A regionally dominant airline can capture sales from rivals by biasing the flight information display[118] or charging high access fees for use of its CRS facility.[119] Furthermore, the dominant air carrier can also charge passengers monopoly prices for tickets. Thus, the dominant airline can earn rents to share with agents who sign the contracts that make such exclusion possible.[120]

5. *Potential entry without penalty*: Northeastern Telephone.

The *Northeastern Telephone* case[121] is the only recent decision approximating careful examination of penalty determination. Northeastern Telephone attempted to sell telephone equipment to customers of AT&T after the market was opened to non-AT&T suppliers by FCC intervention. Northeastern, a small entrant seeking a toehold in the AT&T-dominated market, claimed that AT&T's equipment payment plan was exclusionary and unlawful. Under the two-tier plan, a customer was required to pay both a monthly service charge and charge for capital equipment, payable over a fixed period ranging from one to ten years. After the capital charges were paid, the customer continued to pay the monthly operating charge. If a customer sought termination before the capital charges were completely paid, the customer was required to pay a termination charge equal to the discounted present value of the remaining capital payments. This amount was offset by a credit if AT&T was able to use the equipment for other customers.

Northeastern argued that the plan locked in customers, because a switching customer had to pay not only the entrant's capital charges, but also a termination charge to AT&T. The court rejected Northeastern's argument

117. *See id.*

118. *See* United Air Lines, Inc. v. Civil Aeronautics Bd., 766 F.2d 1107, 1113 (7th Cir. 1985); DOJ 1989 Comments, *supra* note 95, at 17–18.

119. *See Civil Aeronautics Bd.*, 766 F.2d at 1115; *In Re* Air Passenger Computer Reservations Sys. Antitrust Litig., 694 F. Supp. 1443, 1461 (C.D. Cal. 1989), *aff'd sub nom.* Alaska Airlines, Inc. v. United Air Lines, Inc., 948 F.2d 536 (9th Cir. 1991), *cert. denied,* 112 S. Ct. 1603 (1992); DOJ 1991 Comments, *supra* note 95, at 5.

120. We are indebted to Patrick Bolton for this insight. *Cf.* Oliver Hart & Jean Tirole, *Vertical Integration and Market Foreclosure, in* BROOKINGS PAPERS ON ECONOMIC ACTIVITY: MICROECONOMICS 1990, at 205, 252–55 (Martin Neil Baily & Clifford Winston eds.) (arguing that merger between downstream airline and upstream CRS supplier can create ex post monopolization of the downstream airline, allowing price increases detrimental to customers and rival airlines); Janusz A. Ordover, Garth Saloner & Steven C. Salop, *Equilibrium Vertical Foreclosure,* 80 AM. ECON. REV. 127, 127–28 (1990) (demonstrating that a downstream firm can drive up the costs of its rivals by acquiring an upstream firm). *But see* Edward A. Snyder & Thomas E. Kauper, *Misuse of the Antitrust Laws: The Competitor Plaintiff,* 90 MICH. L. REV. 551, 593–95 (1991) (questioning the policy implications of these papers because of the simplicity of the assumptions underlying their economic models).

121. Northeastern Tel. Co. v. AT&T, 651 F.2d 76 (2d Cir. 1981), *cert. denied,* 455 U.S. 943 (1982).

because the residual value credit available under the plan put a switching customer in the same position as a customer who purchased the equipment outright and sought to resell the old equipment. In either case the customer would pay the capital charges of disposing of the old equipment before turning to a new supplier. The court also noted that AT&T's competitors offered almost identical capital equipment payment plans.[122]

Northeastern Telephone demonstrates that courts are capable of scrutinizing penalties. Assuming that expected revenues accurately measured AT&T's losses from early termination and that AT&T fairly administered the credit provision to give customers allowance for all avoidable costs, the termination charge, which the court correctly discounted to present value, would not have exceeded AT&T's actual losses. The court's decision thus appears consistent with the Aghion-Bolton analysis. Furthermore, the case demonstrates how a supplier can formulate a liquidated damage provision that is not a penalty and how courts might effectively review such provisions.

6. *Actual and potential entry*: Barry Wright v. ITT Grinnell.

Barry Wright[123] represents a mixed case, where elements of both actual and potential competition were present, and illustrates the critical need to distinguish between the two effects in an Aghion-Bolton analysis. In *Barry Wright*, Pacific, the dominant seller of mechanical snubbers for nuclear power plants, agreed to supply Grinnell with most of its snubber requirements under a series of year-long supply contracts.[124]

A high cancellation charge, not found in previous contracts between Pacific and Grinnell, was added after Grinnell attempted to create an alternative source of mechanical snubbers through a joint venture with Barry Wright. When Barry Wright encountered delays in performance, Pacific and Grinnell entered into negotiations because Grinnell needed to make deliveries and Pacific had large unused capacity.[125] The resulting contracts included additional 5 to 10 percent price discounts, in excess of Pacific's standard discounts, and a 100 percent cancellation charge if Grinnell failed to take the agreed purchase amounts. Subsequently, Grinnell informed Barry Wright that its performance delays constituted breach of their development agreement. Barry Wright sued Pacific for antitrust damages, arguing that the cancellation charge was exclusionary.[126]

The court held that the supply contracts did not violate the antitrust laws, reasoning that even if the 100 percent cancellation charge was a penalty, contract law would prevent enforcement.[127] The court dismissed the potentially exclusionary effect of the litigation required for Grinnell to prove

122. *Id.* at 79–81, 91–93.
123. Barry Wright Corp. v. ITT Grinnell Corp., 724 F.2d 227 (1st Cir. 1983).
124. *Id.* at 230.
125. *See* Barry Wright Corp. v. Pacific Scientific Corp., 555 F. Supp. 1264, 1267 (D. Mass.), *aff'd sub nom.* Barry Wright Corp. v. ITT Grinnell Corp., 724 F.2d 227 (1st Cir. 1983).
126. *See Barry Wright*, 724 F.2d at 229–30.
127. *Id.* at 238.

the clause unenforceable, emphasizing the size and competence of Grinnell's legal staff.[128] Responding to the allegation that the contracts were exclusionary because nearly all of Grinnell's requirements were covered by the agreements, the court noted that Grinnell's interests were best served by encouraging, rather than foreclosing, competition in the industry. Consequently, Grinnell would not have agreed to a provision hindering competition, given its previous attempts to open up the market.[129] The court concluded that the cancellation charge had no coercive effect on the buyer and the potential anticompetitive effects were speculative.[130]

The conditions for application of the Aghion-Bolton model are largely absent because entry was not uncertain and the contracts were of short duration. By the time Pacific had signed the penalty contracts with Grinnell, Barry Wright had proved unable to perform within the time limits specified, when time was of the essence. Thus, the penalty could not have had a deterrent or rent capture effect on Barry Wright.[131]

The cancellation penalty could, however, have had a deterrent effect on other possible entrants. A 100 percent cancellation charge would surely prevent entry in most circumstances because an entrant would not be likely to have sufficiently low costs to overcome such a large penalty.[132] However, the limited duration of the contracts makes an anticompetitive effect unlikely. The maximum restrictive reach of the contracts, two and a half years,[133] does not seem appreciably longer than the required product lead time. Thus, the penalty contracts would not have bound Grinnell for any appreciable period, thereby allowing access to a lower cost entrant soon after completion of its plant. The main significance of *Barry Wright* is the court's erroneous assertion that the contract law standard meets antitrust needs.

