Using Insights from Game Theory:

Penalty Contracts and Monopolizing Strategies

by Joseph F. Brodley and Ching-to Albert Ma

Forty years separate the recent *Microsoft* consent decree[1] and the *United Shoe* decision.[2] But both cases involve a similar restraint — an agreement by which a monopolist binds its customers to pay a penalty for switching to a rival. In *United Shoe* the penalty consisted of various charges and conditions that imposed higher costs on shoe manufacturers who switched to a rival shoe machinery supplier. In *Microsoft* the penalty consisted of software licensing agreements that required computer manufacturers to pay a royalty for each computer sold, whether or not it contained a Microsoft operating system (e.g., MS DOS or WINDOWS), thus penalizing the computer manufacturer who switched to a rival operating system. In both cases an injunction was entered, prohibiting the penalty and shortening the period of the agreement.

In the four decades between these two cases there has been little antitrust concern about penalty contracts, either in the United States or elsewhere. The successful application of game theory in the economics of industrial organization, particularly as focused on monopoly behavior, has made clear that the use of penalty contracts can be a potent monopolizing device.

Joseph F. Brodley is Professor of Law and Economics and Kenison Distinguished Scholar of Law at Boston University, and Ching-to Albert Ma is Associate Professor of Economics at Boston University. Professor Brodley served briefly as a consultant to the Antitrust Division in the Microsoft *case, but the views expressed here reflect only those of the authors.* *The authors thank Deborah Minehart and Michael Riordan for helpful comments.* *For a more extensive discussion of the antitrust issues of contract penalties, see the authors' article,* Contract Penalties, Monopolizing Strategies and Antitrust Policy, *45* Stan. L. Rev. *1161 (1993).*

Why Are Penalty Contracts Bad?

A penalty contract requires the customer who would switch to a rival supplier during the term of the agreement to pay stipulated damages (or perform other conditions) more costly than the existing supplier's actual losses. While contract law in theory forbids penalties, the penalty rule is weakly enforced and frequently circumvented in common law jurisdictions (often through disguised penalties), and does not exist at all in civil law countries, which limit penalties only when unconscionable.

Operating much like a tax or import tariff, the penalty contract prevents a more efficient producer from entering the market (the *deterrence effect*), and extracts rent from super efficient producers who can enter the market despite the penalty (the *rent capture effect*). Thus, penalty contracts are bad for the same reason that protective import duties are bad. They deter entry by more efficient suppliers, and discourage entry and innovation even by highly efficient suppliers, who face lower investment returns.

Penalty contracts are more restrictive than exclusive dealing contracts. An exclusive dealing contract does not bar the entry of a more efficient firm able to fully supply a customer's needs, nor does it involve any rent capture effects.[3] Under an exclusive dealing contract the customer who breaches the contract in order to switch to an entrant need only compensate its existing supplier for lost profits (the normal contract remedy). Thus, a more efficient entrant can induce exclusive dealing customers to switch by pricing only slightly below the existing seller's costs, and the entrant then earns economic profit to the extent its costs are below those of the seller. Under a penalty contract, however, the entrant must reduce its price by an amount greater than the incumbent's lost profit (to reimburse the buyer for the penalty), resulting in either entry deterrence or rent capture.

It should be noted that penalty contracts are of antitrust concern only in monopolistic markets. In a competitive market penalty contracts have no adverse effects because customers have alternative suppliers, and thus penalty contracts will exist in competitive markets only if they increase efficiency. Hence, from here on we shall call the seller "the monopolist."

Why Do Customers Agree to Penalty Contracts?

It is easy to see why a monopolist supplier would want a penalty contract since the penalty either deters competitive entry or gives the monopolist a share of the entrant's profit. But why would customers ever agree to such an onerous provision, which reduces competition in their supply market? In fact, this is exactly the objection that Posner and Bork made to the *United Shoe* case.[4] They pointed out that even a monopolist cannot charge more than a monopoly price. Since the penalty contract appeared to be simply a disguised price increase, customers already paying a monopoly price would not sign such agreements unless they had some beneficial purpose other than obstructing entry.

The flaw in the Posner-Bork objection lies in its assumption that the burden of the penalty contract falls on the monopolist's customers, who must either pay the penalty or lose the benefit of competitive entry. The new economic analysis shows that the costs of obstructed entry may fall on the entrant.

