

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/257138456>

Procompetitive Theories of Vertical Control

Article in *Hastings Law Journal* · March 1999

CITATIONS

6

READS

63

2 authors:



[Andy C. M. Chen](#)

Taiwan Fair Trade Commission

34 PUBLICATIONS 10 CITATIONS

[SEE PROFILE](#)



[Keith N. Hylton](#)

Boston University

117 PUBLICATIONS 1,177 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Transparency Issues in Competition Law [View project](#)



Standard Essential Patents in Telecommunications Industry [View project](#)

1-1999

Procompetitive Theories of Vertical Control

Andy C. M. Chen

Keith N. Hylton

Follow this and additional works at: https://repository.uchastings.edu/hastings_law_journal



Part of the [Law Commons](#)

Recommended Citation

Andy C. M. Chen and Keith N. Hylton, *Procompetitive Theories of Vertical Control*, 50 HASTINGS L.J. 573 (1999).

Available at: https://repository.uchastings.edu/hastings_law_journal/vol50/iss3/3

This Article is brought to you for free and open access by the Law Journals at UC Hastings Scholarship Repository. It has been accepted for inclusion in Hastings Law Journal by an authorized editor of UC Hastings Scholarship Repository.

Procompetitive Theories of Vertical Control

by

ANDY C. M. CHEN* *and* KEITH N. HYLTON**

Table of Contents

Introduction	575
I. Vertical Control in Competitive Markets: Vertical Integration and the Principal-Agent Dilemma	580
A. The Principal-Agent Dilemma and Adverse Selection	581
B. The Principal-Agent Dilemma and Moral Hazard	584
C. Vertical Integration and the Solution to the Principal-Agent Dilemma.....	585
II. Vertical Control in Competitive Markets: Vertical Integration and Transaction-Cost Considerations	586
A. The Costs of Searching for Qualified Contracting Partners	586
B. The Costs of Reduced Flexibility.....	587
C. The Costs of Opportunism	588
D. Vertical Integration as a Means to Minimize Transaction Costs.....	590
III. Vertical Control in Non-Competitive Markets: Vertical Integration and the Successive Monopoly Problem.....	592
A. Competitive Price and Output Level.....	592
B. Price and Output Levels Under a Monopolized Upstream Market.....	594
C. The Successive Monopoly Problem	595
D. Vertical Integration Under the Technology of Variable-Proportions Production	598
IV. Vertical Restraints as Substitutes for Vertical Integration in Competitive Markets	599
A. Exclusive Dealing Arrangements and the Adverse-	

*Assistant Professor of Law, Tamkang University, Taipei, Taiwan.

** Professor of Law, Boston University, Boston, Massachusetts.

Selection Problem.....	600
B. Vertical Restraints and the Problems of Moral Hazard...	600
(1) Vertical Restraints and the Free-Riding Problem	601
(2) Vertical Restraints and the Maintenance of Product Reputation or Brand Image	604
C. Vertical Restraints and the Mitigation of Transaction Costs	606
(1) Vertical Restraints and the Reduction of the Costs of Finding Qualified Contracting Partners	606
(2) Vertical Restraints and the Reduction of the Costs of Reduced Flexibility	608
(3) Vertical Restraints and the Reduction of the Costs of Opportunism	609
V. Vertical Restraints as Substitutes for Vertical Integration in Non-Competitive Markets	610
A. Franchise Fee and Royalties.....	610
B. Maximum Resale Price Maintenance and Territorial Restriction	611
C. Tying and Output Royalty	612
D. Policy Implications.....	613
VI. The Complementary Nature of Vertical Control—Toward a Theory of Optimal Franchising	614
A. Partial Integration.....	614
B. The Optimal Combinations of Vertical Restraints.....	616
C. Policy Implications.....	619
D. Some Implications for Doctrine.....	620
(1) Resale Price Maintenance.....	621
(2) Territorial Restrictions	622
(3) Vertical Integration	623
(4) The Tying Doctrine.....	624
Conclusion	626
Appendix.....	627
I. Vertical Integration and Variable-Proportions Productions.....	627
II. Vertical Control in Non-Competitive Markets: The Case of Monopsony	628
III. The Successive Monopoly Problem	631
IV. Externalities in the Manufacture-Distribution Relationship	632

Introduction

Few subjects in American antitrust law have undergone as many changes and generated as much debate among economists and lawyers as the regulation of vertical arrangements.¹ Because these arrangements often restrain the freedom of firms to determine their own resale prices or distribution strategies, it has been accepted for some time that a rigorous antitrust policy is necessary, in this area, to maintain competition.²

It was not until the 1960s that a group of economists and lawyers called the "Chicago School" began to question this conventional wisdom.³ They argued that a manufacturer is unlikely to use restrictive vertical arrangements for anticompetitive purposes, because those arrangements tend to increase distributors' market power or decrease consumers' demand for the manufacturer's product and, therefore, work against the manufacturer's self-interest.⁴ Thus, when those arrangements are observed, more frequently than not they are imposed by the manufacturer with reasonable business justifications.⁵ For example, consider minimum resale price maintenance between a manufacturer and its dealers. Once a manufacturer's product is sold to dealers at a wholesale price with a fixed margin, the manufacturer's profits are proportional to the

1. Alan A. Fisher & Richard Sciacca, *An Economic Analysis of Vertical Merger Enforcement Policy*, 6 RES. ON L. & ECON. 1, 3 (1984) ("Of all areas in antitrust, the differences between the legal interpretation and economic analysis seem greatest with respect to vertical relationships (both mergers and vertical restraints)."); John S. McGee & Lowell R. Bassett, *Vertical Integration Revisited*, 19 J.L. & ECON. 17, 18 (1976) ("It appears that the evolution of economic theories about vertical integration has led alternately from deepest fog into brief patches of sunlight, thence plunging again into gloom and confusion.").

2. See generally Robert Pitofsky, *In Defense of Discounters: The No-Frills Case for a Per Se Rule Against Vertical Price Fixing*, 71 GEO. L.J. 1487, 1488-94 (1983) (presenting arguments in support of traditional per se standard with respect to vertical restraints).

3. Among them, Robert Bork, Richard Posner, and Lester Telser are the best known. For the arguments for which these authors are known, see generally ROBERT H. BORK, *THE ANTITRUST PARADOX: A POLICY AT WAR WITH ITSELF* (1978); RICHARD A. POSNER, *ANTITRUST LAW: AN ECONOMIC PERSPECTIVE* (1976) [hereinafter POSNER, *ANTITRUST LAW*]; Lester G. Telser, *Why Should Manufacturers Want Fair Trade?*, 3 J.L. & ECON. 86 (1960). According to Bork, many of these scholars were influenced by the teaching of Aaron Director at University of Chicago Law School. See BORK, *supra* at ix. For a brief presentation of many of Director's views on antitrust, see generally Aaron Director & Edward H. Levi, *Law and the Future: Trade Regulation*, 51 NW. U. L. REV. 281 (1956). On the Chicago School, see generally Richard A. Posner, *The Chicago School of Antitrust Analysis*, 127 U. PA. L. REV. 925 (1979).

4. See BORK, *supra* note 3, at 290-98; POSNER, *ANTITRUST LAW*, *supra* note 3, at 148-50 (discussing retail price maintenance), 160-62 (discussing exclusive territories), 165.

5. BORK, *supra* note 3, at 290-98.

quantities demanded in the final market. Therefore, it is somewhat counter-intuitive, if not self-defeating, for a manufacturer to prohibit its dealers from engaging in price cutting, when such competition tends to increase overall sales of its product. According to the Chicago School, explanations other than the exploitation of market power must exist to explain why a manufacturer would willingly adopt a distribution arrangement that puts it at a competitive disadvantage. For example, Lester Telser, in his widely discussed article, suggests that a manufacturer uses resale price maintenance to avoid free-riding behavior among distributors.⁶ That is, in the absence of a restraint on price discounts, a distributor has a strong incentive to take advantage of any pre-sale services provided by its competitors to attract potential customers without having to bear the costs of service provision. If such an outlook is prevalent among distributors, no service will be provided and the manufacturer may suffer losses if pre-sale services are deemed necessary for its customers.⁷

The procompetitive theories advocated by the Chicago School have not only changed conventional judicial thinking about vertical restrictive arrangements,⁸ but have also improved our understanding of the competitive strategies employed in product distribution.⁹ Recent developments in economic theory have in several respects revised and expanded the Chicagoans' theme to include new elements in a more general analytical framework.¹⁰ Although debate on the

6. See, e.g., Telser, *supra* note 3, at 91-95.

7. For a detailed discussion of the free-riding hypothesis, see *infra* Part VI.B.

8. The Chicago School influence is obvious in the economic rationales provided in several recent Supreme Court antitrust decisions. See, e.g., *Atlantic Richfield Co. v. USA Petroleum Co.*, 495 U.S. 328, 343 n.13, 345 (1990); *Business Elec. Corp. v. Sharp Elec. Corp.*, 485 U.S. 717, 729-30 n.3, 730-31 n.4 (1988); *Matsushita Elec. Indus. Co. v. Zenith Radio Corp.*, 475 U.S. 574, 589, 591-92 n.15, 594-95 (1986); *Continental T.V., Inc. v. G.T.E. Sylvania, Inc.*, 433 U.S. 36, 54-56 (1977). See also Frank H. Easterbrook, *Ignorance and Antitrust*, in *ANTITRUST, INNOVATION AND COMPETITIVENESS* 119, 119 (Thomas M. Jorde & David J. Teece eds., 1992) ("[Judges] have accepted the proposition that antitrust policy divorced from economics would be a calamity, and an antitrust policy conjoined with some inconsistent social policy would be incoherent and ineffectual." (citation omitted)); William H. Page, *The Chicago School and the Evolution of Antitrust: Characterization, Antitrust Injury and Evidentiary Sufficiency*, 75 VA. L. REV. 1221, 1226 (1989) ("[Chicago models] have become a shared conceptual basis for decision in antitrust cases for a substantial part of the federal judiciary.").

9. Most of the modern economic analyses of vertical control methods take the Chicago School analyses as a starting point. See, e.g., ROGER D. BLAIR & DAVID L. KASERMAN, *ANTITRUST ECONOMICS* 349-55 (1985) (surveying the economic literature on minimum resale price maintenance, starting with the Telser analysis).

10. See, e.g., Michael L. Katz, *Vertical Contractual Relations*, in 1 *HANDBOOK OF INDUSTRIAL ORGANIZATION* 655 (Richard Schmalensee & Robert D. Willig eds., 1989) (presenting a theoretical analysis of vertical contracts). In addition to the theoretical

extent to which those theories are applicable still continues, the center of the debate today focuses not so much on their theoretical validity as the normative suggestion that all types of vertical arrangements be treated as per se legal under the antitrust laws.¹¹ It would be a stretch to say that the Chicago School and its progeny will eventually lead the courts to conclude that vertical arrangements are always beneficial to market competition. But by and large, they have reshaped antitrust jurisprudence in this area and will continue to play a part in forming the theoretical basis of court decisions.

In this paper, we survey procompetitive theories of vertical arrangements, and examine how these theories could be organized to aid interpretation of antitrust law. Given the ever-expanding body of literature in this area, this is a daunting task which we cannot hope to complete in a single article. Nevertheless, we find a market-structure-based survey a helpful approach.¹² We have introduced various procompetitive theories in connection with the market structures in which they are likely to be applicable, though we do not claim they could only operate within a specific market context. Our approach

work reflected in the Katz survey, several scholars have taken a more policy-oriented approach toward expanding the Chicago framework. For example, Professor Steven Salop has argued in several articles that many vertical restrictions could be used strategically to increase the costs of competitors. See Steven C. Salop & David T. Scheffman, *Raising Rivals' Costs*, 73 AM. ECON. REV. 267 (1983); Thomas G. Krattenmaker & Steven C. Salop, *Anticompetitive Exclusion: Raising Rivals' Costs to Achieve Power Over Price*, 96 YALE L.J. 209, 268 (1986). Another line of policy-oriented "Post-Chicago" work has stressed the importance of imperfect information in assessing antitrust policy. See, e.g., Richard Craswell, *Tying Requirements in Competitive Markets: The Consumer Protection Issues*, 62 B.U. L. REV. 661, 671-81 (1982). For policy-oriented critiques of the Chicago School, from a "post-Chicago" perspective, see Lawrence A. Sullivan, *Post-Chicago Economics: Economists, Lawyers, Judges, and Enforcement Officials in a Less Determinate Theoretical World*, 63 ANTITRUST L.J. 669 (1995); Herbert Hovenkamp, *Antitrust Policy After Chicago*, 84 MICH. L. REV. 213 (1985).

11. For the classic argument in favor of per se legality in the area of vertical restraints, see BORK, *supra* note 3, at 288-98. Posner, after initially supporting a rule of reason approach to vertical restraints, came around to Bork's view that per se legality should be the rule. See Richard A. Posner, *The Rule of Reason and the Economic Approach: Reflections on the Sylvania Decision*, 45 U. CHI. L. REV. 1 (1977)(discussing Posner's initial, cautious, rule-of-reason argument); Richard A. Posner, *The Next Step in the Antitrust Treatment of Restricted Distribution: Per Se Legality*, 48 U. CHI. L. REV. 6 (1981)(discussing Posner's later, more aggressive stance favoring per se legality). Easterbrook has also moved from a cautious rule-of-reason argument toward favoring something very close to per se legality. See generally Frank H. Easterbrook, *Maximum Price Fixing*, 48 U. CHI. L. REV. 886 (1981)(stating that the rule of reason should apply to maximum price fixing); Frank H. Easterbrook, *The Limits of Antitrust*, 63 TEX. L. REV. 1 (1984)(urging courts to adopt "filters" or bright-line rules which avoid finding vertical restraints unlawful).