III. PROPOSED LEGAL CRITERIA

The new economic insights on contract penalties are relevant to antitrust policy only if workable legal criteria can identify anticompetitive risks, a challenging task given that penalties are often difficult to discover. Thus, we must search for legal criteria to address: (1) market power and foreclosure; (2) the definition of a contract penalty; (3) the evidence needed to identify a penalty; and (4) evaluation of an efficiencies defense.

---

128. *Id.* at 238–39.

129. *Id.* at 237–38.

130. *Id.* at 239.

131. The case essentially involved bilateral bargaining between a powerful buyer and a monopolistic seller. The buyer financed a new entrant into the seller's market to improve its negotiating position. Although the entrant was not successful, the mere threat of entry forced the seller to reduce its price, thereby achieving the buyer's goal.

132. The large size of the penalty would also preclude a rent capture effect because such an effect presumes entry.

133. Although there were actually three consecutive 1-year contracts, when the last two were finalized, the contract period had two and one half years remaining. *Barry Wright,* 724 F.2d at 229, 237.

A. *Market Power and Foreclosure*

Penalty contracts raise anticompetitive risks when a seller with market power uses them on a significant scale to foreclose a future entrant's independent access to the customer base necessary to compete effectively. As discussed previously, the penalty contract operates as a strategic mechanism either to exclude the entrant from the market or to force the entrant to sell at a low price and earn little, if any, economic profit. Nevertheless, this penalty strategy will succeed only if the seller has market power and independent access to customers is foreclosed.

1. *Market power.*

Generally, the market power that raises anticompetitive risks will be single firm monopoly. However, anticompetitive concerns may also arise in oligopolistic markets, especially duopolies or near duopolies, where two firms have a dominant market share. When both duopolists use similar penalty contracts and have reached a stable accommodation enabling both to earn supracompetitive profits, they may effectively act as an Aghion-Bolton monopolist, resulting in entry deterrence and rent extraction from new entrants. Although neither duopolist acting alone could induce customers to sign penalty contracts, if both duopolists offer only such contracts, customers might reasonably believe that the two firms have effectively monopolized the market. Customers would then sign penalty contracts to share in the capture of rents from future entrants.[134] Once a penalty regime is constructed, neither duopolist would have an incentive to curtail use of penalty contracts because to do so would simply allow the other duopolist to steal its customers. Mutual adherence to the penalty system would exploit potential entrants to the advantage of both duopolists.[135]

We do not assert that use of penalty contracts by duopoly firms inevitably injures competition, but simply suggest that the issue merits inquiry where duopolists or near duopolists are using similar penalty contracts in consistently profitable markets.[136] The issue may be particularly suitable for

134. Additional rent could also be captured from other market participants as in the *Airline Reservation* cases. *See* text accompanying notes 117–120 *supra*.

135. Application of the Aghion-Bolton model to duopoly may be criticized as not following its strict requirements. The model assumes, under Bertrand competition, that price will fall to the marginal cost of the higher cost producer. In this scenario, at least one duopolist would have no rent to share with customers, and the Aghion-Bolton model would fail. However, if the duopolists have reached a stable accommodation or collusive equilibrium such that each consistently earns economic rent, the Aghion-Bolton analysis applies. Equilibrium might result from tacit collusion, product differentiation, or oligopolistic competition under less competitive pricing, such as Cournot competition. The Aghion-Bolton model applies, provided there is some rent to share with customers. While full treatment of the duopoly problem would require a formal model, antitrust law should not ignore use of similar penalty contracts in markets where duopolists are consistently able to earn economic rents.

136. The courts in the *Airline Reservation* cases rejected charges that the penalty contracts represented monopolizing conduct because United's 31% share of the national market did not constitute monopoly power. *E.g.*, United Air Lines, Inc. v. Austin Travel Corp., 867 F.2d 737, 742 (2d Cir. 1989). The courts failed to consider, however, that United and American together held 77% of the national market and jointly dominated several regional markets. *See* note 98 *supra* and accom-

FTC inquiry under section 5 of the FTC Act.[137]

2. *Foreclosure.*

Even if the seller has market power, penalty contracts will not raise anticompetitive concerns unless a sufficient share of the seller's customers sign such agreements. If enough customers to support at least one entrant of viable scale are not bound by penalty contracts, the entrant can sell to these customers and bypass the penalty.

When large fixed costs or scale economies are involved, customers who sign penalty contracts drive up the cost of production for all customers. Customers who decline to sign a penalty contract are at a greater risk of paying a higher price, either to a high cost supplier or to an unsympathetic monopolist, who denies the price concession given to loyal buyers.[138] Thus, contract penalties raise antitrust issues when the available customer base is reduced to preclude a new entrant from operating at efficient scale.

3. *Monopsony buyer.*

Initially, it appears that a monopsony buyer would be unwilling to sign a penalty contract to avoid sharing the entrant's economic surplus with the seller, preferring to exercise its buying power to negotiate directly with the entrant seller. However, the strategic advantage of the penalty agreement is also present in the monopsony buyer case. As in the multiple buyer scenario, the penalty contract creates a binding commitment by the buyer to pay the penalty upon switching to the entrant. If the entrant initiates direct negotiations, the buyer can credibly claim that it is bound by the penalty agreement. Only the incumbent seller, who has no incentive to improve the entrant's welfare, can release the buyer from the penalty agreement.

The monopsony buyer case has the same strategic properties as the multiple buyer case. The primary difference is that an incumbent seller is likely to agree to give the monopsony buyer a larger share of the entrant's surplus than in the multiple buyer case.[139] Thus, contract penalties still raise an-

---

panying text. Moreover, the courts ignored the fact that each airline entered into similar 5-year leases with stipulated damage provisions and consistently renewed such leases before expiration. Likewise, each airline faced new entry and toehold expansion from smaller reservation systems, such as System One. Under these circumstances, American and United's mutual interests may have dictated accommodating policies imposing penalties on switching customers, thereby restraining potential competition.

A similar issue of parallel use of penalty contracts arose in *In re* Laidlaw Waste Systems Ltd., Canadian Competition Tribunal, Jan. 20, 1992, at 37, 99-100, where the defendant sought to defend its conduct on the ground that its two leading rivals also used such contracts. The court rejected the argument as irrelevant, stating that two wrongs did not make a right. The court failed to recognize that parallel use of similar penalty contracts in highly profitable markets may have been tacitly collusive, enabling the firms to act jointly as an Aghion-Bolton monopolist.

137. 15 U.S.C. § 45 (1988).

138. The reverse free rider scenario operated in the *Airline Reservation* cases. *See* text accompanying note 117 *supra.*

139. In the absence of a penalty contract, the buyer would be able to exercise bargaining power against the entrant. However, a lower cost entrant would also have bargaining power. Thus, the monopsony buyer has more bargaining power than multiple buyers regardless of contract penalties,

ticompetitive concerns even when an incumbent with market power is selling to a single buyer under a penalty contract.[140]

B. *Definition of Contract Penalty*

A contract penalty raises anticompetitive risks when a monopolist imposes costs on switching customers that exceed the monopolist's actual losses. The penalty is the excess amount above actual losses that the switching buyer must pay, determined at the time of breach. Our definition, designed for use in antitrust cases, differs from the contract law definition by considering only ex post evidence and avoiding the vagaries and imprecision of the common law test. Moreover, actual losses include the costs actually sustained by the incumbent, even those not ordinarily recoverable in contract damages, such as the interest or opportunity costs of deferred income. Thus, the antitrust definition of a penalty more realistically measures the incumbent's losses.