The New Theory

Recent advances in game theory have allowed economists to study a wide range of strategic interactions between the monopolist, its customers, and potential rivals. Thus, using a model rigorously based on game theory, Phillipe Aghion and Patrick Bolton proved in their classic 1987 paper[5] that a penalty agreement can benefit both the monopolist *and* its customers. The penalty agreement enables the monopolist and its customers to form a bargaining coalition by which they can mutually exploit the lower costs of a more efficient entrant in markets where there is some likelihood of future entry. Just how this happens can be explained in simple intuitive terms.

Assume that a single monopolist sells to multiple customers at a monopoly price. If a rival producer fails to appear, clearly the monopolist will continue to enjoy monopoly profit. But if a rival producer tries to enter the market, the monopolist can respond vigorously, meeting any price reduction down to its own marginal cost. Suppose, however, that a more efficient producer enters with costs below those of the monopolist. The worst that can happen to the entrant (short of predation) is that the monopolist will reduce price to its own cost (say \$50). The entrant will then slightly undersell the monopolist (say \$49), capture as much of the market as its capacity permits, and earn economic profit based on its lower costs.

The situation changes drastically under a penalty contract. Suppose that the monopolist signs a penalty contract with its customers, requiring them to pay \$10 for switching to a rival firm. The entrant can no longer enter the market by slightly undercutting the monopolist's lowest price because customers are now bound by the contract to pay the agreed \$10 penalty (and of course the monopolist will not release the customers from the contract). So the entrant must lower its price sufficiently below \$50 to compensate the customer for the \$10 penalty (say to \$39). The result is to deter entry by rival producers with costs above \$39 even though their costs are lower than those of the monopolist. The super efficient entrant with costs below \$39 will still enter, but faces much lower return, discouraging future investment and innovation by this and other entrants.

The monopolist is clearly better off under a penalty agreement because it sets the penalty at an amount greater than its previous monopoly profit. Surprisingly, customers are also better off because the monopolist shares its penalty return with the customers by lowering its monopoly price *whether or not* a rival firm enters the market. The monopolist reduces its price in order to induce customers to sign the penalty contract. This is possible since the penalty exceeds the monopolist's previous monopoly profit, indicating that the monopolist remains better off even after giving some fraction of its profit to the entrant. Although the penalty contract makes it less likely that customers will benefit from entry by a lower cost supplier, customers are compensated by the general reduction in the monopoly price; and the monopolist earns supramonopoly returns

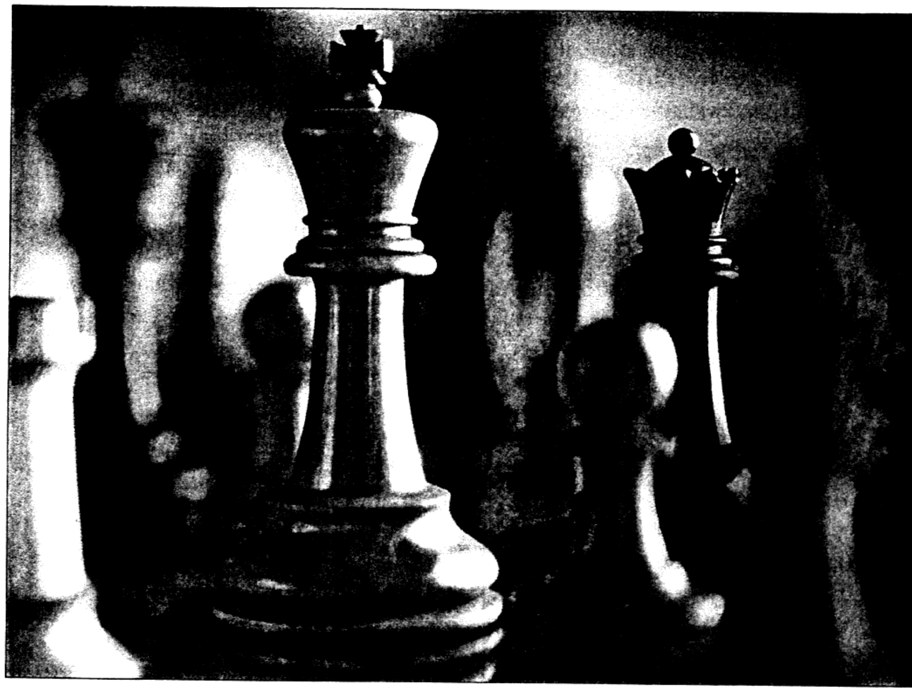

PHOTO: TONY STONE IMAGES

through the rent capture effect. By these means the coalition between the monopolist and its customers expropriates some of the economic return of more efficient potential entrants.