12. For a similar though somewhat dated approach, see generally, FREDERICK R. WARREN-BOULTON, *VERTICAL CONTROL OF MARKETS* (1978).

should aid antitrust courts in setting priorities among the factors to be considered in rule of reason analysis, and provide them a relatively clear path to follow when analyzing cases under the rule of reason.

Throughout the paper, we use the term "vertical control" to denote two general types of vertical arrangement—vertical integration and vertical restraints. Most of the procompetitive theories are based on a fundamental incentive divergence observed in the manufacturer-distributor relationship. In a nutshell, when goods are sold through two markets successively and separately, the manufacturer and its distributors may have different ideas regarding the ideal distribution pattern. For example, what appears to be the optimal price from the perspective of the distributors may be too high for the manufacturer had it sold the goods directly to customers. Vertical arrangements are often employed by the manufacturer to "control" the incentive of the distributor to deviate from the manufacturer's ideal distribution pattern at the downstream market. Accordingly, *vertical integration* might be thought of as a firm's decision to exercise control over the whole production and distribution process through acquisition of ownership, internal expansion, or initial formation.¹³ *Vertical restraints*, on the other hand, denote a market-oriented type of control—usually a discrete contractual arrangement between two independent firms.¹⁴ Specifically, we discuss six types of vertical restraints in this paper: franchise fee, royalties, resale price maintenance, territorial restriction, tying, and exclusive dealing.

In general, Ronald Coase's insight on the existence of firms provides a starting point for understanding why vertical control is observed in the market. In his classic, *The Nature of the Firm*,¹⁵ Coase raised and answered the following question: why do firms vertically

13. See Martin K. Perry, *Vertical Integration: Determinants and Effects*, in 1 HANDBOOK OF INDUSTRIAL ORGANIZATION 183, 186-87 (Richard Schmalensee & Robert D. Willig eds., 1989).

14. Generally, the distinction between vertical integration and vertical restraints may not be as clear as the one used here. The term "vertical integration" has been applied in particular by economists to refer to arrangements with dissimilar structural features. For example, under the definition by Blair and Kaserman, vertical integration refers to the types of transactions in which "[a] firm transfers internally from one department to another a commodity which could be sold in the market without major adaptation." With this definition, they conclude that "[a]ll firms are vertically integrated to some degree [But][n]o firm is totally vertically integrated." BLAIR & KASERMAN, *supra* note 9, at 283-84 (citation omitted). Alternatively, vertical integration has been used by Tirole rather broadly to include vertical restraints. According to Tirole, a firm could be labeled as vertically integrated so long as the firm could "control[] (directly or indirectly) all the decisions made by the vertical structure." JEAN TIROLE, *THE THEORY OF INDUSTRIAL ORGANIZATION* 170 (1989).

15. R.H. Coase, *The Nature of the Firm*, 4 *ECONOMICA* 386, 394-95 (1937).

integrate, rather than use the market, given the efficient market implied by the model of perfect competition? Coase's argument, put simply, is that a firm expands when the cost of expansion is less than that of relying on outsourcing. In the manufacturer-distributor setting, Coase's theory implies that a manufacturer will choose not to use a sales agent when doing so is more costly than direct sale.

Two questions follow from this observation. First, what are the costs associated with use of the market? Second, how can vertical control be applied to mitigate those costs? We examine these questions in Parts I and II. In particular, the costs associated with the *principal-agent problem* and its resolution through the use of vertical integration will be addressed. In contrast, Part III applies the analysis of vertical integration to markets that are not perfectly competitive. Emphasis will be placed on the economic effects of the *successive-monopoly problem*, and how vertical integration can improve welfare.

Parts IV and V focus on showing the equivalence of vertical integration and various vertical restraints in terms of their ability to reduce the costs of distribution. In Part IV, we examine how vertical restraints serve as "substitutes" for vertical integration in reducing distribution costs in competitive markets. Specifically, we describe how vertical restraints serve to mitigate the free-riding problem and to maintain product reputation. Part V is devoted to establishing the equivalence of vertical integration and vertical restraints in controlling the successive monopoly problem. From the manufacturers' perspective, this equivalence or interchangeability is beneficial when there is a need to control the distributor's incentives, but vertical integration is prohibitively costly or fails to pass antitrust law hurdles.

In contrast with the discussion in Parts IV and V, Part VI deals with the complementarity of vertical control mechanisms. Given the equivalence of vertical integration and other contractual restraints as controlling mechanisms, why is mixed use of these arrangements more frequently encountered in practice? As the analysis in Part VI will demonstrate, informational asymmetry and market uncertainty are the major factors that spur a manufacturer to bundle vertical integration with other vertical restraints as a package for the distribution of its products or services.

We close the discussion of complementarity by briefly setting out the implications of this survey for antitrust doctrine. At the risk of oversimplifying this project, perhaps the best way to think about the practical import of this paper is by contrasting it with a rather simplified version of the Chicago School approach. While the Chicago School analysis, under the early influence of Director and

Bork, has generally suggested a *per se* legality approach to vertical control mechanisms,¹⁶ that is not the intended or primary implication of this paper. Our primary aim is to set out the economic considerations in rule of reason analysis of vertical control mechanisms. In other words, instead of pointing courts to a particular conclusion, we have aimed to provide a flashlight, of sorts, for courts to use in navigating the dark tunnels of the reasonableness inquiry.

On a secondary level, this analysis implies that courts should adopt a functionally oriented, rule of reason test for vertical control mechanisms. Antitrust courts have been moving in the direction of rule of reason analysis in the area of vertical restraints, but they still have a long way to go before reaching the stage we think appropriate. Moreover, the rule of reason tests that have been developed in some areas are for the most part not functionally-oriented, in the sense that key legal distinctions are not based on a consideration of the economic function of the vertical control mechanism.

I. Vertical Control in Competitive Markets: Vertical Integration and the Principal-Agent Dilemma

One of the features of a *perfectly competitive* market is that firms will compete based on perfect information.¹⁷ Under this assumption, a manufacturer effortlessly acquires knowledge regarding customer preferences, and uses that information to adjust its competitive strategies in order to offer more attractive terms or to provide better services than its rivals do. Moreover, perfect information enables the manufacturer to respond to any increase or decrease in market demand by entering or exiting the market without delay.

To be sure, this assumption may not always hold. In fact, most of the information necessary for decision making takes time to collect. When the resources required for information collection are significant, selling through an independent dealer who is already operating in the market and is familiar with regional demand conditions is frequently the less costly way of doing business for the manufacturer.

However, sale by independent dealers has its shortcomings too. For example, finding a qualified dealer and negotiating a contract acceptable to both parties is a time-consuming process. And due to the problem of imperfect information, the manufacturer may not be able to find out beforehand the dealer's hidden flaws, such as a tendency to renege on contractual obligations. Consequently, the manufacturer will need to constantly monitor the performance of the

16. See *Director & Levi*, *supra* note 3, at 290; BORK, *supra* note 3, at 288-98.

17. See MICHAEL PARKIN, *ECONOMICS* 281 (1990).

distribution agreement after sale responsibilities are assigned to the downstream firms. This problem complicates Coase's theory,¹⁸ in the sense that the tradeoff between the costs of direct sale and sale by a dealer becomes harder to assess. In cases where the required information for evaluating the tradeoff is unavailable, the outcome of agent sales could be sub-optimal for the manufacturer. The conflict between the need to use a distribution system and the uncertainty arising from informational asymmetry can be described as a *principal-agent* dilemma.¹⁹ The costs inflicted upon the manufacturer can be summarized under two general categories: the costs arising from the process of *adverse selection* and those associated with the problem of *moral hazard*.²⁰

A. The Principal-Agent Dilemma and Adverse Selection

Information asymmetry inherent in the principal-agent dilemma gives rise to the problem of adverse selection. In its most general definition, "adverse selection" describes an unfavorable outcome to a firm emerging from voluntary market transactions where "only consumers with the least desirable characteristics, that are unobservable to the firm, buy the firm's product."²¹ To demonstrate

18. See Coase, *supra* note 15, at 390-92. (We refer here to the general theorem that in the absence of transactions costs, parties will bargain themselves to the economically efficient allocation of resources.)

19. See ROBERT S. PINDYCK & DANIEL L. RUBINFELD, MICROECONOMICS 608 (3d ed. 1995). The term "principal-agent" here refers loosely to an employment relationship in which the welfare of one party (manufacturer-principal) is dependent on how the other party (distributor-agent) will act. *Id.* Under this definition, the principal-agent dilemma could simply be understood as the manufacturer's failure to prevent its distributors from doing something against its interest. Milgrom and Roberts, on the other hand, offer a definition of the principal-agent relationship that is more consistent with that in law. See PAUL MILGROM & JOHN ROBERTS, ECONOMICS, ORGANIZATION AND MANAGEMENT 170 (1992)(defining principal-agent relationships as "situations in which one individual (the agent) acts on behalf of another (the principal) and is supposed to advance the principal's goals").

20. A more detailed classification of principal-agent costs (or more commonly, "agency costs") has been provided by Michael Jensen and William Meckling. Using corporate managerial decisions as an example, they divide the costs into three groups. First, *monitoring costs*. This group includes the costs incurred by the principal (stockholders) in order to control the behavior of its agent (manager) through the devices such as budget or compensation restrictions. Second, *bonding costs*. It refers to the expenses that have to be incurred by the agent to signal its intention to act in the best interest of the corporation. Finally, *residual loss*. It is defined as the dollar value of the disparity between the agent's decision and the principal's maximized profits. See Michael C. Jensen & William H. Meckling, *Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure*, 3 J. FIN. ECON. 305, 308 (1976). See also Eugene Fama, *Agency Problems and the Theory of the Firm*, 88 J. POL. ECON. 288, 289 (1980).

21. DENNIS W. CARLTON & JEFFREY M. PERLOFF, MODERN INDUSTRIAL ORGANIZATION 919 (2d ed. 1994).

the problem more clearly, consider a transaction in which the buyer is the better-informed party. The insurance industry provides a good example.²² In most cases, consumers know more about their own health conditions and management, when they purchase insurance, than do the insurance companies. Suppose that an insurance company offers an insurance program that includes the expenses of pregnancy and delivery. It is reasonable to predict that people who are planning to have children will be more likely to purchase the insurance than people who are not. But the decision to have a child is hidden information that is frequently unavailable to the insurance company beforehand. As a result, in setting the premiums, the insurance company cannot *price-discriminate*²³ against its customers based on their respective needs for the insurance and can only charge an *average* premium to all its customers. The premium thereby determined tends to be lower than the benefit in dollars that could be derived from the program by customers who are planning to have children but it will be higher than the benefit to those who are not. Eventually, a large fraction of the insurance program will be comprised of customers belonging to the former group. This generates a less favorable outcome for the insurance company in the sense that it leads to lower profits for the company than would have been the case had the population of the insured been randomly selected.

Consider next the case where the buyer is the less-informed party in a transaction. That adverse selection can occur in this situation was demonstrated by George Akerlof with the example of the used-car market.²⁴ Assume that the used-car market consists of two types of cars, good and bad used cars ("lemons").²⁵ Good used cars are valued by consumers at the price of \$200, and lemons at \$100. Suppose that sellers know the quality of the cars they are selling but such knowledge would be available to buyers only after purchasing and driving the cars for a while. Assume further that buyers know their chances of getting a good used car and a lemon is half-half. With these assumptions, how much will a buyer pay for a used car?

Since the information regarding the quality of cars is unknown to the buyers when buying them, they would assume the cars to be of

22. See, e.g., MILGROM & ROBERTS, *supra* note 19, at 149; PINDYCK & RUBINFELD, *supra* note 19, at 596-97.

23. "Price discrimination" refers to the practice of charging different prices to different buyers, for the same item. For the various economic definitions, see JACK HIRSHLEIFER, *PRICE THEORY AND APPLICATIONS* 244-49 (4th ed. 1988).

24. See George A. Akerlof, *The Market for "Lemons": Quality Uncertainty and the Market Mechanism*, 84 Q.J. ECON. 488, 489-92 (1970).

25. This simplified example of Akerlof's paper is based on CARLTON & PERLOFF, *supra* note 21, at 560.

average quality. To a risk-neutral buyer, he or she will be willing to pay up to \$150 for a used car, as if the car were randomly selected.²⁶ Intuitively, consumers would prefer paying \$150 and hoping that they will get a good used car to paying \$200 and taking the risk of buying a lemon. That is, constrained by the lack of information regarding the quality of used cars, consumers tend to overvalue lemons but undervalue good used cars, which will in turn discourage the sellers from selling good used cars. Over time, the used-car market will become dominated by lemons. And after realizing that, consumers will further reduce the prices they are willing to pay for a used car to a point that only lemons will be sold on the market.²⁷

In the context of distribution arrangements, the problem of adverse selection occurs when the potential value of the manufacturer's product to a market is unknown to or difficult-to-acquire for the manufacturer. As a result, the constituents of the distribution chain are not optimally selected. Consider the case where a manufacturer of a famous brand decides to franchise its product into new markets. Suppose that precontractually, the prospective franchisees are better informed on the potential value of the franchise at each market than is the manufacturer. Incomplete knowledge on the manufacturer's side may cause the prices for the franchise right to be hard to determine. In a similar vein as the insurance-market example, the manufacturer may choose to sell its franchise at an average price on the assumption that demand at each market will be medium. Likewise, with the price so set, dealers facing low demand for the product will be more likely to turn down the offers to become a franchisee and the chain will consist predominantly of high-demand franchisees. This may not be the optimal outcome for the manufacturer, in the sense that the overall sales could be lower than when the chain includes both low-demand and high-demand franchisees.²⁸

26. The highest price that a buyer will pay for a randomly selected car will be his or her expected value from the transaction. In the present case, it will be equal to $\$200 \times 0.5 + \100×0.5 .