Contract provisions imposing switching costs exist in many forms, including liquidated damages, take-or-pay provisions, forfeiture of collateral, or loss of future discounts and other benefits. In each case, the buyer who switches to a rival supplier has agreed to surrender something of value above the incumbent's actual losses. The relevant antitrust issue is whether the provision may be a strategic vehicle for monopolizing conduct.

The contract law definition of penalties is inadequate for antitrust analysis because of its permissiveness in specifying liquidated damages and allowing disguised penalties. Our criticisms of contract law do not deny that in competitive markets and even in markets of imperfect information with strategic bargaining, the parties can write efficient contracts more successfully than can courts.[141] However, when firms with market power enter into strategic coalitions with buyers to exploit potential entrants, the market is no longer effectively competitive; courts must then closely examine contract terms to ensure they do not become vehicles for restraining trade. Effective analysis requires methods to identify contract penalties, relying on both objective evidence and, where such evidence is not decisive, on carefully limited evidence of intent.

C. *Objective Evidence of a Penalty*

Determining whether liquidated damages or a similar provision is a penalty under the antitrust definition poses practical difficulties, which may ex-

---

but the monopsonist's bargaining power increases even more upon forming a coalition with the incumbent seller.

140. When market power exists solely at the buyer level and the supply market is competitive, the analysis mirrors that of the monopoly seller case. The monopsony buyer writes contracts with its suppliers that require suppliers to pay a penalty for shifting sales from the incumbent buyer to a rival firm. As in the monopoly supplier case, the monopsonistic buyer, acting jointly with its suppliers, will create a strategic coalition to extract surplus from the monopolist's potential rivals. *See* Rasmusen, *supra* note 17, at 383–85.

141. *See* Schwartz, *supra* note 47, at 383–87; text accompanying notes 35–59 *supra.*

plain the reluctance of courts to undertake careful examination of such provisions. However, the analysis can be simplified by focusing on three objective factors: (1) differential treatment of customers who switch to a rival firm, or between customers in monopoly and competitive markets; (2) refusal of a sale or short-term lease option; and (3) the level of stipulated damages in relation to actual damages.

1. *Adverse treatment of switching customers.*

A stipulated damages provision is likely to be a penalty if it (1) imposes a larger payment on customers who switch to a rival supplier than on customers who terminate their relationship with a seller for other reasons; or (2) requires switching customers in monopoly markets to make discriminatorily larger payments than switching customers in competitive markets. Differential treatment of customers is a powerful indicator of a penalty because, in a standard competitive market, when a lease ends, a switching customer is treated the same as a nonswitching customer. Because the leased machine is returned and the customer thereafter cannot affect its value, the supplier has no reason to treat the switching customer differently than the nonswitching customer. Thus, discriminatory treatment of switching customers provides strong evidence of a penalty intended to preclude entry.

United Shoe and *Tetra Pak* provide vivid examples of adverse treatment of customers who switched to a rival firm. In *United Shoe*, customers who returned leased United Shoe machines to substitute rival machines were subject to higher termination charges and more stringent enforcement of the leasing contract than customers who replaced one United Shoe machine with another.[142] Similarly, customers who reduced usage of United Shoe machines in favor of rival machines were treated more harshly than customers who reduced usage for other reasons.[143] In *Tetra Pak*, customers who tried to dispose of a purchased machine to switch to a rival machine received lower repurchase prices than customers who replaced old machines with new Tetra Pak machines.[144] While such differential treatment might be justified on efficiency grounds or rebutted by other factors, these practices clearly merit close scrutiny for potential antitrust violations.

Selective application or enforcement of stipulated damage provisions provides another means of identifying the possible presence of a penalty contract. As previously noted, contract penalties are unlikely in competitive markets because competition itself will eliminate such costly practices. If an incumbent supplies the same or a similar product in both competitive and monopoly markets, the competitive market provides a nonpenalty bench-

142. United States v. United Shoe Mach. Corp., 110 F. Supp. 295, 319–20 (D. Mass. 1953), *aff'd per curiam*, 347 U.S. 521 (1954).

143. *Id.* at 323.

144. Elopak Italia Srl v. Tetra Pak, 4 C.M.L.R. Antitrust Rep. 551, 600 (E.C. Comm'n 1991); *see also* Laidlaw Waste Sys. Ltd., Canadian Competition Tribunal, Jan. 20, 1992 (holding that differential treatment of switching customers as compared with customers terminating contracts for other reasons was probative of abuse of dominant position).

mark against which to assess the incumbent's behavior in the monopoly market. Behavior by the incumbent suggesting the possibility of a penalty includes: (1) entering into stipulated damage contracts only in the monopoly market; (2) while including stipulated damages in all contracts, enforcing these provisions against switching customers only in the monopoly market or discriminating, only in monopoly markets, against buyers who refuse to agree to stipulated damages; or (3) setting stipulated damages in the monopoly market at significantly higher levels than in the competitive market, such that the differential substantially exceeds any lost profit variation between the two markets. However, these indicators are not conclusive, because the different economic characteristics of the monopoly market itself could result in distinctive treatment.

2. *Refusal of sale or short-term lease option.*

A penalty may also be present when the seller refuses to offer its product at any price that would make purchase or a short-term lease a viable option. In the absence of a penalty, a long-term, lease-only policy would ordinarily not make economic sense.

If there were no penalty, a lower cost entrant could offer the leasing customers a price low enough to enable it to cancel the long-term lease, pay the original supplier its lost profits, and still gain. Under these conditions when switching suppliers is relatively easy, the incumbent has no reason to refuse a sale or short-term lease option, which some customers may prefer to a long-term lease; because these customers will be willing to pay extra for the added options, the incumbent seller sacrifices revenues by refusing to offer alternatives.

Refusal of a sale or short-term lease option may be justified when long-term leasing involves peculiar economies. Furthermore, an incumbent monopolist may justify refusal as necessary to exploit an otherwise lawful monopoly of a durable good. However, this is a less plausible explanation. While a durable good monopolist might insist on leasing, there is no convincing reason to refuse a *short-term* leasing alternative.[145] Unless there is an efficiencies explanation, or much less likely a persuasive durable goods

145. The producer of a durable good might insist on leasing for at least three reasons. First, leasing may allow highly effective price discrimination, as in *United Shoe,* where rentals were based on machine output measured by meters installed on the machine. *See United Shoe,* 110 F. Supp. at 316–17. Second, the Coase Conjecture suggests that a durable good monopolist might use short-term leasing to induce buyers with immediate needs to pay the monopoly price without fear that later buyers would pay less. *See* John Shepard Wiley, Jr., Eric Rasmusen & J. Mark Ramseyer, *The Leasing Monopolist,* 37 UCLA L. REV. 693, 719–24 (1990). Third, a durable good monopolist might use leases to eliminate the possibility of a second-hand market, which might drive down the monopoly price of the original equipment market. *See generally* Michael Waldman, Eliminating the Market for Secondhand Goods: An Alternative Explanation for Leasing (May 1992) (unpublished manuscript, on file with the *Stanford Law Review*).

Although these reasons explain why a durable good monopolist might refuse a sale option, none sufficiently explains why the monopolist would reject a *short-term* lease. Indeed, the Coase Conjecture implies that a short-term lease would be much more effective than a long-term lease in exploiting the monopoly.

explanation, the incumbent supplier has no apparent justification for refusing a sale or short-term lease option other than to impede entry.