The penalty agreement enables the monopolist and its customers to form a bargaining coalition by which they can mutually exploit the lower costs of a more efficient entrant. . .

Why doesn't the entrant foil this arrangement by negotiating directly with customers, taking advantage of its lower costs? The reason is timing. The penalty contract is made *before* the entrant has appeared. Thus, when the entrant comes onto the scene, the customers are bound by the agreement. That is to say, the bargaining coalition works only as against potential entrants (or toehold firms that can serve added customers only by new investment). Existing suppliers with lower costs can counter any penalty contract strategy by undercutting the monopolist's best offer. For monopoly markets (or dominant firm markets with only a small competitive fringe) this is not an important qualification since the main force of competition is from potential entrants.[6]

Reverse Free Rider Effect

There is still another reason why customers sign penalty contracts with a monopolist. They may justly fear the consequences of rejecting the contract. Suppose that despite the monopolist's inducements, a customer prefers to switch to an entrant. If the industry has substantial sunk costs or other economies of scale, an entrant's viability will depend on its ability to recruit a sufficient customer base. Thus, unless the customer can be assured that enough other customers will also switch, the entrant will not survive. In that event the customer must return to the monopolist, who will be motivated to punish the customer as an example to other would-be defectors.

Thus, each customer is likely to hold back from quitting the monopolist, secretly hoping that enough other customers will sign up with the entrant to make entry viable. Entry then fails because too few customers switch to the entrant. (This is called a "reverse" free rider effect because the customers rather than attempting to board a figurative streetcar without paying, are refusing to get off a streetcar they no longer wish to ride until enough other riders have left to assure a viable alternative.)[7] Examples of reverse free rider effects appear in *United Shoe* and in two recent foreign antitrust decisions.[8]

Case Law Examples

Several vivid instances of anticompetitive penalty contracts exist. In the *United Shoe* case, an undoubted monopolist leased its important shoe manufacturing machines under ten year leases, imposed higher charges and other burdensome conditions on customers who sought to shift to rival suppliers, and refused to offer any sale or short term lease option to the ten year contract. This amounted to a penalty because the charges and undertakings that United Shoe imposed on customers who terminated leasing contracts for other reasons (such as machine upgrade and bankruptcy) provided an objective measure of United's actual losses. Thus, the additional costs that the United Shoe leases imposed on switching customers reflected a penalty because they exceeded United's actual losses as evidenced by its differential charges to switching and non-switching customers. At least partly as result of these leasing restrictions, United was able to maintain its monopoly position in the shoe machinery market for fifty years.

In *Tetra Pak*,[9] an EU antitrust case, the dominant manufacturer of milk packaging machines entered into long term leases (and sold machines with resale limited) with penalties for switching to rival suppliers. The European Commission held the agreement to be an abuse of dominant position in violation of Article 86, but without reference to the new economic theory of penalty contracts. In fact, the penalty nature of the agreements was indicated by the high level of the charges and the fact that customers who switched to rival machines faced much higher payments or other charges than those who returned machines for other reasons. The result was blockaded entry and continued high prices — a clear entry deterrence effect. In addition, rent capture almost certainly occurred when on several occasions the monopolist acquired firms that were preparing to enter the market. The purchase price was negotiated in the shadow of a penalty contract that obstructed new entry, so that the monopolist necessarily bought out the entrant at penalty-depressed prices.

A third striking example of a penalty contract appears in the *Hazeltine* cases,[10] where the owner of 200 radio patents licensed them as a block to radio manufacturers, charging a percentage of sales royalty for each radio the manufacturer sold, *whether or not* it incorporated any of the licensed patents. Thus, if a manufacturer-licensee wished to use competing patents, it would have to pay two royalties — one owing to Hazeltine and also one for the competing patents. This amounted to a penalty because a manufacturer would presumably switch to a rival patent only if the new patent was superior to the Hazeltine patent. In that event the manufacturer's sales would increase, raising Hazeltine's sales-based royalty above what it would have been had the manufacturer continued to use the Hazeltine patents. Strikingly, the better the rival technology, the higher would be Hazeltine's royalty.