27. For a graphical illustration of this outcome here, see PINDYCK & RUBINFELD, *supra* note 19, at 595.

28. This argument is similar to the reasons offered by some economists for believing that price discrimination enhances consumer welfare. If a firm could charge each customer according to his reservation price, i.e., the highest price he is willing to pay for an additional unit of good, the overall sales would be larger than that when the firm charges a uniform price for all customers. The increase in sales means more customers will be served and therefore implies a positive welfare effect. See, e.g., JOAN ROBINSON, *THE ECONOMICS OF IMPERFECT COMPETITION* 206 (2d ed. 1969); PINDYCK & RUBINFELD, *supra* note 19, at 364-66; ROBERT H. BORK, *THE ANTITRUST PARADOX* 398 (1993). Analogously, if the manufacturer can distinguish high-demand from low-demand franchisees and charge them accordingly for the licensing of the franchise, more of its

Following the above analysis, it is easy to see how the "lemons problem" could occur in manufacturer-dealer relationships. Consider the case where several franchisors of the same industry take bids for franchisees in a certain market. Assume that a "good" franchise is defined as a chain selling a high-quality product and, therefore, will be licensed at a higher price than a "bad" franchise. Assume further that the quality of the franchised product is known to the franchisee only after he becomes part of the chain and after the product is sold on the market for a while. Again, due to informational asymmetry, the price that a prospective franchisee is willing to pay for a franchise would be too low for the "good" franchise and too high for the "bad" ones. Eventually, "good" franchises will be driven out of the market and only "bad" ones will be available in the market.

B. The Principal-Agent Dilemma and Moral Hazard

Unlike adverse selection that occurs before a vertical relationship is formed, the problem of moral hazard refers to a form of postcontractual opportunism that arises from the disparity between the manufacturer and distributors in the objectives pursued. And due to the difficulty in observing the disparity, one of the parties has an incentive to pursue its own interest at the expense of the other.²⁹ The term "moral hazard" (like adverse selection) comes from the insurance industry.³⁰ In the insurance context, it refers to the tendency of the insured not to take sufficient precaution to minimize losses or to avoid accidents from happening after purchasing insurance. Since it is difficult for the insurance company to know in advance the propensity of the insured to engage in risky behavior, it will not be able to correctly specify in the contract what types of precaution should be taken by the insured. As a result, the insured may have an incentive to be lax about precaution and consequently to file larger claims against the insurance company. For example, when an individual is fully insured against all costs for visits to the doctor, the lack of responsibility to share medical expenses may diminish his incentive to stay healthy and induce him to use all kinds of medical services as frequently as he can.³¹

In a vertical relationship, moral hazard arises in the form of renegeing on contractual obligations by the parties in the agreement. Similar to an insurance contract, when the information regarding the

products could be available to consumers at more locations.

29. See MILGROM & ROBERTS, *supra* note 19, at 167; PINDYCK & RUBINFELD, *supra* note 19, at 604-05.

30. On the definition of "moral hazard," see KARL H. BORSCH, *ECONOMICS OF INSURANCE* 317-18 (Knut K. Aase & Agnar Sandmo eds., 1990).

31. See MILGROM & ROBERTS, *supra* note 19, at 167-68.

characteristics of both parties is incomplete, the agreement made tends to be subject to manipulation and exploitation after taking effect. For example, a dealer may determine, based on its own cost consideration, to shirk on its duty under the distribution agreement to maintain product quality. This could be the result of decreasing market demand for the manufacturer's product, which makes shirking more profitable than abiding strictly by the contract.³² And because the potential for moral hazard cannot be fully verified in advance, additional costs have to be incurred by the manufacturer to ensure that the contract will be faithfully enforced. Despite the express terms in the agreement, the manufacturer may still have to make efforts to monitor or detect shirking on product quality or other obligations. When those costs are substantial, moral hazard could render the distribution arrangement inefficient.

C. Vertical Integration and the Solution to the Principal-Agent Dilemma

Since the problems of adverse selection and moral hazard arise from informational asymmetry, the solutions to these problems naturally involve the use of various institutional designs to induce the provision of the required information.³³ From this perspective, vertical integration could be the most direct and obvious method to serve this function. By bringing business operations under a unified governance structure, a firm can avoid the difficulty of obtaining information regarding the profitability of a particular market and, therefore, gain additional confidence in the estimates of the value of a franchise. The problem of adverse selection disappears. Also, the divergence of interests and the need to monitor and prevent moral hazard from occurring can be significantly eliminated by merging two firms with potentially divergent goals into one and consolidating the decision-making process.³⁴ In sum, vertical integration resolves the principal-agent dilemma by enabling firms to bypass the obstacle of

32. Equally applicable are those considerations to cases where a supplier is obliged to provide its dealers inputs with a minimum quality standard but the pursuance of self-interest induces it to do otherwise.

33. For a detailed discussion of the likely solutions to these problems, see MILGROM & ROBERTS, *supra* note 19, at 154-59 (adverse selection), 185-90 (moral hazard).

34. The problem of moral hazard, however, cannot be completely cured by vertical integration. Moral hazard could arise within a vertically integrated firm, leading to a misallocation of benefits within the organization. For example, internal transactions could encourage rent-seeking activities within an integrated firm. That is, a division manager could lobby the central manager to re-distribute more profits to his own division than would have been the case had the distribution been based on the profit-maximizing considerations. Or he could ask for "forgiveness" for any mistakes his division had made. See DAVID BESANKO ET AL., *THE ECONOMICS OF STRATEGY* 87-88 (1996). Inefficiencies of these types are collectively called by Milgrom and Roberts the "influence costs." For a detailed discussion, see MILGROM & ROBERTS, *supra* note 19, at 192-94.

informational asymmetry.

II. Vertical Control in Competitive Markets: Vertical Integration and Transaction-Cost Considerations

In the hypothetical world in which the model of perfect competition operates, firms incur no additional costs, other than their own production costs, in using independent dealers in their distribution channels. Within this environment, the now widely celebrated Coase Theorem predicts that both the upstream and downstream firms will negotiate and strike a deal that maximizes joint profits.³⁵ Nevertheless, transaction costs are frequently important factors that cannot be neglected by firms in designing the ideal distribution pattern. For example, searching for a qualified distributor that will fit the need of the manufacturer is usually time-consuming. Drafting a contract that sets out as broadly as possible all future contingencies is notoriously laborious, not to mention the resources invested in haggling during the negotiation process and the interpretation of contractual terms when disputes arise.

Based on Coase's insight, Williamson developed a framework to examine how transaction-cost considerations affect institutional design.³⁶ In particular, Williamson focuses on the costs associated with *ex ante* idiosyncratic or transaction-specific investments and *ex post* opportunistic actions and how these costs affect the choice between internal production and outsourcing. We discuss three types of transaction costs frequently mentioned by commentators,³⁷ and examine how they fit into the Williamsonian framework in this part. We will also examine how vertical integration could be employed to reduce those costs.

A. The Costs of Searching for Qualified Contracting Partners

As Blair and Kaserman have stated: "[a] sizable portion of the budgets of the purchasing and marketing departments of many corporations may be due to the presence of search costs involved in buying inputs and selling outputs."³⁸ The reasons why search costs are encountered in markets where candidates for distribution agreements are competitively supplied are related to the phenomenon of informational asymmetry, particularly the problem of

35. See generally R.H. Coase, *The Problem of Social Cost*, 3 J.L. & ECON. 1 (1960).

36. See generally OLIVER E. WILLIAMSON, *MARKETS AND HIERARCHIES* (1975).

37. See, e.g., BLAIR & KASERMAN, *supra* note 9, at 292-93 (1985); E. THOMAS SULLIVAN & JEFFREY L. HARRISON, *UNDERSTANDING ANTITRUST AND ITS ECONOMIC IMPLICATIONS* 267 (2d ed. 1994).

38. BLAIR & KASERMAN, *supra* note 9, at 292.

moral hazard. The costs incurred postcontractually by firms to monitor shirking on product quality by the other party could appear in the form of more cautious selection of partners precontractually. In other words, the concern for moral hazard may drive the manufacturer to conduct deeper investigations on the hidden characteristics of prospective dealers and thus prolong the period for the completion of contracts.

On the other hand, search costs could also arise from the idiosyncratic nature of the product. A typical model of a perfectly competitive market assumes the product sold to be homogeneous across the market.³⁹ Under this hypothesis, buyers are indifferent to the sources from which they purchase the product. But even if the physical attributes of the goods produced by firms are identical, they may not be homogeneous in all respects. For example, locational considerations could place the good produced by the neighboring supplier on top of the others. Uncertainty among buyers concerning the reputations of the sellers might lead them to favor the product sold by a particular seller as well.⁴⁰ When these possibilities are considered, the number of qualified candidates for a distribution agreement declines and search costs increase.

B. The Costs of Reduced Flexibility

Even if one assumes that search costs are trivial in competitive markets, the formation of a distribution agreement imposes the costs of reduced flexibility on both parties in the sense that the manufacturer and dealer are "[locked] into a predetermined pattern of behavior in order to assure the other party that it has not misrepresented its intended postsale [sic] performance."⁴¹ In general, two reasons may contribute to the emergence of such costs.

First, inflexibility results primarily from the limit of human minds in foreseeing future contingencies and the failure to take these contingencies into consideration when measuring the value of an agreement. In the words of Williamson, this is a misfortune caused by "bounded rationality."⁴² The adverse consequences of bounded

39. See PARKIN, *supra* note 17, at 281.

40. Through repeated purchases, for example, a buyer may develop a personal relationship and trust with a certain seller, and could treat the product sold by the seller to be different from its competitors.

41. BLAIR & KASERMAN, *supra* note 9, at 293.

42. See Oliver E. Williamson, *The Economics of Antitrust: Transaction Cost Considerations*, 122 U. PA. L. REV. 1439, 1444 (1974) (referring to "bounded rationality" as a phenomenon in which "[t]he capacity of the human mind for formulating and solving complex problems is very small compared with the size of the problems whose solution is required for objectively rational behavior in the real world" (quoting H. SIMON, *MODELS OF MAN* 198 (1957))).

rationality, in the manufacturer-distributor context, are realized when incidents occurring after the relationship forms render the performance of the original contract unprofitable, but the parties cannot reposition themselves accordingly. For example, when unanticipated changes in market conditions make direct sale the more appropriate marketing strategy, adhering to the original distribution agreement becomes costly for the manufacturer, and yet the manufacturer may practically be unable to renegotiate the agreement.

Second, inflexibility also occurs when investments during the existence of a contract are specifically made by a party to cater to the need of the other.⁴³ Due to their idiosyncrasies and low salvage value, the party that made the investments will in some situations be compelled to remain in the contractual relationship even if the other party has breached the contract. This phenomenon is described in economic parlance as the creation of "appropriable quasi rents" by specific investments (which will be discussed in more detail later).⁴⁴ At this point, a simple illustration is sufficient to understand how it generates costs, due to inflexibility. Consider again the case of franchising.⁴⁵ Suppose that McDonald's purchases and then leases the buildings for the operation of a business to its franchisees. In order to create its own brand image, the buildings are distinctively constructed and designed. Assume that McDonald's later finds out some of its franchisees fail to live up to the quality standard specified in the franchise contract. It is clear that McDonald's will feel more restrained in terminating the relationship with those franchisees than would have been the case had the buildings not been so specifically designed and could have been put to use as computer stores or office buildings.

C. The Costs of Opportunism

The costs of opportunism inherent in vertical relationships are related to the costs of reduced flexibility. Once the specific investments are made and the parties whom the investments serve realize that such investments are unique and will be difficult to be fully reimbursed by their "next best" alternative uses, they could easily "hold up" the parties who made the investments by forcing

43. See OLIVER E. WILLIAMSON, *THE ECONOMIC INSTITUTIONS OF CAPITALISM* 95-96 (1985) (classifying specific investments into four categories: site specificity, physical asset specificity, human asset specificity, and dedicated assets); BESANKO ET AL., *supra* note 34, at 113-14 (utilizing the same categories).

44. See Benjamin B. Klein et al., *Vertical Integration, Appropriable Rents, and the Competitive Contracting Process*, 21 J.L. & ECON. 297-307 (1978) and the numerical illustration in Part II.C.