However, an incumbent monopolist intent on forming a coalition with customers to impede potential entry has strong motivation to refuse customers a sale or short-term lease option. As discussed earlier, the critical factor in implementing the coalition is the penalty imposed on buyers who attempt to withdraw from the coalition and buy directly from the entrant. A sale or short-term lease option undermines the coalition by allowing the entrant to bypass the penalty and sell directly to the customers. Although the entrant must compensate sale or short-term leasing customers for any loss from disposal of their old machines, such compensation includes no penalty.

Therefore, an incumbent monopolist's refusal to offer a sale or short-term lease option in a long-term supply contract suggests the presence of a penalty. Moreover, the option must represent a viable economic alternative, rather than being simply a formality. In the *Airline Reservation* cases, United Airlines apparently offered customers no alternative to its 5-year leases, which continually were renewed before expiration. While sale of the CRS to travel agents would not have been a realistic alternative, the court should have examined the viability of a shorter term lease. In the *Telex* case, by contrast, IBM gave customers both sale and short-term leasing options. Such alternatives, if economically viable, provide strong indication of the absence of a penalty contract.

3. *Stipulated damages exceeding actual loss.*

The final, and often critical, objective test for identifying contract penalties is to examine the actual economic impact of the penalty provision at the time a rival firm attempts to enter the market. The key issue is whether stipulated damages or alternative performance provisions exceed the actual losses the incumbent would sustain from loss of a customer, as determined ex post. If the stipulated damages are excessive, the court should closely scrutinize the contract penalty for antitrust violations.

Although courts review ex post consequences in contract penalty cases, the norm is to determine the parties' ex ante reasonable expectations.[146] This ex ante, or reasonable expectation, standard provides an ineffective restraint on stipulated damages as it encompasses inherently subjective assessments of anticipated future losses.

Moreover, the divergence between a supplier's anticipated and actual losses is likely to increase in monopoly cases because the monopolist may overestimate the price reduction resulting from entry. The stipulated damages may well presume that price will fall to competitive levels if a new firm enters the market. However, the price will frequently not decrease substantially, because new firms will enter only if price wars are not probable and profits are available to justify risky investment. Consequently, the stipulated

146. *See* 3 FARNSWORTH, *supra* note 35, § 12.9.

damages will exceed the monopolist's actual losses, and the resulting penalty is likely to have anticompetitive effects.

Using an ex post standard to judge the reasonableness of stipulated damages provides an objective measure of a penalty. A court can simply examine the market after entry to determine whether the stipulated damages imposed on switching customers exceed the monopolist's actual loss.[147]

An objective test based on actual losses introduces a complex issue into the legal analysis, but the complexities appear manageable. The monopolist need not abandon liquidated damages provisions to avoid antitrust scrutiny; it must only use facially reasonable liquidated damages in its contracts. When enforcing such provisions against a customer who switches to a rival, the monopolist should simply limit its demand to the actual losses it sustained. For instance, if the contract assesses liquidated damages based on loss of the full monopoly profit, but in fact losses consist of only a portion of that profit, the monopolist can avoid liability for imposing a penalty by limiting its demand to the smaller amount.[148]

D. *Intent Evidence*

Evidence of intent becomes relevant when objective evidence is not decisive. Although intent evidence is generally inferior to objective evidence because competitive and anticompetitive motivations are often indistinguishable, this difficulty is less acute for entry-impeding penalty contracts. Furthermore, in recent decisions, the Supreme Court has recognized the relevance of intent evidence under both sections 1 and 2 of the Sherman Act.[149]

Intent evidence is more reliable regarding penalty contracts because the concerns of a supplier and its customers in writing an ordinary liquidated damages clause differ sharply from their concerns in creating an entry-impeding penalty contract. In the ordinary liquidated damages provision, the parties anticipate *their own* future losses, while in an entry-impeding penalty contract, the parties consider the entry effect on *other firms*.

Intent to create an anticompetitive penalty contract is shown by evidence

147. The *Northeastern Telephone* court followed such an approach in upholding the legality under the antitrust laws of a termination charge for leased telephone equipment. The customer had to pay only the discounted present value of future lease payments, offset by the alternative revenues AT&T would earn from early return of the leased equipment. Northeastern Tel. Co. v. AT&T, 651 F.2d 76, 93 (2d Cir. 1981), *cert. denied*, 455 U.S. 942 (1982). In the *Airline Reservation* cases, however, the courts examined both anticipated and actual losses and did not clearly determine whether stipulated damages exceeded actual losses.

148. It may appear curious that this analysis permits a higher stipulated damage amount when price falls to the competitive level than when price remains at the monopoly level. Antitrust law contains no prohibition against monopoly pricing; the law prohibits monopolizing conduct. Thus, a stipulated damage provision that awards a monopolist an amount not exceeding the recovery normally obtained for breach of contract involves no anticompetitive use of monopoly power.

149. *See* Broadcast Music, Inc. v. Columbia Broadcasting Sys., Inc., 441 U.S. 1, 19–21 (1979) (using evidence of purpose of licensing scheme to find no violation of § 1); Aspen Skiing Co. v. Aspen Highlands Skiing Corp., 472 U.S. 585, 602 (1985) (using evidence of intent to harm competitor to find violation of § 2).

that one or more contracting parties have set stipulated damages either to directly impede new entry or indirectly increase the entrant's waiting costs. Unless parties intend to exclude entrants using a penalty, they have no reason to focus on the entry-related effects of stipulated damages. Therefore, courts may reasonably presume that if a monopoly supplier writes a stipulated damage clause with express concern for these factors, the monopolist demonstrates an anticompetitive intent.

Evidence that parties formulated stipulated damages to raise an entrant's waiting costs also shows anticompetitive intent. Aghion and Bolton define waiting costs as the forcing mechanism deterring new entrants or extracting a fee for entry. These waiting costs include the capital and other costs that an entrant incurs after its plant is functional and before an adequate customer base is available free of penalty.[150] In the absence of waiting costs, the entrant and its customers could avoid the penalty by simply waiting until existing contracts have expired. As a result, penalties in short-term contracts will have little deterrent effect on entry. But as the term of the contract lengthens, waiting becomes a less attractive option and the entrant must choose either to reduce price by the amount of the penalty or to bear substantial waiting costs.

An entrant's waiting costs thus depend on the length of the contract. The court in *United Shoe* apparently recognized this relationship in granting relief that included a shortening of the 10-year contract term to five years, thereby creating an *ad hoc* standard for the permissible length of a penalty contract.[151] However, the correlation between contract length and an entrant's waiting costs depends entirely on the particular cost levels and production functions of the entrant. Thus, contract length is not decisive in assessing the competitive effects of a penalty contract. However, the *method* by which an incumbent monopolist determines contract length may reveal an intent to increase an entrant's waiting costs, thereby impeding entry.

Evidence showing that a monopolist supplier has intentionally manipulated contract length to impede entry or increase the entrant's waiting costs includes the following: (1) lengthening the contract to influence the likelihood or rate of new entry; (2) determining contract length using an assessment of the customer base required by an entrant to achieve an efficient output level; and (3) selecting a contract term longer than the product lifecycle, or estimated future product lifecycle, in industries of rapid technological change.

Lengthening the contract to influence new entry. Extending the contract solely or primarily in response to probable entry suggests an intent to narrow

150. The penalty also imposes waiting costs on *customers* by extending the time before they can switch to a more efficient supplier. The customer's waiting cost is the likely profit it would have earned by switching to the more efficient supplier without delay. As previously discussed, however, the customer may benefit from rent sharing with the incumbent. The incumbent's overriding anticompetitive intent is to increase the *entrant's* waiting cost to deter or tax new entry.