The *Hazeltine* cases represent a clear case of rent capture that would exclude or reduce the economic return of an entrant with superior technology.[11] The *Microsoft* case involved similar facts. Computer manufacturers paid royalties to Microsoft for each computer shipped, whether or not they incorporated Microsoft software. The decree does not indicate whether computer manufacturers paid a flat royalty for each computer shipped or a percentage of sales royalty, but in either case Microsoft's royalty would rise as the value of the competing software increased. While in both *Hazeltine* and *Microsoft* a royalty based on end product output provides an effective method of price discrimination — in itself competitively ambiguous — by measuring the value of the license, facilitation of price discrimination cannot justify anticompetitive restraints, such as penalty contracts that suppress rival technologies.

Proving a Penalty Contract Case

Penalty contracts may constitute monopolizing conduct under Sherman Act Section 2, agreements restraining trade under Section 1, possible exclusive dealing under the Clayton Act, and unfair methods of competition under Section 5 of the FTC Act. Under any of these statutes the same type of proof would be needed.

Monopoly Power and Foreclosure. A penalty contract strategy can succeed only if the seller has monopoly power and the coalition between the seller and its customers can effectively foreclose independent access to customers. Normally, this requires proof of single firm monopoly power, but antitrust concerns can also arise in duopoly or near duopoly markets if the duopolists reach a stable accommodation (not necessarily agreement) to use similar penalty contracts. The duopoly case appears more suitable for challenge under Section 5 of the FTC Act.

The penalty contract must also foreclose a large enough percentage of customers to deprive the entrant of the customer base needed to operate at efficient scale. Critical here will be not only the percentage of customers bound by the penalty contract, but also the length of the contract, which measures the time an entrant must wait until enough contracts expire to provide a viable base of customers not bound by penalty contracts.

Proof of Penalty. The plaintiff must establish that a penalty exists. By a penalty we mean an agreed requirement that the switching customer pay stipulated damages or assume other burdensome conditions not imposed on non-switching customers, which are more costly than the losses actually suffered by the seller from early termination of the contract. Proof that the penalty exceeds the seller's losses presents practical difficulties, particularly in government enforcement actions where the issue concerns the general effect across all customers. Thus, we suggest several tests, both objective and subjective, to prove the existence of a penalty.

Objective Tests.

• ADVERSE TREATMENT OF SWITCHING CUSTOMERS. Customers who switch to rival firms are treated more adversely under the penalty clause than customers who terminate for other reasons. This evidences a penalty because normally the seller's losses are the same, whatever the reason for breach. This test was satisfied in *United Shoe*, where customers who reduced usage of a leased United Shoe machine in favor of a rival machine were treated more harshly than customers who reduced usage for other reasons, such as substitution of hand labor.

The test was also satisfied in *Tetra Pak*, where customers who sought to dispose of a purchased machine to switch to a rival machine received a lower repurchase price than customers who replaced their old machines with new Tetra Pak machines. In somewhat different context, the test was satisfied in the *Hazeltine* cases because a manufacturer-licensee who stopped producing radios paid no royalties, whereas a manufacturer who switched to a rival technology continued to pay royalties to Hazeltine.

• REFUSAL OF SHORT TERM SALE OR LEASE OPTION. Rigid adherence to a long term lease-only policy also evidences a penalty. If a switching penalty is not intended, the monopolist would ordinarily be willing to offer customers a viable sale or short term lease option at a sufficiently remunerative price. The customer would then be free to switch to an entrant offering the desired option; and the monopolist simply sacrifices revenue by being inflexible. Thus, the (unexplained) refusal to offer such an option at any feasible price evidences a penalty.

The test was satisfied in *United Shoe*, where customers were offered no alternative to the ten year leasing contract. A central provision of the *United Shoe* decree required that customers be offered a viable sale option to the United Shoe lease (which was shortened to five years). The test was also satisfied in *Tetra Pak*, where customers were offered no alternative to the long term lease or restrictive sales agreement.

• STIPULATED DAMAGES EXCEEDING ACTUAL LOSS. Proof that stipulated damages or alternative performance conditions exceed the monopolist's actual losses from termination of the contract directly evidences a penalty. However, proof of excessive stipulated damages presents difficulties, especially if courts follow contract law procedures, which assess stipulated damages both in terms of reasonable expectations of loss before breach and actual losses sustained after breach. Thus, we suggest that so far as possible courts limit their inquiry in antitrust cases to actual losses sustained after contract termination. This removes the speculative element that complicates breach of contract cases.