45. The example is based on James A. Brickley & Frederick H. Dark, *The Choice of Organizational Form: The Case of Franchising*, 18 J. FIN. ECON. 401, 406-07 (1987).

renegotiation of a new agreement to pursue their interests.⁴⁶ To see this point, let's revisit the concept of appropriable quasi-rents, and the McDonald's example.⁴⁷

Suppose that the cost for McDonald's to purchase and design the building is \$10,000. To simplify our discussion, assume that it is the only cost that the franchisor has to incur to license an additional franchisee. Assume also that the next best use of this \$10,000 for the franchisor is to deposit it in a bank with an annual interest rate of 12%. To the franchisor, the *minimum* price or franchise fee that will induce it to *enter* into a franchise relationship will therefore be \$1,200 per year. Under this scenario, *economic rents* refers to the difference between this minimum amount and the *actual* franchise fee received by the franchisor. A franchisor can demand a positive economic rent if it has bargaining power over the franchisee. However, when the number of franchisors bidding for franchisees is large, we would expect that competition would drive McDonald's rent down to zero. In this light, the concept of rent is equivalent to that of "economic profits" in economics.⁴⁸

Consider next the meaning of "quasi-rents." Suppose that due to the specificity of the building, it now has few alternative uses. Suppose that the next best use of the building once it is constructed would generate annual earnings of \$1,000 for McDonald's. It is obvious that McDonald's will be better off continuing to franchise its business. And the *minimum* amount of revenues for McDonald's to remain in this relationship will also be \$1,000. The meaning of "quasi-rents" could be defined as the difference between the *actual* revenues (i.e., the franchise fee) that would be received under the original franchise contract and this minimum amount that would *prevent the franchisor from terminating and exiting* the contractual arrangement. In our example, the quasi-rents are \$200.

The existence of positive quasi-rents creates the incentive for the franchisees to act opportunistically. Knowing that McDonald's will accept an offer between \$1,200 to \$1,000, the franchisees may push McDonald's to either reduce the franchisee fee after the initial contract is formed or take the consequence of termination of the franchise contract. To McDonald's, this is similar to a "hold-up" by the franchisees in the sense that it would be deprived of the protection from the original contract.⁴⁹

46. Cf. WILLIAMSON, *supra* note 43, at 47 (defining opportunism as "self-interest seeking with guile"); Benjamin Klein, *Why Hold-Ups Occur: The Self-enforcing Range of Contractual Relationships*, 34 ECON. INQUIRY 444, 446 (1996).

47. The following discussion is based on BESANKO ET AL., *supra* note 34, at 114-16.

48. See *id.* at 115 n.19.

49. Conversely, franchisees could also be held up due to their investments in human

D. Vertical Integration as a Means to Minimize Transaction Costs

Williamson provides an analytical framework that explains how vertical integration could minimize transaction costs. According to Williamson, transaction costs consist of two main factors: *environmental factors* and *human factors*.⁵⁰ Market uncertainty and the costs of finding qualified contracting partners are grouped into the former category. Bounded rationality coupled with the costs of opportunism, on the other hand, is classified under the latter. Williamson shows how the interaction among these four factors creates transaction costs that handicap the functioning of the market.

Consider first market uncertainty, a problem to a large extent the product of informational asymmetry. To overcome the inefficiency created by this, a manufacturer could enter into a long-term contract with its distributors and thereby reduce the costs of having to find qualified contracting partners frequently. On the one hand, long-term contracts could help the firms internalize their specific investments in physical or human capital, and could thus reduce the amounts of appropriable quasi-rents.⁵¹ But on the other hand, due to the constraint of bounded rationality in stipulating future contingencies, transactors could take advantage of the incompleteness or vagueness of long-term contracts to act opportunistically and to hold up the parties who are locked in the relationship because of their specific investments.⁵² And needless to say, long-term contracts will increase the costs of reduced flexibility.

The second option to reduce market uncertainty would be to sign a short-term contract subject to periodic renewals by both parties. Such arrangements allow transactors to take new information into account during the renewal periods and, therefore, to redraw the provisions in a way that could better satisfy mutual needs. The costs of reduced flexibility and opportunism can thereby be reduced. However, the main drawback of using a short-term contract is that it creates what Williamson calls the problem of "small numbers."⁵³ Aside from the problem of finding a qualified transacting partner precontractually, the term "small numbers" is used by Williamson to refer specifically to the limited supply of qualified partners during the renewal periods. For example, as cooperation proceeds, economies of scope gained through specialized production or cost savings in

capital that are specific to meet the franchisor's need. See Amy Barrett, *Indigestion at Taco Bell*, BUS. WK., Dec. 14, 1992, at 66-67.

50. See generally WILLIAMSON, *supra* note 36, at 20-40.

51. See Brickley & Dark, *supra* note 45, at 409.

52. See WILLIAMSON, *supra* note 36, at 17; Benjamin Klein, *Vertical Integration as Organizational Ownership*, 4 J.L. ECON. & ORG. 199, 201 (1988).

53. See WILLIAMSON, *supra* note 36, at 26-28.

connection with "first-mover" advantages may induce both parties to favor the original associate when renewing the contract.⁵⁴ Even if bidders are competitively supplied *ex ante*, the advantage thereby gained could differentiate the products or services provided by the initial winner from those of competitors *ex post*, and increase their costs of finding alternative transacting partners.

Vertical integration can reduce these costs.⁵⁵ The cost of finding qualified contracting partners need not be incurred with integrated production and thus no problem of "small numbers" is encountered. Internal production also yields the advantage of more extensive control over input combinations, eliminating the need for firms to attempt to construct a fundamental contract governing all likely contingencies, and increasing the flexibility to adapt to changes in market conditions. And since the relationship between the affected parties is turned from rivalry to cooperation, vertical integration could diminish the incentive to take advantage of specific investments to hold up contracting partners and, as a result, reduce the costs of monitoring opportunism.⁵⁶

54. *See id.* at 28. The concept of "first-mover advantage" refers to a phenomenon that the initial contracting partner gains a competitive advantage over late comers through the learning-by-doing process. *See id.* at 34-35. For example, being the first to sell a manufacturer's product in an area, a distributor could, based on the information not available to outsiders, develop a marketing technique that is best for the sale of the manufacturer's products and could thereby propose a better offer than its competitors at the next bidding.

55. *See id.* at 29-30, 35-37. *See also* Timothy J. Muris et al., *Strategy and Transaction Costs: The Organization of Distribution in the Carbonated Soft Drink Industry*, 1 J. ECON. MGMT. & STRATEGY 83 (1992)(showing, through empirical study, that the consideration to economize on transaction costs in response to the new competitive environment, not the desire to raise final prices, has influenced Coca-Cola and Pepsi-Cola to alter their distribution pattern from that of using independent bottlers to using its own subsidiaries).

56. *See, e.g.,* Irwin M. Stelzer & Richard Schmalensee, *Potential Costs and Benefits of Vertical Integration*, 52 ANTITRUST L.J. 249, 252 (1983)(using R & D as an example to demonstrate that vertical integration avoids the need to stipulate all contingencies and could protect the parties from being held up by their specific investments such as those of knowledge and other human capital devoted to the R & D).

However, it is worth noting here that some commentators have questioned the idea that vertical integration functions more efficiently than market transactions in reducing the costs of opportunism. For example, Coase argues that because firms will take into account the adverse impact on their future business opportunities by acting opportunistically, the marketplace should be sufficient to internalize those costs. *See* R.H. Coase, *The Nature of the Firm: Meaning*, 4 J.L. ECON. & ORG. 19, 30-31 (1988). *See also* John Stucky & David White, *When and When Not to Vertically Integrate*, 1993 SLOAN MGMT. REV. 71, 73-74 (1993)(suggesting that in addition to the consideration of costs inherent in the specific investments, frequency of transaction is also a decisive factor in inducing firms to adopt vertical integration). *But cf.* Klein, *supra* note 52, at 203 (criticizing Coase's arguments).

III. Vertical Control in Non-Competitive Markets: Vertical Integration and the Successive Monopoly Problem

In addition to the principal-agent dilemma and the transaction-cost considerations, firms face different types of concerns about using independent dealers as sales agents in non-competitive downstream markets. Perhaps the most frequently discussed issue in this regard is the "successive monopoly" problem.⁵⁷ Put briefly, the successive monopoly problem refers to a situation in which the final prices charged by a monopolist-distributor will be higher and, therefore, the final sales of the manufacturer's product lower than the levels under direct sales. The problem arises in principle from the profit-maximizing processes undertaken by the two monopolists. In this part, we explain this theory more formally and show how vertical integration can solve the problem and improve overall welfare.

A. Competitive Price and Output Level

To demonstrate how the successive monopoly problem arises, we begin first with an examination of the benchmark case; that is, when both the upstream and downstream markets are competitive. For expositional purposes, we assume the following:

1. There are two stages of production. The intermediate good I , produced at stage one by the upstream industry is used to produce the final good F , at stage two by the downstream industry;

2. The technology of production at stage two is "fixed-proportions," in the sense that production on one unit of F requires exactly one unit of I ;

3. Firms at both stages face constant marginal costs denoted by MC_i and MC_f respectively;

4. The downstream firms incur a constant cost of MC_t to transform I into F ;

5. Assume there is no transaction cost.

Figure 1 shows how the price of F is determined when both stages are in perfect competition.

57. The successive monopoly problem is discussed in most industrial organization textbooks. The original proof that vertical integration solves the problem was presented by Spengler, and the economic analysis we present in the text follows Spengler's. See Joseph J. Spengler, *Vertical Integration and Antitrust Policy*, 58 J. POL. ECON. 347, 350-51 (1950). For a review of the development of this theory, see generally Fritz Machlup & Martha Taber, *Bilateral Monopoly, Successive Monopoly, and Vertical Integration*, 27 ECONOMICA 101 (1960). See also CARLTON & PERLOFF, *supra* note 21, at 523-25; TIROLE, *supra* note 14, at 174-75; WILLIAM F. SHUGHART II, *THE ORGANIZATION OF INDUSTRY* 317-24 (1990). The analysis described here is based on Shughart's framework.

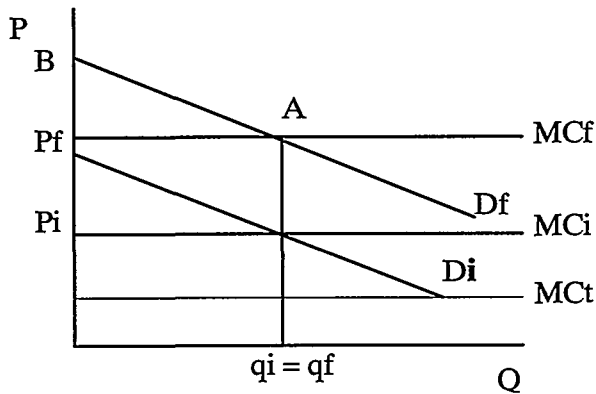


Figure 1

Suppose that D_f is the industry demand curve for F and D_i represents the *derived demand* for I .⁵⁸ Under the perfect-competition assumption, the upstream industry is expected to sell I at its marginal cost MC_i and produce at the level of q_i . Within the unintegrated transactions setting, P_i then becomes part of the production cost for the downstream industry, and the price for F is denoted in Figure 1 as P_f where it is equal to $P_i + MC_t$. The final demand for F thus becomes q_f . Under the technology of fixed-proportions production, q_f is equal to q_i .

The implication of this illustration is that it makes no difference whether a manufacturer uses independent distributors to sell its product or vertically integrates the downstream firms and sells the product itself when both market levels are competitive. Since both the upstream and downstream firms are selling at their marginal costs respectively, they make zero economic profits from the distribution arrangement. And given the competitive environment at the downstream market, a manufacturer is incapable of raising the price of F by vertically integrating one or some of the downstream firms. By selling I at MC_i to its downstream division, the highest price the integrated firm could charge the consumers will still be P_f . Accordingly, consumer welfare, in the form of consumer surplus, remains at the level of the triangular area BAP_f before and after

58. The term *derived demand* refers to a demand for a product "not for its own sake but in order to use it in the production of [other] goods or services." PARKIN, *supra* note 17, at 384. In the present case, the demand for I is not for the use of I itself but for the production of F , and therefore is *derived* from the demand of F . One of the determinants of the derived demand is that an increase in the price of F results in an increase in the price of I . See SHUGHART, *supra* note 57, at 318 n.14.

vertical integration.⁵⁹

B. Price and Output Levels Under a Monopolized Upstream Market

Next, let us relax the assumption of a competitive upstream market and continue to hold the others. Suppose that one firm monopolizes the supply of I . MR_i is introduced in Figure 2 to represent the marginal revenue curve facing the upstream monopolist. Following the profit-maximizing process, the input supplier is expected to produce I up to the point where its marginal cost is equal to its marginal revenue.⁶⁰ That is, the upstream monopolist will produce q_i units of I at the price of P_i for each unit sold. In turn, P_i becomes part of the downstream firm's marginal cost of producing F which is equal to P_i plus MC_t . With a competitive downstream market, it is also the price that F will be sold to the consumers. In sum, the final price and output levels under this circumstance can be represented by P_f and q_f in Figure 2.