151. *See* Telex Corp. v. IBM, 510 F.2d 894, 920 (10th Cir.), *cert. dismissed,* 423 U.S. 802 (1975).

the market available to the entrant free of penalty. This factor alone will rarely be sufficient to show an entry-impeding purpose because many other factors influence contract length. However, any extension of the contract in response to probable entry adds force to other evidence bearing on intent.

For example, in *Tetra Pak*, use of a 9-year lease in the Italian market, where Tetra Pak was zealously combating new entrants, strongly suggested an intent to impede entry, because leases in several other countries were as short as three years.[152] Similarly, in the *Airline Reservation* cases, the court should have examined why United Air Lines consistently sought early renewal of its 5-year leasing contracts, especially because the contracts had been limited to 5-year terms only after governmental pressure.[153]

Determining contract length based on the entrant's required customer base for scale efficient output. These calculations are highly revealing of an intent to impede market entry by raising waiting costs. Suppose an entrant requires a 25 percent market share to operate at an efficient output level and the incumbent monopolist substantially increases its contract period from five to ten years. Assuming the contracts are evenly distributed in time, the entrant must now wait two and one half years before the minimum 25 percent of the market will be available without penalty.[154] Evidence that the incumbent deliberately set the contract length to delay an entrant's access to the minimum required market share for efficient operations strongly demonstrates intent to impede entry.

Selecting a contract term longer than the product lifecycle or estimated future lifecycle. Unless otherwise explained by efficiency considerations, a contract term exceeding the product lifecycle—the length of time a customer uses a machine before replacement—shows an intent to increase waiting costs. When the contract length is limited to a single machine generation, the penalty may reimburse informational or other advantages transferred to the customer with the machine. However, when the contract extends over a longer period, such machine-specific economies no longer apply. Moreover, uncertainty costs increase because the contracts must encompass undeveloped products, further inhibiting efficiency gain.

In the *Airline Reservation* cases, United's solicitation of early renewal of the leases increased the effect of the contract term far beyond five years. Assuming five years as the average product lifecycle, this practice indicates an intent to raise a future entrant's waiting costs. On the other hand, in the *Barry Wright* case, the length of the contract probably did not exceed the

152. *See* Elopak Italia Srl v. Tetra Pak, 4 C.M.L.R. Antitrust Rep. 551, 579–82, 604 (E.C. Comm'n 1991).

153. *See* DOJ 1989 Comments, *supra* note 95, at 56.

154. Even if the entrant had access to 25% of the market, that would not guarantee an efficient output level. Customer inertia and the incumbent's reputational advantages virtually assure that not all accessible customers will immediately switch to the entrant. Thus, the entrant will ordinarily require access to a greater percentage of the customer base than the minimum needed to operate efficiently if all of the customers switched.

start-up time for a new entrant, which tends to negate any evidence of an anticompetitive intent.

Similar considerations apply when the contract length exceeds estimated future product lifecycles. This factor is particularly important in industries with rapid technological innovation, where short product life greatly increases the cost of delay. The EEC Commission recognized this effect in the *Tetra Pak* case, holding that a minimum lease term exceeding the technological, but not the physical life of the product would violate Article 86, and noting that even a 3-year lease could be excessive for industries involving rapid technological development.[155] Thus, without a persuasive efficiencies defense, a contract extending beyond a single product generation demonstrates an intent to raise an entrant's waiting cost.

In summary, where objective evidence is inconclusive, stipulated damage contracts should be suspect when intended to impede entry. Evidence of anticompetitive intent can be shown in the following ways:

(1) Stipulated damages are fixed with a purpose to deter new entry or increase the existing profits of incumbent firms through payments by customers, subsidized by price concessions from new entrants;

(2) the contract term is lengthened solely or primarily to deter new entry;

(3) contract length is set to delay a potential entrant's access to the necessary market share to achieve an efficient level of output; or

(4) the contract extends beyond the product lifecycle, or anticipated future product lifecycle, without a clear showing of efficiency gains.

E. *Efficiencies Defense*

If the above criteria are satisfied and if, as a result, competition is substantially injured, penalty contracts should be subject to antitrust challenge. A firm could rebut the presumption of an antitrust violation by showing that the contract terms are necessary to achieve substantial efficiencies and that no less restrictive alternative is reasonably available.

Allowing an efficiencies defense to restrictive penalty contracts is legally required and economically justified. Although potentially unlawful under sections 1 or 2 of the Sherman Act, or section 5 of the FTC Act, penalty contracts would not be illegal per se. Under modern antitrust law, the resulting Rule of Reason analysis includes an efficiencies defense.[156] Examining efficiency implications is also justifiable on economic grounds. Because rigorous economic analysis of penalty contracts is a recent development, possible exceptions to the general conclusions of the economic model may not be fully understood. Moreover, courts and enforcement agencies have

155. *See Tetra Pak*, 4 C.M.L.R. Antitrust Rep. at 604–05.

156. *See* National Collegiate Athletic Ass'n v. Board of Regents of the Univ. of Okla., 468 U.S. 85, 103 (1984); Charles F. Rule, Remarks at the Meeting of the International Trade Section & Antitrust Committee of the D.C. Bar (Nov. 29, 1988), *in* 7 Trade Reg. Rep. (CCH) ¶ 50,013.

little experience evaluating such contracts. Thus, full inquiry into the possible efficiency justifications for penalty contracts is desirable.

Although no authoritative statement of the elements of an efficiencies defense exists, courts and enforcement agencies have recognized three conditions, with the approval of many commentators.[157] First, the transaction underlying the penalty contract must produce substantial efficiencies. Second, no alternative less restrictive of competition should be reasonably available.[158] Third, the parties to the transaction bear the burden of proof.

IV. APPLICATION OF PROPOSED ANALYSIS: *UNITED SHOE*

In this Part, we apply our proposed criteria to the facts of the *United Shoe* case. United Shoe, a manufacturer of shoe machinery, entered into long-term leases with shoe manufacturers. The leases contained several restrictive provisions that raised the cost of switching to a rival supplier.

A. *Competitive Analysis*

1. *Market power and foreclosure.*

United Shoe clearly had market power, holding approximately 85 percent of the shoe machinery market, and from 95 to 100 percent with respect to certain machines. The market included no close substitutes, and had high entry barriers, as well as pervasive price discrimination.[159]

The leasing contracts foreclosed entry for ten years. Access-denying foreclosure was almost certain in machine lines where United Shoe's market share approached 100 percent. Because the demand for shoe machinery was stagnant, an entrant needed either to induce customers to switch from United Shoe, or to make sales to new manufacturers. However, new manu-

157. *See* 7 PHILLIP E. AREEDA, ANTITRUST LAW §§ 1502–1505 (1986); *see also* Joseph F. Brodley, *The Economic Goals of Antitrust: Efficiency, Consumer Welfare, and Technological Progress,* 62 N.Y.U. L. REV. 1020, 1037–41 (1987); Franklin M. Fisher, *Horizontal Mergers: Triage and Treatment,* ECON. PERSP., Fall 1987, at 23, 36-39; Robert Pitofsky, The Renaissance of Antitrust, Remarks at the House of the Association, Handler Lecture (Oct. 4, 1990), *in* 45 REC. OF THE ASS'N OF THE B. OF THE CITY OF N.Y. 851, 876–77 (1990).