An ex post standard also facilitates antitrust compliance. The monopoly firm can avoid inadvertently entering into a penalty contract by (1) setting stipulated damages at a facially reasonable amount, and (2) limiting its demand in cases of breach to its actual losses. The feasibility of measuring stipulated damages in antitrust cases is illustrated in the *Northeastern Telephone* case,[12] where an antitrust court upheld a stipulated damages clause as non-exclusionary based on careful assessment that the stipulated damages did not exceed actual losses. (Of course, the monopoly firm should also take care not to violate any of the other tests for a penalty contract.)

The plaintiff must establish that a penalty exists. By a penalty we mean an agreed requirement that the switching customer pay stipulated damages or assume other burdensome conditions not imposed on non-switching customers, which are more costly than the losses actually suffered by the seller from early termination of the contract.

Intent Evidence. When objective evidence is indecisive, the court should consider other evidence tending to show that the monopolist set stipulated damages (or alternative performance conditions) with the deliberate intent to impede new entry. Intent evidence is often ambiguous in antitrust cases. But it is more reliable in contract penalty cases because the conduct showing an intent to impede entry differs sharply from non-exclusionary conduct. Unless obstruction of entry is intended, the monopolist has no reason to focus on entry effects when setting stipulated damages, but will only be concerned about its own losses from contract termination.

By impeding new entry and innovation, penalty contracts inhibit the dynamic process by which social wealth is created. Enforcement agency attention is called for because the absence of enforcement in the past may have spawned the widespread use of such contracts, especially in civil law countries, where no doctrinal rule bars penalty contracts even in theory.

Evidence that the monopolist intentionally manipulated contract length to influence entry (or to raise the entrant's waiting costs) includes:

(1) lengthening the contract term with deliberate purpose to impede entry;

(2) determining contract length based on an assessment of the customer base needed by an entrant to achieve a viable scale of operations; and

(3) selecting a contract term longer than the product life cycle.

In *Tetra Pak*, for example, the company entered into nine year leases of its packaging machines in the Italian market where it was zealously resisting new entry, but used leases as short as three years in other markets where entry was not threatened. In addition, the existence of rapid technological change in the packaging machine market would drastically shorten the product life cycle, possibly making even a three year lease exclusionary.

It is also of interest that the *Microsoft* consent decree limits the term of future licensing agreements to one year. While the decree does not indicate the length of existing licensing contracts, clearly the product life cycle in this dynamic industry is very short.

Anticompetitive Effects and Efficiencies Defense

If the evidence shows that a firm with monopoly power or jointly acting duopoly firms have used penalty contracts and that these contracts effectively foreclose access to sufficient customers to support viable entry, the remaining issues would be injury to competition and an efficiencies defense. Whether the case is examined under Section 1 or 2 of the Sherman Act, Section 3 of the Clayton Act, or Section 5 of the FTC Act, the inquiry would follow a rule of reason approach. The government or other plaintiff would have to show that the penalty contracts served to create or maintain monopoly conditions or substantially restrained trade, either standing alone — as in *Hazeltine* — or in combination with other practices — as in *United Shoe* and *Tetra Pak*. An efficiencies defense would of course be available in a penalty contract case, as in other rule of reason cases.[13]

Policy Implications

By impeding new entry and innovation, penalty contracts inhibit the dynamic process by which social wealth is created. Enforcement agency attention is called for because the absence of enforcement in the past may have spawned the widespread use of such contracts, especially in civil law countries, where no doctrinal rule bars penalty contracts even in theory. Thus, enforcement agencies should systematically investigate penalty contracts when used in monopoly, duopoly, or near duopoly markets.

Enforcers in the United States should recognize that the common law contract rule that prohibits penalty contracts does not effectively bar such contracts. Enforcement investigations should use the new economic theory of penalty contracts, building empirical understanding through case-by-case assessment.

There are several screens enforcement authorities and corporate advisors can use to identify anticompetitive penalty contracts:

(1) The supplier has monopoly power, either alone, or as part of a two-firm duopoly.

(2) The contracts substantially foreclose

access to customers not bound by penalty contracts.

(3) The contracts contain stipulated damage or other performance conditions that could amount to a penalty, and one or more of the following criteria are present:

- Switching customers are treated adversely as compared with non-switching customers.
- The supplier refuses to offer customers a sale or short term lease option when there is reason to think customers would value such an option.
- Stipulated damages or other performance conditions exceed the monopolist's actual losses from terminating the contract, as determined after termination.
- The length of the contract exceeds the product life cycle.

(4) A facial examination of possible efficiencies defenses does not indicate that plausible efficiencies are likely to outweigh anticompetitive harm.