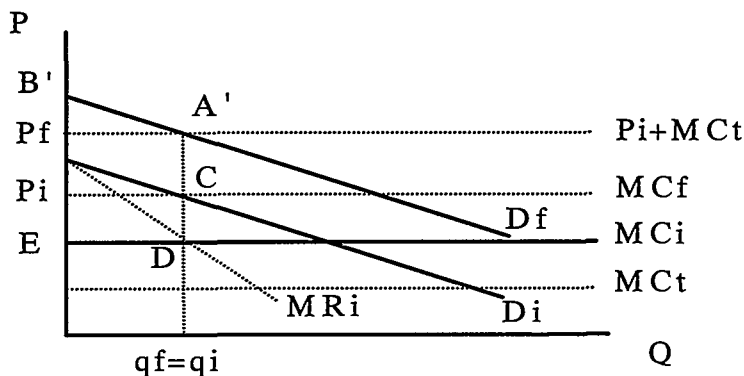


Figure 2

With a monopolized upstream market and a competitive downstream market, we again observe the same welfare implications with or without the employment of vertical integration by the

59. In economics, the concept of consumer surplus is defined as the difference between the highest price a consumer is *willing* to pay for a product (reservation price), which is represented by the higher point on the demand curve, and the price he or she *actually* pays for it times the quantities purchased. Adding up the surplus over the buyers in a market yields the total consumer welfare or aggregated benefits from the transaction. See PINDYCK & RUBINFELD, *supra* note 19, at 113-14. Therefore, consumer surplus in the present case is the triangular area surrounded by D_f , MC_f , and the vertical axis in Figure 1.

60. See *id.* at 322-24.

upstream firm. In the case of "separate sale," the downstream firms will continue to make zero economic profits. The upstream firm will capture a monopoly profit represented by the rectangular area of P_iCDE from selling I .⁶¹ Consumer welfare in the form of consumer surplus will be $B'A'P_f$.

Now, consider the case of integrated operation. Suppose the upstream monopolist acquires *one* of the downstream distributors. Instead of charging the profit-maximizing price P_i , the upstream monopolist sells I to its downstream division at its marginal cost, MC_i . Will the internal transfer of input increase the integrated firm's profits at the downstream market? The answer is no. Given the price of MC_i , the downstream division's production cost of F becomes MC_f which is equal to $MC_i + MC_t$. It is obvious from Figure 2 that MC_f is lower than the competitive price P_f , and, therefore, the downstream division could earn a profit of $(P_f - MC_f) \times q_f$.⁶² However, under the construction of Figure 2, this profit will be totally offset by the loss that the upstream monopolist has to incur by selling I at its marginal cost instead of at the monopoly price. By the same token, even if the manufacturer could in some way acquire a *substantial portion* or *all* of the downstream firms, the increase in the downstream profits is simply an internal transfer of the monopoly rents from the upstream firm.⁶³ Taking the integrated firm as a whole, the benefits that could be captured are unaffected by integration. Also, consumer welfare will remain at the level of $B'A'P_f$. The overall welfare effect after vertical integration is the same as that in the "separate sale" case. To be sure, in comparison with Figure 1, the amount of consumer surplus is decreased; but this is due to the exercise of the pricing power by the upstream monopolist, which in essence is a horizontal problem and usually cannot be cured by prohibiting vertical integration.

C. The Successive Monopoly Problem

Let us further relax the assumption of a competitive downstream market and examine the scenario when both the upstream and downstream markets are dominated by two separate monopolists. Respectively, let MR_i and MR_f in Figure 3 represent the curves of

61. A monopolist's profits from sales is the aggregation of the difference between the marginal cost of the product and the price *actually* paid by the consumer. In the present case, they are equal to $(P_i - MC_i) \times q_i$.

62. Insofar as the downstream market remains competitive, P_f will still be the optimal price for the integrated firm to sell F . It cannot take advantage of the cost savings from internal transfer of I to raise the final price above the level of P_f because it will lose all its sales. On the other hand, pricing F below P_f is irrational because given the large number of sellers in the market, the integrated firm cannot thereby increase the demand of F from itself.

63. See SHUGHART, *supra* note 57, at 319-20.

marginal revenue schedules facing the upstream and downstream firms. To simplify our analysis, assume that the upstream monopolist sells I to industries other than that of the downstream monopolist (thereby avoiding the need to consider the case of monopsony here).⁶⁴ (Figure 3 looks shockingly complex at first glance, but we must ask the reader to take a few seconds to stare at it. It is much simpler than it looks, and we will need to refer to it here and at one point later in the text.)

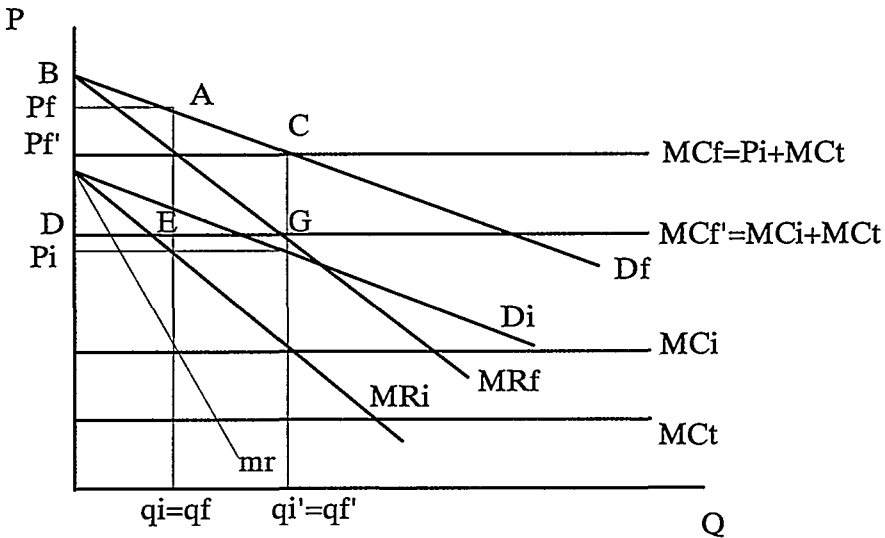


Figure 3

Consider first the case of unintegrated operation in which both firms independently make price and output decisions based on their own profit-maximizing objectives. The upstream monopolist starts by producing I at the level where MR_i is equal to MC_i , and would charge the downstream firm P_i for each unit of I . With that price, the downstream monopolist's marginal cost of production will be raised to MC_f and the final output of F will be reduced to the level q_f , where MC_f is equal to its marginal revenue MR_f at the price of P_f per unit. In effect, it is as if the *actual demand* for I is the curve MR_i in Figure

64. The relationship between vertical integration and monopsony will be discussed in Part IV.

³⁶⁵ and the upstream monopolist will begin producing by equating MC_i with the new corresponding marginal revenue curve mr . In sum, given that both the upstream and downstream markets are monopolized, profit-maximizing considerations by the two monopolists will lead to an increase in the final price and a reduction in the sale of the final products. Coincidentally, the derived demand for the input will also fall.

To gain an intuitive understanding of this phenomenon, consider the following example by Justice (then Judge) Breyer.⁶⁶ Suppose that the market of aluminum ingot is monopolized and it costs \$40 to produce. The transformation cost of fabricating ingot into aluminum sheet is presumed to be \$35. Assume the profit-maximizing price for the sheet is \$100. With a competitive market for aluminum sheet, the ingot monopolist will charge \$65 for the ingot to ensure that the final price will maintain at the profit-maximizing level. On the other hand, if the fabricating market is dominated by an independent dealer and the input is priced at \$65, the dealer will mark up the final price by more than \$35 because the dealer will want to maximize its profits as well. By so doing, however, the dealer increases the final price for the sheet and reduces the quantities demanded for both the sheet and the ingot.

The phenomenon described here is commonly referred to by economists as the *successive monopoly* problem. It is a "problem" for the upstream firm because the downstream firm's self-interest conflicts with its own. And similar to the resolution of the principal-agent dilemma, one of the solutions to this problem is to bypass the distributors and sell the product directly via vertical integration. Unlike what we have seen in Sections A and B, vertical integration employed in a successively monopolized environment has the effect of improving overall social welfare.

Consider first the change in profits to the monopolists. Before integration, the profits that could be made by the upstream and downstream firms from the sale of I and F respectively are $(P_i - MC_i) \times q_i$ and $(P_f - MC_f) \times q_f$. Together, both monopolists earn a profit of $(P_f - MC_f') \times q_f$ which is equal to the rectangular area of $P_f AED$ in Figure 3.⁶⁷ After integration, the integrated firm internally transfers I to its downstream division at its marginal cost and thereby induces the downstream division to produce F at the cost of MC_f' . With the

65. Since the optimal output level of F occurs at the point where $MR_f = MC_f = P_i + MC_t$, we can rewrite the equation as $P_i = MR_f - MC_t$. This could be viewed as the demand function for I .

66. See *Town of Concord, Mass. v. Boston Edison Co.*, 915 F.2d 17, 24 (1st Cir. 1990).

67. *Proof:* The total monopoly gains for the upstream and downstream firms are $(P_i - MC_i) \times q_f + (P_f - MC_f) \times q_f = [P_i - MC_i + P_f - (P_i + MC_t)] \times q_f = (P_f - MC_f') \times q_f$.

same profit-maximizing consideration as that of a monopolist, the optimal price-quantity combination for the integrated firm now becomes Pf' and qf' . Compare the changes in total profits that could be earned by the monopolists. The upstream division of the integrated firm now makes zero economic profit and will capture all the monopoly gains through its downstream division. With the same calculation conducted previously, the total profits for the integrated firm will be $(Pf' - MCf') \times qf'$ and can be represented by $Pf'CGD$. It is obvious from Figure 3 that $Pf'CGD$ is larger than $PfAED$. Furthermore, since the final price is reduced from Pf to Pf' after integration, consumer welfare in the form of consumer surplus is increased from $BAPf$ to $BCPf'$. In sum, vertical integration makes both the producers and consumers better off.

D. Vertical Integration Under the Technology of Variable-Proportions Production

The positive welfare implication associated with using vertical integration to solve the successive monopoly problem has led proponents of the Chicago School to promote the view that vertical integration should be subject to a less stringent legal standard.⁶⁸ But such a proposal is not without challenge. In particular, commentators have questioned the conclusion that vertical integration will be beneficial to the producers and consumers, noting that the positive welfare result may not hold when one relaxes the assumption of fixed proportions production.

In the real world, industries often use production technologies in which input combinations are variable. That is, different portions of inputs can be adjusted from time to time, based on fluctuations of their relative prices, to produce the same amount of output. For example, within a fixed expenditure constraint, a producer could switch between using either five units of labor and two units of capital, or four units of labor and three units of capital to manufacture a computer, depending on relative prices of capital and labor. In the extreme case where one of the inputs is monopolized and the other competitively supplied, the technology of variable-proportions production may lead the downstream firm to use too much of the competitive input and too little of the monopolized one.

The welfare effect of vertical integration under variable-proportions production is indeterminate. In general, it depends on the balance of two factors.⁶⁹ First, for the integrated firm, vertical

68. See, e.g., POSNER, *supra* note 3, at 196-201 (1976); BORK, *supra* note 3, at 226-30.

69. See generally HERBERT HOVENKAMP, *FEDERAL ANTITRUST POLICY*. See also Perry, *supra* note 13, at 191; BLAIR & KASERMAN, *supra* note 9, at 304; F.M. SCHERER & DAVID ROSS, *INDUSTRIAL MARKET STRUCTURE AND ECONOMIC PERFORMANCE* 523-

integration eliminates inefficient production at the downstream market. Input combinations can be restored to the level that minimizes production costs and, thereby, increases productive efficiency at the downstream market.⁷⁰ Second, after integration, the integrated firm could control all the downstream industry's inputs, not just the one input it makes. It could further lead the integrated firm to abuse its monopoly power to reduce output, increase the price for the final product, and decrease allocative efficiency.⁷¹ Contrary to the proposition advanced by the Chicago School, this analysis implies that antitrust enforcement agencies may need to make a tradeoff between the efficiency gains associated with the removal of distorted input combinations and the welfare loss attributable to the augmentation of monopoly power.⁷²

IV. Vertical Restraints as Substitutes for Vertical Integration in Competitive Markets

In some cases, vertical restraints offer the same virtues as those of vertical integration in alleviating the costs of separate distribution in competitive markets. In general, these costs are specific realizations of the principal-agent dilemma and the transaction costs inherent in vertical relationships. That is, in pursuance of their own interests, distributors create additional costs, or "externalities," that are borne by the manufacturer. And as Tirole explains, to "internalize" those externalities, competition among downstream dealers has to be reduced or eliminated by the manufacturer, through the imposition of various contractual restraints to give them a property right against any predation of their investments in promoting the manufacturer's product.⁷³

24 (Houghton Mifflin Co. 3d ed. 1990). See Part I of the appendix.

70. See generally HOVENKAMP, *supra* note 69.

71. See SCHERER & ROSS, *supra* note 69, at 523; HOVENKAMP, *supra* note 69 at 337-38.

72. Nevertheless, such a "trade off" model for the evaluation of the welfare effect of vertical integration with variable-proportions technology is criticized by Professor Shughart as misleading. In particular, the model is based on the assumption that vertical integration will in some ways change a competitive output market into a monopolistic market which will lead to an increase of the final price. On this assumption, efforts from antitrust enforcement agencies to balance various elements concerning efficiency is warranted. However, Shughart points out that it is not only unlikely but also rare that an input monopolist could use vertical integration to change a competitive downstream market. See SHUGART, *supra* note 57, at 324. Moreover, even if vertical integration does lead to price increases, the adverse effects are attributable to horizontal market power at the downstream market, and could be more appropriately attacked by horizontal than by vertical antitrust doctrines. See *id.*

73. See TIROLE, *supra* note 14, at 183.

A. Exclusive Dealing Arrangements and the Adverse-Selection Problem

Generally, an *exclusive dealing arrangement* can be broadly defined to include both *exclusive dealing contracts* and *requirement contracts*. A *requirement contract* is a commitment from a downstream firm to purchase a minimum amount or all of its required inputs from the upstream supplier. Its function in resolving the successive monopoly problem will be addressed later. On the other hand, an *exclusive dealing contract* refers to an agreement between the upstream and downstream firms in which one promises to deal only with the other. One of its advantages, equivalent with vertical integration, is that it facilitates the reduction of costs related to adverse selection.