158. The less restrictive alternative criterion has been a source of confusion. The concept does not require that a firm must choose the absolutely least restrictive mechanism to achieve the efficiency. As stated in United States v. Addyston Pipe & Steel Co., 85 F. 271, 282–83 (6th Cir. 1898), *aff'd,* 175 U.S. 211 (1899), a justifiable restraint under the Rule of Reason is one that is reasonably necessary. Determining what is reasonably necessary depends on the availability of significantly less restrictive alternatives, but a restraint is clearly unnecessary if the benefits from the transaction can be achieved without the anticompetitive restraint. *See generally* Kevin J. Arquit, Efficiency Considerations and Horizontal Restraints, Remarks to the National Health Care Lawyers Ass'n (Feb. 14, 1991) (on file with the *Stanford Law Review*) (discussing both acceptable and unacceptable efficiency claims, as well as reasonable restraints); Timothy J. Muris, *The New Rule of Reason,* 57 ANTITRUST L.J. 859 (1989). The ultimate test of legality is whether the procompetitive efficiencies outweigh anticompetitive harm; the greater the competitive restraint claimed necessary to induce the efficiencies, the less likely the efficiencies will overcome the restraint. *See generally* Antitrust Division, Dep't of Justice, Antitrust Guidelines for International Operations, 53 Fed. Reg. 21,584, 21,590 (1988) (discussing the role of efficiencies in the Antitrust Division's enforcement discretion).

159. *See* United States v. United Shoe Mach. Corp., 110 F. Supp. 295, 299–307 (D. Mass. 1953), *aff'd per curiam,* 347 U.S. 521 (1954); KAYSEN, supra note 69, at 46 (finding market shares of 96–100% in several machine lines).

facturers were not sufficiently numerous to sustain an entrant.[160] Furthermore, new firms would probably be reluctant to purchase machinery from an untried entrant. Foreclosure of customers open to potential entrants was also likely on other machine lines where United Shoe had lower market shares because United Shoe's rivals also used long-term leases, which, if they contained penalties, would preclude new entrants from bypassing United Shoe's penalty contracts even when a competitive alternative existed.

2. *Objective evidence of a penalty.*

Objective evidence of a contract penalty demonstrated that the leasing contracts imposed higher costs on customers who switched to a rival firm than on those who terminated leases for other reasons, and that United Shoe refused to offer customers a sale or short-term lease option.

Adverse treatment of switching customers. Customers who returned a leased machine to substitute a rival machine were required to pay a "commutation charge" equal to 25 to 50 percent of the remaining monthly lease rentals.[161] Customers who returned leased machines for other reasons did not bear such costs.[162] This discrepancy strongly suggests a contract penalty, because the preferential treatment of nonswitching customers provides an objective baseline for measuring the costs imposed on switching customers. Commutation charges were in fact paid on early returns of at least seventy machines.[163]

In addition, customers who did not comply with a lease requirement that machines be used at full capacity before shifting operations to a rival manufacturer's machine were required to pay larger rentals than customers who reduced usage for other reasons, such as reduced demand or replacement with hand labor.[164]

Other terms and conditions in the lease, which appeared neutral on their face, could be administered to penalize switching customers. For example, United Shoe gave customers a large rebate upon returning a machine, which increased with the length of the lease.[165] However, this "Right of Deduction" was lost if the customer breached the lease, for example, by switching to a rival; the lost rebate might be as high as 40 percent of United Shoe's original cost.[166]

The leases also contained an *in terrorem* clause, which gave United Shoe the right to cancel the lease as to all machines included in a leasing agree-

160. *See* KAYSEN, *supra* note 69, at 55 (despite annual customer turnover as high as 10–12% in some years, in most cases entrants would have to break existing lease ties to enter market successfully).

161. *United Shoe,* 110 F. Supp. at 320. Although the monthly rental was a minimum charge, United Shoe collected more revenue from the unit charges for each pair of shoes produced by its most important machines. *Id.* at 318–20.

162. *Id.* at 320.

163. KAYSEN, *supra* note 69, at 67.

164. *Id.* at 106.

165. *United Shoe,* 110 F. Supp. at 320–21.

166. *Id.*

ment upon breach as to any one machine.[167] Customers were vulnerable to claims of breach because the lease placed obligations on customers, such as the duty to repair, which United Shoe itself performed.[168]

Customers received various noncontractual benefits, such as free repair and service, which could be withdrawn at will. Customers contemplating switching justifiably worried that United Shoe might withhold these benefits, especially given United Shoe's avid concern about competition, close surveillance of customers, and stringent enforcement of the leasing contracts against switching customers.[169]

Refusal of sale or short-term lease option. United Shoe offered its important machines only on a long-term lease basis.[170] Unless a penalty scheme was involved, a supplier would ordinarily have no reason to refuse customers a desired option because they would presumably pay more for that choice. However, under a penalty leasing system, refusing a sale or short-term lease option prevents rival firms from bypassing the buying coalition between the monopolist and its customers. United Shoe's lease-only policy is subject to the alternative explanation that it facilitated price discrimination based on intensity of use. Although such an explanation justifies refusal of a *sale* option, it does not justify the refusal to offer a short-term *leasing* option, which would be equally effective in metering demand intensity.[171]

Stipulated damages exceeding actual damages. The *United Shoe* court did not consider whether the various charges and burdens placed on switching customers exceeded United Shoe's actual losses; the facts do not contain data for such a cost comparison. However, the evidence suggests that cus-

167. Although it is uncertain how many machines one lease might cover, a single lease could easily encompass multiple machines. *Id.* at 314. In earlier years, the leasing agreements explicitly authorized United Shoe to cancel *all leases* in the event of breach of *any* condition in *any* lease. *Id.* at 317.

168. Justice Brandeis, a former director and legal counsel for United Shoe, later explained that "the terror" underlying the *in terrorem* clause was caused by the lack of competing machinery manufacturers. *See Nomination of Louis D. Brandeis: Hearings Before the Subcomm. of the Senate Comm. on the Judiciary,* 64th Cong., 1st Sess. 222 (1916) [hereinafter *Brandeis Hearings*].

169. *See* KAYSEN, *supra* note 69, at 106. Kaysen notes that when a customer used a competitor's machine, "all applicable lease terms [we]re 'enforced' to the full," while "there [wa]s no enforcement" of lease provisions against customers who solely used United Shoe's machines. *Id.*

170. *Id.* at 32, 33.

171. As previously discussed, Professor Waldman suggests that an additional reason for United Shoe's refusal of a sale option was to prevent the development of a second hand market that would compete with the original equipment market. *See* Waldman, *supra* note 145, at 4, 28–31. However, a *short-term* leasing option would also have achieved this goal. *See* text accompanying note 145 *supra.*

One might argue that United Shoe would have found that a short-term lease would decrease incentives for consumer maintenance and would raise contracting costs as compared with a long-term lease. Neither of these reasons sufficiently explains United Shoe's behavior. United Shoe itself assumed the duty of primary maintenance and repair. Furthermore, United Shoe's hundreds of "roadmen," who closely scrutinized shoe manufacturers in frequent visitations, could have quickly detected inadequate customer care of its machines. *See* United States v. United Shoe Mach. Corp., 110 F. Supp. 295, 332 (D. Mass. 1953), *aff'd per curiam,* 347 U.S. 521 (1954). Although the periodic renewal of short-term contracts entails some costs, these costs are not necessarily large. IBM, for example, used 30-day contracts in computer equipment leases, lengthening the term only in response to competitive pressures from rival manufacturers. *See* Telex Corp. v. IBM, 510 F.2d 894 (10th Cir.), *cert. dismissed,* 423 U.S. 802 (1975).

tomers were threatened with charges exceeding United Shoe's actual losses, and the court should have examined the facts more closely. Given the strength of other objective factors, however, the absence of proof on this element does not preclude finding a penalty contract.