Cases that pass through these screens should then undergo more detailed investigation. Scrutiny of penalty contracts is important because monopoly firms may use them to defeat entry by more efficient producers or to take away the economic profit super efficient producers would otherwise earn. Prevention of anticompetitive penalty contracts thus helps to maintain incentives for new investment and innovation in monopoly markets, improving the performance of both the monopolist and its rivals, and enhancing social welfare. ●

---

[1] United States v. Microsoft Corp., Trade Reg. Rep. (CCH) ¶ 50,764 (D.D.C. 1994) (consent decree).

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

[3] However, if an entrant cannot fully supply the customer's needs, e.g., because some customers have a strong preference for the existing seller's brand, then an exclusive dealing contract may well exclude a more efficient rival.

[4] Richard A. Posner, ANTITRUST LAW: AN ECONOMIC PERSPECTIVE 203–05 (1976); ROBERT H. BORK, THE ANTITRUST PARADOX: A POLICY AT WAR WITH ITSELF 140–44, 325–26 (1978).

[5] Phillipe Aghion & Patrick Bolton, *Contracts as a Barrier to Entry*, 77 AM. ECON. REV. 388 (1987).

[6] The Aghion-Bolton analysis is now well accepted by other economists, and indeed has spawned a considerable sub-literature. *See, e.g.*, Richard J. Gilbert, *Mobility Barriers and the Value of Incumbency*, and Michael J. Katz, *Vertical Contractual Relations*, *in* HANDBOOK OF INDUSTRIAL ORGANIZATION 501, 709 (Richard Schmalensee & Robert D. Willig, eds. 1989); JEAN TIROLE, THE THEORY OF INDUSTRIAL ORGANIZATION 196–98 (1988); Tai-Yeong Chung, *On the Social Optimality of Liquidated Damages*, 8 J.L., ECON. & ORG. 280 (1992); Kathryn E. Spier & Michael D. Whinston, Legal Restrictions for Breach of Contract (Working paper, Harvard Dept. of Econ., May 17, 1993). *But see* Richard Innes & Richard J. Sexton, *Strategic Buyers and Exclusionary Contracts*, 84 AM. ECON. REV. 566 (1994) (proposing intricate and implausible counter-strategy by customer-entrant coalition).

[7] For economic authorities discussing the reverse free rider effect, see Aghion & Bolton, *supra* note 5, at 396–98; Eric B. Rasmusen et al., *Naked Exclusion*, 81 AM. ECON. REV. 1137 (1991); Stephen C. Salop, *Practices that (Credibly) Facilitate Oligopoly Coordination*, *in* NEW DEVELOPMENTS IN THE ANALYSIS OF MARKET STRUCTURE 265, 272–73, 278, 284 (Joseph E. Stiglitz & G. Frank Mathewson, eds. 1986).

[8] Thus, in *United Shoe* customers could have been induced to sign penalty contracts by fear of retaliation if they refused to sign and other customers failed to join them in their refusal. In view of United's extensive customer surveillance activities, customers dependent on United for a wide variety of shoe machines would be reluctant to switch to an entrant until the entrant's viability was assured, and United would have reason to reinforce that reluctance by punishing deviants as an example to other would-be defectors. *See also* Elopak Italia Srl v. Tetra Pak, 4 C.M.L.R. Antitrust Rep. 551 (Comm'n 1991) (similar case involving milk packaging machines); Laidlaw Waste Systems Ltd., Canadian Competition Tribunal, Jan. 20, 1992 (similar case involving solid waste collection).

[9] Elopak Italia Srl v. Tetra Pak, 4 C.M.L.R. Antitrust Rep. 551 (Comm'n 1991).

[10] 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).

[11] In the second Hazeltine case the Supreme Court held the agreement unlawful but not for the reasons discussed here. However, in a separate opinion Justice Harlan noted the incentive reducing effects of the royalty provision on new entrants. 395 U.S. 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)).

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

[13] *See* Scott E. Masten & Edward A. Snyder, *United States Versus United Shoe Machinery Corporation: On the Merits*, 36 J.L. & ECON. 33 (1993) (penalty contracts in *United Shoe* defensible on efficiencies grounds); *cf.* Joseph F. Brodley & Ching-to Albert Ma, *Contract Penalties, Monopolizing Strategies and Antitrust Policy*, 45 STAN. L. REV. 1161, 1209–11 (1993) (challenging efficiencies justification, especially since less restrictive alternatives were available to United Shoe).