When the quality of the manufacturer's product varies over units, it is foreseeable that this will induce the distributors to spend time and effort sorting out products with higher quality. In comparison with the case where all distributors have the same expectation regarding the average quality of the manufacturer's product, such inspection is usually costly and of no collective value.⁷⁴ This oversearching, and the associated costs, could be greatly reduced if both parties agreed to have the distributors pay an average price for the manufacturer's promise to provide the product of average quality, an arrangement that could be secured through a requirements contract.⁷⁵ However, such an arrangement might in turn induce the manufacturer to sort out the high-quality items and sell them to other dealers at higher prices. Exclusive dealing contracts constrain the manufacturer's freedom to transact elsewhere and, as a result, eliminate the adverse selection costs that are likely to be borne by the distributors.⁷⁶

B. Vertical Restraints and the Problems of Moral Hazard

As far as using vertical restraints to resolve the problem of moral hazard in vertical relationships, two specific cases are frequently

74. See EDWARD C. GALLICK, FTC STAFF REP., BUREAU COMPETITION, EXCLUSIVE DEALING AND VERTICAL INTEGRATION: THE EFFICIENCY OF CONTRACTS IN THE TUNA INDUSTRY 12 (1984).

75. In a similar vein, Kenney and Klein have argued that the consideration of reducing oversearching costs is the main reason why De Beers, the famous diamond supplier, sells its diamond by packages and on a take-it-or-leave-it basis. See Roy W. Kenney & Benjamin Klein, *The Economics of Block Booking*, 26 J.L. & ECON. 497, 502-05 (1983).

76. By selling only one manufacturer's product, however, the appropriable quasi-rents created by exclusive dealing could be significant. Therefore, it might induce the manufacturer to act opportunistically. Gallick suggests that this is a problem of tradeoff. If the net effect between the benefit of exclusive dealing and the cost it will impose is positive, exclusive dealing could still be an efficient arrangement. See GALLICK, *supra* note 74, at 12-13, 18.

discussed by commentators: the *free-riding problem* and the *dilution* of product reputation or brand image. In both cases, vertical restraints are instituted by the manufacturer with a view toward protecting investments made by the distributors in providing services or other promotional efforts and, thereby, maintaining the quality of the product and keeping its sales at a high level.

(1) *Vertical Restraints and the Free-Riding Problem*

In general, manufacturers favor more services or promotional efforts from their distributors because they will increase market demand for their products. To consumers, services and promotional efforts are indispensable in some circumstances. For example, a buyer may need a detailed explanation or presale testing before deciding to purchase an expensive hi-fi stereo.⁷⁷ To accommodate these needs, additional investments may have to be made by the distributors in hiring presale service personnel or installing show rooms. Furthermore, intensive advertisements or enthusiastic recommendations from the distributors may be crucial for successful marketing when the product is new to local markets or when the manufacturer is located far (geographically or even culturally) from the market that it wishes to penetrate. Equipped with their superior knowledge regarding local demand conditions and customers' preferences, the required promotional efforts could be more effectively provided by a local distributor than the manufacturer itself.

Nevertheless, in ensuring that the distributor will provide the required services or promotional efforts, a normal contract stipulating the distributors' responsibilities to perform those functions will be subject to the "free-riding" problem. Simply put, free-riding will occur when the good being supplied is a "public good."⁷⁸ Because the benefits generated from the supply of a public good are nonexcludable, people have an incentive to wait for others to incur the cost of providing the goods and then take a free ride on the benefits.⁷⁹ In the context of vertical distribution, the provision of

77. Presale services are especially needed when the good involved are durable goods or what Porter calls the "inconvenience goods." Buyers of inconvenience goods will rely more heavily on technical support, the ease of obtaining the required services, and are more willing to shop around before purchasing. See MICHAEL E. PORTER, *INTERBRAND CHOICE, STRATEGY, AND BILATERAL MARKET POWER* 106 (1976).

78. "A commodity is called a 'public good' if its consumption by any one person does not reduce the amount available for others." HIRSHLEIFER, *supra* note 23, at 478. For a more detailed discussion of the theory, see *id.* at 478-82. In the context of a retail market, information is often the best example of a public good. If one retailer advertises heavily, the information he feeds to potential buyers will benefit competing retailers.

79. See PINDYK & RUBINFELD, *supra* note 19, at 649.

presale services and promotional efforts create an important type of public good -- information. A distributor understands that the provision of information will increase sales of its product. But it also realizes that once the information is created by one of its rivals, it could easily take advantage of that information to increase its own sales. Therefore, it would be preferable to the distributor, thinking only of itself, not to provide the services. If such thinking is prevalent among the distributors, no services or promotional efforts will be provided.

Mathewson and Winter's externality hypothesis provides a theoretical framework for understanding the free-riding problem.⁸⁰ Under their model, a simple vertical contract creates three types of externalities.⁸¹ First of all, it creates a *vertical externality*. By vertical externality, we refer to the failure of a distributor to collect all of the increase in profits in connection with its sale efforts because part of the profits will go to the manufacturer in the form of wholesale margin. This, in turn, reduces the distributor's incentive to provide services or promotional efforts.⁸²

Consider a numerical example. Suppose there is one distributor, and the quantity sold depends on that distributor's effort in promoting the product. The retail price of the product is \$2, the wholesale price is \$1, the cost of the distributor's initial effort is \$.10, and the unit cost of the product is \$.20. If Q_0 is the quantity sold when the integrated manufacturer invests \$.10 worth of effort into promotions, then the profit for the integrated unit at the initial level of effort is $(\$2 - \$.10 - \$.20)(Q_0) = (\$1.70)(Q_0)$. Under the separate-sales agreement, the distributor's profit at the initial level of effort is $(\$2 - \$.10 - \$1)(Q_0) = (\$.90)(Q_0)$. Consider the effects of a change in effort levels from \$.10 per unit to \$.20 per unit with sales quantity rising from Q_0 to Q_1 . The integrated unit would see the value of the increase as $(\$1.60)(Q_1 - Q_0) - (\$.10)Q_0$, while the distributor alone would see the change as $(\$.80)(Q_1 - Q_0) - (\$.10)Q_0$. It is clear that the gain from the increase in effort is smaller for the distributor, and indeed may be negative for him, when it would be positive for the integrated manufacturer-distributor.

Second, a simple vertical contract creates a *horizontal externality* among distributors. Recall the previous stereo example. If discounts are allowed, a customer could simply go to a store that provides the pre-sale services, obtain the information regarding the product from

80. See generally G.F. Mathewson & R.A. Winter, *The Economics of Vertical Restraints in Distribution*, in NEW DEVELOPMENTS IN THE ANALYSIS OF MARKET STRUCTURE 211-36 (Joseph E. Stiglitz & G. Frank Mathewson eds., 1986).

81. See *id.* at 220-221.

82. See *id.* at 221.

the distributor, and then walk a few blocks away to purchase the same product from a discounter. Furthermore, the horizontal externality could also occur at the upstream market. To avoid the free-riding problem, a manufacturer might wish to provide the required promotional efforts for all its distributors. For example, it might adopt a nation-wide advertising program to promote its product. But if the distributors are not prohibited from selling other competing brands, those efforts could be used by distributors as "bait" to lure customers to visit their stores and then recommend the customers purchase the brands that do not have such services but are sold at lower prices. To the distributors, this "bait-and-switch" strategy could provide higher profits than when only the manufacturer's product is sold.⁸³

The third externality is the *informational externality*. This refers to the spillover effect that one distributor provides to adjacent rivals through its own efforts of providing product information. For example, a customer could be informed of a new product from an advertisement provided by one distributor, but purchase the product at another store. Through the information it provided, the distributor increases the demand for the new product from its competitors.⁸⁴

Vertical restraints could serve to internalize these externalities. To encourage the downstream firms to provide the required services, the manufacturer could create a minimum amount of profits for the distributors by imposing a resale price floor.⁸⁵ The vertical externality could thereby be attenuated. Moreover, to the extent that the resale prices of the same brand are capped from below across the market, the incentive to free ride on other distributors' services by price discounts is diminished. Thus, the horizontal externality could also be held in check.⁸⁶

Another mechanism that could be used to control the downstream free-riding behavior is a territorial restriction arrangement. By assigning an exclusive sale territory to a specific distributor and prohibiting its competitors from selling into that area, the manufacturer could effectively deter other distributors from soliciting customers across their borders and "skimming the cream" created by its neighbors' own sale efforts.⁸⁷ In addition, there are two

83. See Howard P. Marvel, *Exclusive Dealing*, 25 J.L. & ECON. 1, 7 (1982); Victor P. Goldberg, *The Free Rider Problem, Imperfect Pricing, and the Economics of Retailing Services*, 79 NW. U. L. REV. 736, 745 (1984).

84. See Mathewson & Winter, *supra* note 80, at 220.

85. See Telser, *supra* note 3, at 91.

86. See *id.*

87. On "cream-skinning" generally, see William A. Brock & David S. Evans, *Creamskimming*, in *BREAKING UP BELL: ESSAYS ON INDUSTRIAL ORGANIZATION AND REGULATION* 61-94 (David S. Evans ed., 1983). The "cream-skinning" entrant is able to

measures to resolve the bait-and-switch dilemma. An upstream firm, a franchisor for example, could charge its franchisees royalties based on their *total* sales revenue or units sold across all brands to extract part of the franchisees' profits from selling other brands that is attributable to the franchisor's promotional efforts.⁸⁸ The other way to avoid upstream free riding is to adopt an exclusive dealing arrangement. An exclusive dealing contract forecloses opportunities for the distributor to sell other brands and preserves the distributors' undivided loyalty to promote only the product produced by the manufacturer. It thus ensures that the profits generated from the promotional efforts will be captured by the manufacturer.⁸⁹

(2) *Vertical Restraints and the Maintenance of Product Reputation or Brand Image.*

Arguments justifying the use of vertical restraints to maintain product reputation or brand image are closely related to the analysis of the free-riding problem. When shirking on product quality occurs, the adverse impact it creates often permeates the whole distribution network and harms the manufacturer's reputation. For example, McDonald's reputation in the fast-food industry could be tarnished if one of its franchisees located near the exit of an interstate highway shirked on the quality of its hamburgers. A customer who bought a rotten hamburger from the franchisee may never return. However, the impression of "McDonald's lousy hamburgers" could spill over to franchisees at other locations and reduce the demand for their hamburgers.⁹⁰ As the Fifth Circuit in *Kentucky Fried Chicken Co. v.*

offer a lower price than its competitor, but only because it has managed to avoid some cost that is associated with selling in a given market. Thus, the cream-skimming entrant is not a cheaper producer—it competes only because it is able to take advantage of certain infrastructure expenditures borne by its rivals. The classic example involved the long-distance carriers who were able to compete against AT&T largely because AT&T was forced by regulators to keep its long-distance charges above the competitive level. See, e.g., William J. Baumol & Janusz A. Ordover, *Use of Antitrust to Subvert Competition*, 28 J.L. & Econ. 247, 257 (1985).

88. See Katz, *supra* note 10, at 655, 717.

89. See Marvel, *supra* note 83, at 6-7; Goldberg, *supra* note 83, at 745 (quoting Marvel's view); Richard M. Steuer, *Exclusive Dealing in Distribution*, 69 CORNELL L. REV. 101, 130 (1983). According to Katz, however, the anticipated benefits from using exclusive dealing contract under this circumstance must be offset by the increase in the distribution costs. Two types of distribution costs are specifically mentioned by Katz. First, by requiring distributors to sell only one single brand, it could lead to an increase of distribution costs associated with the lost economies of scale in downsize selling. Second, consumers may therefore choose to go to multi-brand stores first in order to minimize their search costs. The distributors could suffer diminishing sales. See Katz, *supra* note 10, at 697.

90. Paul H. Rubin, *The Theory of the Firm and the Structure of the Franchise Contract*, 21 J.L. & ECON. 223, 228 (1978). Shirking or cheating on product quality could occur in a

*Diversified Packaging Co.*⁹¹ noted: "A customer dissatisfied with one Kentucky Fried outlet is unlikely to limit his or her adverse reaction to the particular outlet; instead, the adverse reaction will likely be directed to all Kentucky Fried stores."⁹²

In addition, the product reputation justification is similar to the resolution of the bait-and-switch dilemma described above. Just as with the downstream distributors in bait-and-switch cases, upstream suppliers or manufacturers have an incentive to free ride on their competitor's brand name. For example, they could request their distributors to place their products adjacent to more famous brands. Customers who are attracted by those brands, but are more price conscious, might therefore switch to the low-price products that are closest to the prestigious brand on the shelf.