3. *Intent evidence.*

The evidence demonstrates that United Shoe sought to deter entry, but not to capture rent from potential entrants. Kaysen maintains that the various charges United Shoe imposed on switching customers were not designed to produce income, and did not do so.[172] Thus, rent capture does not appear to have been United Shoe's purpose in writing the contracts. However, United Shoe clearly sought to prevent entry by rival firms;[173] Kaysen notes that the costs imposed on switching customers were real and substantial deterrents to substitution of rival machines.[174] Thus, the justifiable inference is that United's contracts were intended to achieve their result of deterring entry.

United Shoe's 10-year lease term arguably reveals an intent to raise entrants' waiting costs. An initial determination must be made whether the lease period exceeded the length of time a customer would use a machine before replacement. In *United Shoe*, ten years was the average time a customer kept a machine, but approximately 20 to 25 percent of the customers returned leased machines within five years.[175] Thus, the lease bound these customers to United Shoe for twice as long as their actual use, thereby raising the waiting costs for entrants seeking access to these customers.

4. *Economic effects.*

United Shoe's leases had both deterrence and rent capture effects, but deterrence appears to have been the primary purpose. In addition, as discussed below, there was also a reverse free rider effect. The leasing restrictions were instrumental in deterring new entrants, preserving United Shoe's monopoly for over fifty years. Aghion-Bolton rent capture also occurred, but the effect was not substantial. In seventy documented instances, United Shoe required payments from switching customers that were not imposed on nonswitching customers. However, as noted above, these types of charges did not produce much revenue. Rather, the switching penalties operated more as draconian threats, indicating a deterrence goal.[176]

172. KAYSEN, *supra* note 69, at 69.

173. United Shoe was intensely, even obsessively, concerned with preventing competition. *See id.* at 114. When entry occurred, United Shoe frequently reduced prices, often for long periods and sometimes below total cost; once the entrant left the market, United Shoe raised its prices again. *United Shoe,* 110 F. Supp. at 326–27.

174. KAYSEN, *supra* note 69, at 69–70.

175. *United Shoe,* 110 F. Supp. at 319.

176. The use of penalty contracts to deter entry into a monopolized market, unless fully justified on efficiencies grounds, would constitute deliberately exclusionary conduct in violation of § 2 of the Sherman Act, while the contracts themselves would restrain trade in violation of § 1. *See id.* at 344–45.

Probably the key factor motivating customers to sign penalty contracts was a reverse free rider effect which made refusal costly and hazardous for any individual customer. Because joint action among United Shoe's hundreds of customers was not feasible, it was unlikely that an adequate customer base could be comprised of individuals refusing to sign the contracts.[177] Customers needed no positive inducements to sign the penalty contracts because they had no real alternative.

In addition, United received spillover gains from entry deterrence that were costless to customers, but created added surplus for United. Deterrence of entry in one machine line would tend to make it more costly to enter other machine lines by reducing cross-product synergies and economies of scope. In the short run this would not harm a customer who did not need the other machines that an entrant might produce, but it would directly benefit United by raising entry barriers on other product lines. Because most customers used relatively few of the machines manufactured by United Shoe, as few as thirty-seven out of a total of 342 machines,[178] United Shoe would be likely to gain more from deterrence of entry in a particular machine line than an individual customer would lose. United Shoe's additional revenues from these spillover effects created a surplus from which it could compensate customers.

B. *Efficiencies Defense*

The final step in applying our proposed criteria is to determine whether the contracts created socially productive efficiencies that outweighed their anticompetitive effects. A recent economic study by Masten and Snyder concludes that the restrictive leases in *United Shoe* increased United Shoe's incentive to maintain product quality and facilitated the flow of valuable information and know-how from United Shoe to its customers.[179] Using the Masten and Snyder findings, we can illustrate the application of an efficiencies defense.

1. *Product quality.*

The leasing contracts arguably increased United Shoe's incentive to supply high quality machines by overcoming the risk customers would otherwise face in buying a long lasting machine where quality cannot be ascertained in advance and enforcement of warranties is difficult. United

177. In 1911, an association of shoe manufacturers, comprising one-third of the industry, attempted joint action to assure competition in the shoe machinery market. Louis D. Brandeis, former director and legal counsel for United Shoe, represented the association. Brandeis strongly criticized the leasing contracts and United Shoe's attempts to raise the costs of a firm contemplating entry into the full line of shoe machinery manufacturing through coercion of the entrant's banks. *Brandeis Hearings, supra* note 168, at 219–20. The association's failure to achieve a vigorous competitive market in shoe machinery manufacturing probably discouraged further joint action by United Shoe's customers.

178. *See* Scott E. Masten & Edward A. Snyder, *United States v. United Shoe Machinery Corporation*: On the Merits 7 (Aug. 1990) (unpublished manuscript, on file with the *Stanford Law Review*).

179. *Id.* at 2–3.

Shoe's incentive to maintain quality would be weaker under a sale because the product is paid for at the time of sale and the high durability and infrequency of purchase reduce reputation effects.[180] The lease-only policy and the basing of rental payments on machine output overcame the moral hazard that United Shoe might fail to maintain quality by making United Shoe an economic partner with the shoe manufacturer in the successful operation of the machines. Accordingly, Masten and Snyder found that United Shoe used leasing and output-based rental charges more often for its complex machines where the probability and cost of breakdown were presumably greater.[181]

The product quality justification defense is problematic because other economic motivations would have compelled United Shoe to maintain quality. As Kaysen observes, the desire to sell additional machines to an old customer or the same type of machine to a new customer would also motivate quality control.[182] Moreover, as Masten and Snyder recognize, the manufacturer's enhanced incentive to maintain quality under a lease system is countered by the customer's reduced incentive to use the machines carefully.[183] Finally, United Shoe's more frequent use of leasing and output-based pricing on its important machines could be explained by discriminatory monopoly pricing and prevention of a second-hand market.[184]

Less restrictive means of promoting product quality were available to United Shoe. Short-term leases and leases without switching penalties would have provided more opportunities for customers to discipline United Shoe for failing to maintain quality, thereby promoting product quality more effectively than long-term penalty leases.[185] Moreover, the lease-only policy appears unnecessary to maintain quality, provided customers concerned about United Shoe's commitment to quality had an option to lease.

2. *Information transfer.*

The leasing contracts arguably facilitated the transfer of valuable information and know-how concerning shoe manufacturing to United Shoe's customers. This information was valuable whether or not customers used United Shoe's machines, because it related to factory-wide shoe manufacturing processes, such as factory configuration, rather than the operation of particular machines. Thus, United Shoe was subject to exploitation by customers who might switch to rival firms after receiving valuable manufacturing information. To prevent such exploitation, Masten and Snyder argue

180. *See* Wiley et al., *supra* note 145, at 714–15.

181. Masten & Snyder, *supra* note 178, at 23–27.

182. KAYSEN, *supra* note 69, at 194.

183. Masten & Snyder, *supra* note 178, at 15–16.

184. Output-based discriminatory pricing would naturally lead United Shoe to adopt a lease-only policy on its most valuable machines. Similarly, United Shoe's efforts to prevent the development of a second-hand market and erosion of its monopoly return would also focus on the most valuable machines. *See* Waldman, *supra* note 145, at 6–13.