To avoid free riding on product quality, firms need mechanisms to reduce or to eliminate the possibility of substitution between high-quality and low-quality products. Viewed in this light, the desire of firms to use tying or exclusive dealing arrangements seems obvious. Tying is a common arrangement in franchising where a franchisor bundles the licensing of its trademark (the "tying" product) with the purchase of inputs (the "tied" product) from the franchisor itself or from a supplier designated by the franchisor.⁹³ The purpose of this arrangement is to ensure that by using inputs of the same quality, a franchisee's incentive to free ride on the efforts by other franchisees could be eliminated.⁹⁴ Moreover, under the tying arrangement, when the shipment of inputs from sources other than the franchisor or its

"reverse" fashion. Goldberg has noted that a manufacturer might cheat on its product quality as well if the distributor is a famous outlet in the market and carries more than the manufacturer's brand. Customers may rely on the distributor's reputation to purchase a brand that is unknown to them but is sold at the distributor's store. The manufacturer may take advantage of this opportunity to promote their low-quality products and dilute the distributor's reputation. See Goldberg, *supra* note 83, at 748.

This problem explains why it is more common to observe franchisors using both integrated operation and franchising in practice. In markets where customers seldom make repeated purchases, the franchisor might prefer company-owned outlets to franchising because of quality concerns. Therefore, the co-existence of integrated operation and franchising could be viewed as a strategic response to the divergence in demand conditions. Cf. Brickley & Dark, *supra* note 45, at 418 (stating that the higher costs of monitoring employees at outlets near the interstate highways may lead the franchisor to use franchising system).

91. 549 F.2d 368 (5th Cir. 1977).

92. *Id.* at 380.

93. See J. Thomas McCarthy, *Trademark Franchising and Antitrust: The Trouble with Tie-ins*, 58 CAL. L. REV. 1085, 1110 (1970).

94. See Benjamin Klein & Lester F. Saft, *The Law and Economics of Franchise Tying Contracts*, 28 J.L. & ECON. 345, 349-51 (1985); Scott Makar, *In Defense of Franchisors: The Law and Economics of Franchise Quality Assurance Mechanisms*, 33 VILL. L. REV. 721, 729-31 (1988).

designated supplier is observed, the franchisor could reasonably suspect that shirking might be occurring, and thus tying makes the detection of shirking easier.⁹⁵ Alternatively, to recoup investments in creating a brand image, firms with more famous brand names could enter into exclusive dealing contracts with distributors, prohibiting them from selling other brands. This denies competitors the opportunity to capitalize on the investments of the famous brand holder.

Aside from product reputation, brand image could also be subject to free-riding behavior. In a general sense, a high price signals high quality and psychologically creates the image of superiority. As Richard Caves has described, by charging a higher price, the manufacturer "provides a signal of high quality . . . by assuring that the product is distributed only by [the distributors] whose costly premises signal a quality image."⁹⁶ Some customers may be willing to pay a high price for a product in exchange for its presumed higher quality or for its association with a more prestigious social status. Price cutting by a distributor may increase its own sales; but it could also dilute the value of brand image and reduce the probability of repeat purchases.⁹⁷ On balance, this could reduce the total profits of the whole chain. By discouraging price competition, minimum resale price maintenance can prevent the high-quality brand image from being dissipated.

C. Vertical Restraints and the Mitigation of Transaction Costs

(1) *Vertical Restraints and the Reduction of the Costs of Finding Qualified Contracting Partners*

Instead of spending time and effort investigating the hidden

95. See Klein & Saft, *supra* note 94, at 353. Tying itself may not totally eliminate the problem of detecting shirking if the franchisees are allowed to use a non-franchise input in conjunction with a non-franchise product. Other types of restraints such as exclusive dealing are required to supplement tying to control product reputation. The combined use of tying and exclusive dealing to ensure product quality may be understood under the "optimal franchising theory" described in Part VI discussing optimal franchising theory. And the extent to which the policing costs will be reduced by tying depends on how much effort will need to be exercised by the franchisor itself after tying. See Makar, *supra* note 94, at 739.

96. Richard E. Caves, *Vertical Restraints in Manufacturer-Distributor Relations: Incidences and Economic Effects*, in ANTITRUST AND REGULATION 29, 40 (Ronald E. Grieson ed., 1986).

97. This phenomenon is described by economists as the "snob effect," which refers to the desire to own unique products. A consumer's demand for a snob good will be higher when fewer people own it. See PINDYCK & RUBINFELD, *supra* note 19, at 120; Harvey Leibenstein, *Bandwagon, Snob, and Veblen Effects in the Theory of Consumers' Demand*, 64 Q.J. Econ. 183, 199 (1950) (presenting a theoretical analysis of the snob effect).

characteristics of distributors and to find qualified contracting partners, an upstream firm may induce the downstream dealers to voluntarily abide by their contractual obligations and eliminate their incentive to shirk through appropriately designed contractual arrangements that will render default unprofitable for the dealers.⁹⁸ The franchise fee and minimum resale price maintenance arrangements offer illustrations.

Consider the franchise fee first. Unlike charging distributors a uniform price for each additional unit they purchase, a two-part tariff⁹⁹ pricing scheme is the most unique feature that distinguishes franchising from other distribution methods. This refers to a pricing arrangement that requires a franchisee to pay an up front lump-sum fee (the first part of the tariff) to secure the right to purchase the franchisor's tangible goods at an agreed upon unit price for each purchase (the second part of the tariff). In situations where informational asymmetry exists and the resources committed to monitor the franchisees' performance of their contractual obligations are significant, a franchisor could require the franchisees to pay a higher franchise fee to compensate for the reduction in the value of its product should reneging occur. The same rationale applies to the situations where a famous franchisee requests a lower franchise fee to offset the losses arising from the franchisor's chiseling on its duty to supply a high-quality product. In both, the franchise fee serves a similar function; as a "bond" used by the franchisor or franchisee to force the other party to live up to their contractual obligations. Such a bonding mechanism reduces the need to carefully screen contracting partners.¹⁰⁰

Minimum resale price maintenance could be used to reduce the costs of finding contracting partners, in a similar fashion. Insofar as the profits from reneging outweigh the capitalized value of those that

98. In a broader sense, this is the "self-enforcing contract" theory proposed by Professor Klein. See Benjamin Klein & Keith B. Leffler, *The Role of Market Forces in Assuring Contractual Performance*, 89 J. POL. ECON. 615, 635-37 (1981); Benjamin Klein & Kevin M. Murphy, *Vertical Restraints as Contract Enforcement Mechanisms*, 31 J.L. & ECON. 265, 270-76 (1988); Benjamin Klein, *The Economics of Franchise Contracts*, 2 J. CORP. FIN. 9, 17-22 (1995); Klein, *supra* note 46, at 449-50.

99. "Two-part tariff," "quantity-dependent pricing," and "nonlinear pricing" have been used interchangeably by economists to describe this pricing pattern. See TIROLE, *supra* note 14, at 143; CARLTON & PERLOFF, *supra* note 21, at 462; Hal R. Varian, *Price Discrimination*, in 1 HANDBOOK OF INDUSTRIAL ORGANIZATION 597, 604 (Richard Schmalensee & Robert D. Willig eds., 1989); LOUIS PHILIPS, *THE ECONOMICS OF PRICE DISCRIMINATION* 166 (1983).

100. See Benjamin Klein, *Transaction Cost Determinants of "Unfair" Contractual Arrangements*, 70 AM. ECON. REV. 356, 359 (1980); Kabir C. Sen, *The Use of Initial Fees and Royalties in Business-Format Franchising*, 14 MANAGERIAL & DECISION ECON. 175, 177 (1993).

could be earned had the contract been duly performed plus the cost of penalty, the incentive to renege will still remain even when the obligation to provide the required services and to maintain product quality has been unequivocally stipulated in the contract. Two contractual arrangements might be necessary for the contract to be self-enforcing. First, the manufacturer could enlarge the future stream of the premium for honoring the contract to the extent that it will be greater than the benefits from reneging.¹⁰¹ Second, the manufacturer must be able to credibly terminate the distributorship should reneging occur. Replacing the punishments from the courts, the right of termination will impose immediate economic losses on the distributors and make reneging unprofitable. And according to Benjamin Klein, the threat of being terminated will provide sufficient stimulus for the distributors to perform the contract.¹⁰² Minimum resale price maintenance stifles retail price competition, making it less attractive to distributors to shirk.¹⁰³

On the other hand, when the gains from reneging are greater, the manufacturer may have to lower its wholesale price to increase the future profits for its distributors. Without fixing a price floor, however, this profit margin may not be dissipated by a service provision but could be used by the distributors to lower their retail prices to attract more customers from one another. Minimum resale price maintenance under this circumstance is the means for the manufacturer to protect the future premium it creates for the distributors.¹⁰⁴ And again, once the price floor is properly set, the need to probe into the characteristics of distributors is greatly reduced.

(2) *Vertical Restraints and the Reduction of the Costs of Reduced Flexibility*

Unlike vertical integration which provides a manufacturer the opportunity to obtain more information regarding market demand, tying arrangements or requirement contracts mandate a sharing between the manufacturer and its distributors of the losses resulting from incidences unforeseeable before the contract is formed, and could thereby alleviate the costs of "reduced flexibility."¹⁰⁵ Consider the case where a manufacturer producing two products intends to license the whole production line to distributors. Assume the wholesale prices are set before the demand or popularity of the products is known and cannot be perfectly adjusted according to the

101. See Klein & Leffler, *supra* note 98, at 617.

102. See Klein, *supra* note 98, at 18.

103. See Klein & Murphy, *supra* note 98, at 276.

104. See *id.*

105. On "reduced flexibility," see *infra* text accompanying notes 41-45.

change in market conditions for each product.

Usually, investments have to be made for each product, to start producing, before demand conditions are known to the manufacturer. But once the demand for each product becomes observable, the downstream firms will be inclined to order more of the popular product and less of the unpopular one. Meeting this demand may not be in the manufacturer's best interest. Because of the difficulty of transforming the production technology of the unpopular product into that of the popular one, the surge in the demand for the popular product may impose two types of costs on the manufacturer. First, the demand for the unpopular good might decrease to the extent that it fails to cover its allocation of fixed costs. Second, when the capacities of both products are constrained, the increase in the demand for the popular product may eventually increase its marginal cost to a point where marginal cost begins to exceed marginal revenue. The manufacturer may be better off if it continues to produce the unpopular product to share the setup costs during the transitory period.¹⁰⁶ Tying the purchase of the unpopular with the sale of the popular product, or a requirement contract mandating the purchase of the specific amount of the unpopular product, could be helpful for the manufacturer to internalize these costs.

(3) *Vertical Restraints and the Reduction of the Costs of Opportunism*

Under certain conditions, tying and exclusive dealing offer the parties a chance to deter hold-up behavior and thereby reduce the costs of opportunism. Consider the case where a franchisor agrees in a long-term contract to supply its franchisees' inputs at a constant price. In the absence of arrangements that could effectively prohibit the franchisees from defaulting on their purchase obligations, the franchisees' incentive to hold up the franchisor during the periods of lower spot prices by threatening to switch to a rival is apparent. In a similar vein, when an unexpected surge in market demand results in an inventory shortage, the manufacturer may threaten to renege on its sale obligation and to sell to other distributors that are willing to pay higher prices for the product. The existence of quasi-rents created by specific investments may cause either the upstream or downstream firm to accept less favorable contractual terms.

Tying the licensing of the franchisor's trademark with the purchase of the input offers a solution to downstream hold-up. By bundling both, it is as if the franchisor has demanded that franchisees credibly commit to a specific asset, and uses that as a "hostage" to

106. See Goldberg, *supra* note 83, at 754-55.

guard itself against any future hostile actions by the franchisees.¹⁰⁷

Exclusive dealing arrangements, together with the penalty of violation, foreclose the opportunity for a distributor to purchase from another manufacturer, or for a manufacturer to sell to other distributors, and thereby raise the costs for a party who is considering acting opportunistically. To a certain extent, such arrangements safeguard either the upstream or the downstream firm from being exploited by the other when market conditions change.¹⁰⁸

V. Vertical Restraints as Substitutes for Vertical Integration in Non-Competitive Markets

As we have seen, the main functions of vertical integration in non-competitive markets are to solve the successive monopoly problem and to avoid inefficient substitution of inputs. The discussion in this part follows this dichotomy. We intend to show how various types of vertical restraints could also be used to achieve the same goals.

A. Franchise Fee and Royalties

One of the alternatives for an upstream monopolist to solve the successive monopoly problem is to use the franchise fee to extract downstream monopoly profits and then sell its product at the marginal cost to induce the downstream firm to produce or sell at the optimal level.¹⁰⁹ In Figure 3, for example, the upstream monopolist could require its franchisee to pay a lump-sum franchise fee in the amount of $Pf'CGD$, and then sell I at MCi .¹¹⁰ To the manufacturer, the franchise fee enables it to capture the monopoly profits that could be owned under vertical integration. And since the franchise fee is a fixed cost to the franchisee which is not variant with final output level, the downstream firm will still base its output decision on its marginal cost, MCi . As a result of this arrangement, the final price and output level will remain at Pf' and qf' . Furthermore, the franchise fee arrangement makes the downstream firm a "residual

107. See Williamson, *Credible Commitments: Using Hostage to Support Exchange*, 73 AM. ECON. REV. 519 (1983); Katz, *supra* note 10, at 699.