185. *See* Masten & Snyder, *supra* note 178, at 17 (noting the superior incentives of short-term leases).

that United Shoe entered into long-term leasing contracts requiring customers to pay substantial damages for switching.[186]

Assuming this explanation is correct,[187] United Shoe's losses from consumer opportunism would not justify a 10-year contract with switching penalties. Depending on the timing of information transfer, separate sale of the information or a shorter lease without penalty are less restrictive alternatives. For example, if the information was supplied in full at the beginning of the lease, United Shoe could have sold the information separately. The customer could then not avoid paying for the information by switching to a lower cost supplier who does not bear the information supply cost.

Alternatively, if the information was supplied in small increments over time in proportion to machine use, then a leasing agreement with a rental fee based on machine output might have strengthened United Shoe's incentive to supply the information.[188] But, even then, a long-term lease with a switching penalty was not necessary. A short-term lease without penalty would provide a better incentive for the efficient supply of information by allowing the customer to verify the quality of the information and to discipline United Shoe by switching to a rival if United Shoe did not maintain information quality.[189] Thus, an efficiencies defense does not appear justified.

V. POLICY IMPLICATIONS

Antitrust authorities should be concerned about exclusionary penalty contracts.[190] By impeding new entry and innovation, penalty contracts severely constrain the dynamic process by which social wealth is created. Moreover, the economic theory of exclusionary penalty contracts is powerful, generally accepted, and rests on the settled strategic principle of credible commitment.[191] Recent antitrust decisions by U.S. courts, discussed above,

186. *Id.* at 18–19.

187. In fact, it is not entirely clear that United Shoe possessed superior knowledge about shoe manufacturing, as opposed to knowledge regarding the operation of its machines. It is questionable that United Shoe would have had superior manufacturing knowledge if it did not itself engage in manufacturing. Although United Shoe maintained a large staff of "roadmen" who visited shoe factories frequently, these representatives were primarily responsible for repair and maintenance, as well as for surveillance of customers to ensure compliance with contract terms. *See* United States v. United Shoe Mach. Corp., 110 F. Supp. 295, 332 (D. Mass. 1953), *aff'd per curiam*, 347 U.S. 521 (1954).

188. *See* Masten & Snyder, *supra* note 178, at 14–15.

189. An explanation based upon the transfer of information is problematic for another reason. Customer expropriation of United Shoe's information by switching to rival suppliers appears unlikely in view of United Shoe's dominant market share. Although the court claimed a shoe manufacturer could manufacture shoes without using any of United Shoe's machines, *United Shoe*, 110 F. Supp. at 339, there is no evidence that this ever occurred. Shoe manufacturers remained dependent on United Shoe for at least some of their essential machines; such dependency would constrain customer opportunism with respect to the information United Shoe supplied.

190. *See generally* Rasmusen, *supra* note 17 (discussing efficiency, price discrimination, and strategic reasons for exclusion).

191. *See, e.g.*, SCHELLING, *supra* note 9, at 24, 36; Mathias Francois Dewatripont, On the Theory of Commitment, with Applications to the Labor Market 117–24 (1986) (unpublished manuscript, on file with the *Stanford Law Review*).

which have sustained stipulated damage contracts without economically informed inquiry, may lead to greater use of such contracts. As suggested by the recent *Tetra Pak* case, the general legality of penalty contracts in European and other civil law countries creates the risk that monopolists may be using such contracts on a widespread basis. Both economic theory and judicial experience justify closer scrutiny of penalty contracts when used by firms with market power. We suggest the following guidelines for antitrust policy.

First, enforcement agencies and courts should recognize that penalty contracts in monopolistic markets raise antitrust concerns. Enforcement agencies should systematically investigate such contracts when used in monopoly, duopoly, or near duopoly markets. Both courts and agencies should recognize that penalty contracts in such markets may deter entry altogether or discourage entry through rent capture. In such cases, penalty contracts injure competition and maintain monopoly in possible violation of both sections 1 and 2 of the Sherman Act, as well as section 5 of the FTC Act.

Second, in examining exclusionary penalty contracts, antitrust enforcers should abandon myopic and mistaken assumptions that the contract law rule against contract penalties meets all antitrust needs. Antitrust officials should recognize that the contract rule does not effectively exclude penalty contracts. Although a permissive penalty rule may be appropriate in predominantly competitive markets, the needs of contract law in competitive markets differ fundamentally from the requirements of antitrust law in monopolized markets. Competition authorities in the EEC and Japan should recognize that the civil law acceptance of penalty provisions raises the risk of serious anticompetitive problems in contractual relationships between a monopolist and its customers.

Third, enforcement authorities should abandon the permissive approach used in recent antitrust cases, decided without knowledge of the economic theory of contract penalties. Instead, antitrust enforcers should investigate penalty contracts in monopolistic markets using the Aghion-Bolton analysis, gradually building empirical understanding through case-by-case adjudication. Section 5 of the FTC Act provides a particularly suitable vehicle for investigating penalty contracts, especially duopoly use of penalty contracts. Furthermore, FTC proceedings culminate in cease and desist orders, thereby avoiding the estoppel effect of a Sherman Act verdict, which becomes prima facie evidence of liability in subsequent private suits.[192] These limitations are desirable where findings of anticompetitive restraints are based on newly developed economic theory. However, when penalty contracts are coupled with other antitrust violations, as in *United Shoe*, Sherman Act prosecution is appropriate.

192. Under § 5 of the Clayton Act, a final judgment against a defendant in any case brought by the United States under the Sherman Act or the Clayton Act is prima facia evidence against such defendant in any subsequent case as to all matters in which the judgment would be an estoppel between the parties. 15 U.S.C. § 16 (1988).

Fourth, enforcement authorities should focus particular attention on regulated monopolies. Because profits in such industries are constrained, the regulated monopoly may find rent extraction from potential entrants an attractive alternative. A penalty contract that extracts rent from potential entrants, but does not prevent entry, enables the regulated monopoly to increase its return without current price increases or predatory tactics against existing rivals; customers will not object because they receive benefits. However, as described previously, such tactics harm consumers in the long run by reducing the return to innovation and new investment.

Fifth, the criteria outlined in Part III provide several useful filters or screens for identifying anticompetitive penalty contracts. The first screen requires showing that the monopolistic seller, or monopsonistic buyer, has market power and that the contracts substantially foreclose independent access to customers (or suppliers) not bound by the penalty contract. The second screen would utilize our proposed objective criteria (including differential treatment of switching customers, refusal of sale or short-term lease options, and stipulated damages exceeding actual losses) to identify cases of threshold anticompetitive risk. At least one type of intent evidence should also be considered: evidence that contract length exceeds the product's technological life. The final screen would involve a facial determination that the substantial efficiencies produced by the transaction are not outweighed by its anticompetitive harms, with the burden of proof on the participants. The enforcement agencies should undertake a more detailed inquiry only when the transaction passes through all of these screens.

CONCLUSION

An analysis of exclusionary penalty contracts deepens our understanding of the mechanisms by which strategic conduct can injure competition. The Aghion-Bolton model of bargaining through credible precommitment illustrates how, in an important class of transactions, anticompetitive strategic behavior can exclude or impede potential competitors. These transactions involve the use of monopoly power to defeat entry by more efficient firms or to capture the economic rent they would otherwise earn from their superior efficiency. The analysis of these issues demonstrates that strategic behavior and efficiency are alternative explanations for exclusionary effects, and only careful factual investigation, informed by sound economic theory, can determine which explanation prevails. That insight may lead to renewed respect for decisions such as *United Shoe* which sought through painstaking factual investigation, informed by skilled economic advice, to resolve that duality.