108. See Klein, *supra* note 46, at 445-46; GREGG FRASCO, EXCLUSIVE DEALING 7-8 (1991).

109. The seminal work on this topic is Walter Y. Oi, *A Disneyland Dilemma: Two-Part Tariffs for a Mickey Mouse Monopoly*, 85 Q.J. ECON. 77 (1971).

110. This arrangement permits the manufacturer to garner the entire monopoly profit, while allowing the retailer to earn a competitive return. In other words, the retailer does not really go without a profit altogether. The retailer earns just enough to cover the costs of capital and labor, and the opportunity costs of other resources used in production (e.g., managerial skill).

claimant" in the sense that it will be able to appropriate any increase in the joint profits in connection with its own promotional efforts.¹¹¹

Obviously, it makes no difference in terms of their effects in solving the successive monopoly problem and their welfare implications whether the upstream firm uses vertical integration or franchise fee arrangement. But the franchisee fee arrangement offers an additional advantage, the saving of transaction costs. The franchise fee arrangement is likely to be cheaper than vertically integrating.¹¹²

Alternatively, an upstream firm could sell its input to its distributors at marginal cost, and require them to pay royalties proportional to the total sales of the final product. By similar analogy, marginal pricing induces the distributors to sell at the optimal level and the royalty payment transfers the monopoly profits at the downstream market to the manufacturer as if the product had been distributed directly by the manufacturer.¹¹³

The royalty solution may seem counterintuitive initially because it has the same effect as taxing the franchisee. Standard economic analysis indicates that a revenue tax will reduce the quantity produced by a firm. This is indeed what happens to the franchisees that are required to pay a royalty to the franchisor; the royalty reduces their demand for the input supplied by the franchisor. However, this demand reduction is precisely what the franchisor intends, for his goal is to replicate the output-price combination that would be achieved by an integrated monopolist.

B. Maximum Resale Price Maintenance and Territorial Restriction

For resale price maintenance to function as the equivalent of vertical integration in solving the successive monopoly problem, the manufacturer could sell the input according to its own profit-maximizing price but set a maximum resale price requiring its distributors not to sell the final product at or below a certain level. Using Figure 3 as an illustration, the manufacturer could sell I at the price of P_i per unit, and then set the maximum price level at P_f' . P_i enables the manufacturer to capture the profits at the upstream

111. See TIROLE, *supra* note 14, at 176; Katz, *supra* note 10, at 664.

112. However, one caveat is necessary for this conclusion. It is based on the assumption that the downstream firm's profits are observable, and therefore, extractable, by the upstream firm. Once the final market demand is uncertain, two-part tariff arrangements may impose too much risk on the downstream firm and the upstream firm may have to adjust between the unit prices for the product and the amount of franchise fee to achieve the same outcome. See TIROLE, *supra* note 14, at 176-77; G.F. Mathewson & R.A. Winter, *An Economic Theory of Vertical Restraints*, 15 RAND J. ECON. 27, 35 (1984).

113. See Frederick S. Inaba, *Franchising: Monopoly by Contract*, 47 S. ECON. J. 65, 71 (1980).

market through monopoly pricing. On the other hand, Pf' prevents the downstream firm from raising the final price above its marginal cost, and reduces the downstream firm's derived demand for I as a result. It is as if the manufacturer were facing a competitive rather than monopolistic downstream market.¹¹⁴

The use of territorial restrictions to avoid the successive monopoly problem seems to be counter-intuitive. By assigning a sales territory exclusively to a distributor, a territorial restriction creates a monopolist at the downstream market. How, then, can territorial protection achieve the same result as vertical integration? The answer lies primarily in the bidding process for the exclusive territorial right. Under the exclusive distribution relationship, candidates will bid for the exclusive monopoly. The winning bid will equal the expected monopoly profits that could be gained from the chain, assuming the input is transferred at marginal cost from the manufacturer. Again, the manufacturer is able to capture the whole monopoly profits through the bidding process and marginal-cost pricing by the manufacturer prevents the final output from deviating from the optimal level.¹¹⁵

C. Tying and Output Royalty

In contrast with the solution of the successive monopoly problem, the equivalence between tying and vertical integration, in non-competitive markets, can be seen in the ability of both to correct inefficient input substitution under the technology of variable-proportions production. Similar to vertical integration, under which the input monopolist could internally adjust to the most efficient utilization of inputs A and B , the input monopolist could also manipulate the relative price of A with respect to B through tying, and thereby force the downstream firms to purchase more of A and less of B . The input combination could therefore be brought back to

114. See TIROLE, *supra* note 14, at 177; CARLTON & PERLOFF, *supra* note 21, at 526. Due to the uncertainty in market demand and the difficulty in observing the differences in consumers' preferences, some commentators have argued that the effectiveness of using maximum resale price maintenance to solve the successive monopoly problem may be limited. See, e.g., Richard L. Smith II, *Franchise Regulation: An Economic Analysis of State Restrictions on Automobile Distribution*, 25 J.L. & ECON. 125, 128 (1982); Katz, *supra* note 10, at 676.

115. Since the more capable a distributor is, the more willing it is to outbid other candidates, the bidding process could also increase the probability of selecting qualified contracting partners. In addition, once the exclusive territory is authorized, it becomes a valuable investment to the distributor. It may be more willing to cooperate with the manufacturer regarding any change in the method of operation. See SCHERER & ROSS, *supra* note 69, at 558.

the efficient ratio.¹¹⁶

One might ask why an input monopolist would be concerned at all over the input combination adopted by downstream firms. In general, the input monopolist would not be concerned. But the important feature of the tying arrangement is that it can be used by the input monopolist to achieve the same result as forward vertical integration, provided the input monopolist is able to tie his monopolized good to all other input substitutes used by downstream firms.

Analogously, an output royalty arrangement under which a distributor is required to pay the manufacturer a proportion of the total quantities sold in addition to the unit price can also be used to avoid inefficient input substitution. One might view the output royalty as a tax, t , on each unit of the final good sold by the distributor. The supplier of input A could sell A at its marginal cost to induce the downstream firms to use the same input combination as that under vertical integration.¹¹⁷ At the same time, it could set t equal to the difference between the monopoly price and the marginal cost of A to extract the downstream profits. The output royalty arrangement does not subject the manufacturer to diminishment of its appropriable rents from sale to the downstream market. The incentive on the distributors' side to maintain the optimal output level could also be preserved.

D. Policy Implications

The substantive equivalence of vertical integration and many vertical restraints provides support for a uniform legal standard for the vertical restraints examined in this paper. Under American antitrust law, minimum resale price maintenance and tying arrangements fall under the scrutiny of the more stringent per se rule.¹¹⁸ On the other hand, territorial restrictions and exclusive dealing are reviewed under the rule of reason. But in view of their ability to substitute for vertical integration in solving the principal-agent dilemma or the successive monopoly problem, such discriminatory treatment appears to be unwarranted. In particular, it will become increasingly difficult for the Court to justify the claim

116. For a formal proof of the equivalent effect between tying and vertical integration, see Roger Blair & David Kaserman, *Vertical Integration, Tying, and Antitrust Policy*, 68 AM. ECON. REV. 397, 397-99 (1978).

117. See ROGER D. BLAIR & DAVID L. KASERMAN, LAW AND ECONOMICS OF VERTICAL INTEGRATION AND CONTROL 58-63 (1983).

118. See *Dr. Miles Med. Co. v. John D. Park & Sons Co.*, 220 U.S. 373, 408-09 (1911)(discussing minimum resale price maintenance); *International Salt Co. v. United States*, 332 U.S. 392, 396 (1947)(discussing tying).

that resale price maintenance and tying arrangements seldom prove beneficial for market competition when their substantive equivalence to other vertical control mechanisms becomes more widely accepted.

The recent Supreme Court decision in *State Oil Co. v. Khan*,¹¹⁹ overruling the application of the per se rule to maximum resale price maintenance¹²⁰ seems to accept this reasoning. In *State Oil*, a case concerning a resale price ceiling agreement between an oil company and gas stations, the Court said that it is difficult to maintain that maximum resale price maintenance could harm competition to the degree that would justify the application of the per se rule.¹²¹ The Court also endorsed Judge Posner's view that maximum resale price maintenance is a means to prevent downstream monopolists from exploiting their positions.¹²² In doing so, the upstream supplier is not acting out of malice but in its own commercial interest because "[t]he higher the price at which the gasoline is resold, the smaller the volume sold, and so the lower the profit to the supplier if the higher profit per gallon at the higher price is being snared by the dealer."¹²³

VI. The Complementary Nature of Vertical Control—Toward a Theory of Optimal Franchising¹²⁴

A. Partial Integration

The analysis so far has focused on how vertical integration mitigates costs associated with the principal-agent dilemma, or solves the successive monopoly problem, and how various types of vertical restraints independently could serve the same function. The analysis has focused on theory. *In practice*, we are likely to encounter hybrid combinations of vertical control methods, especially in the business of franchising. For example, McDonald's franchises most of its outlets but still maintains twenty-eight percent of all its retailers as company-owned outlets.¹²⁵ Why is this necessary if vertical integration and vertical restraints are equivalent in many respects?

We think the answer rests in large part in the firm's desire to

119. 522 U.S. 3 (1997).

120. See *Albrecht v. Herald Co.*, 390 U.S. 145, 152-54 (1968) (establishing the application of the per se rule to maximum resale price maintenance).

121. 522 U.S. at 15.

122. See *id.* at 15-16.

123. *Id.* (quoting *Khan v. State Oil Co.*, 93 F.3d 1358, 1362 (7th Cir. 1996)).

124. "Optimal" refers to the minimally sufficient combinations of various types of vertical control to maximize the vertically integrated profits. TIROLE, *supra* note 14, at 173; Mathewson & Winter, *supra* note 112, at 33.

125. Severin Borenstein & Richard Gilbert, *Uncle Sam at the Gas Pump*, 2 REGULATION 63, 66 (1993).

overcome the problems of informational asymmetry and market uncertainty.¹²⁶ The validity of the claim that vertical restraints can serve as a perfect substitute for vertical integration depends on the availability of information concerning the demand schedule at the downstream market. For example, the effective employment of the "franchise fee plus marginal unit price" arrangement¹²⁷ requires that the franchisor be able to observe the demand curve of the final good and, consequently, the downstream monopoly profits. This is also true of claims concerning the mutual benefits from adoption of maximum resale price maintenance or requirement contracts. Lacking the knowledge about consumers' reservation prices for the final goods, it is unlikely that the upstream firm will figure out the optimal price and output levels for the downstream firm, let alone the minimum quantities of the inputs that the downstream firm is required to purchase or the maximum price that it is allowed to charge its customers. Given ever-changing market conditions, it is difficult, perhaps impossible, for the firm to acquire this information.

In addition, market uncertainty imposes other costs on the manufacturer who integrates downward or the distributor who integrates upward. As Dennis Carlton has argued, the competitive market assumption that firms can perfectly respond to any price change and immediately adjust their output levels is usually unattainable.¹²⁸ Decisions regarding production capacity often have to be made before crucial market information can be verified. Once this information is verified, capacity adjustment takes time. Full vertical integration increases the risk that there will be more unsold goods if market demand unexpectedly plummets. The use of partial integration by the upstream firm is tantamount to treating the downstream market as an insurance market for the costs of overproduction.¹²⁹ That is, the manufacturer will continue to produce by itself but simultaneously utilize its distributors to supply the extra demand or to bear the risk of fluctuations.¹³⁰

126. Aside from this consideration, partial integration offers other benefits for managerial purpose. For example, a franchisor could use market transaction as a threat to motivate more efficient internal production. See BESANKO ET AL., *supra* note 34, at 156. Conversely, internal production provides solid protection for the upstream firm from any hold-up by the independent suppliers. See *id.* However, it might create some disadvantages as well. In particular, shared production between internal and external units might cause coordination and monitoring costs to increase, which could result in less efficient production. See *id.*

127. See text accompanying notes 99-101.

128. See Dennis Carlton, *Vertical Integration in Competitive Market Under Uncertainty*, 27 J. IND. ECON. 189 (1979).

129. See *id.* at 199.

130. See *id.* See also MICHAEL E. PORTER, *COMPETITIVE STRATEGY* 319-20 (1980). Porter further points out that partial integration also reduces the costs arising from a lock-

B. The Optimal Combinations of Vertical Restraints

Apart from partial integration, combinations of two or more types of contractual restraints are more frequently observed distribution patterns than the single use of any particular restraint. However, the economic function of many contractual vertical restraints still remains a new subject of inquiry for economists. Moreover, some of the recent studies in this area appear to be limited by their assumptions. The two studies introduced below are intended only to provide a snapshot of the recent developments in this area.

Blair and Kaserman focus on how market uncertainty that causes the franchisor and franchisees to place different values on the franchise will induce the franchisor to use a "package" of contractual restraints.¹³¹ The essence of their argument is that since market demand is subject to fluctuation, both the franchisor and franchisees may apply different discount rates to calculate the franchise's net present value.¹³² This divergence in franchise value will be enlarged if the quality of the franchisor's effort *ex post* to maintain the value of the franchise (the franchisor's trademark, trade name, for example) is taken into consideration *ex ante* by the franchisee. For example, if the franchisee has little confidence in the franchisor's determination to prevent free-riding by other franchisees, it would apply a much higher discount rate than would be applied by the franchisor (if the franchisor thinks otherwise). Constrained by bounded rationality from writing a complete contract regarding the degree of effort that should be exercised by the franchisor to avoid free-riding, this divergence in expectations could make a mutually acceptable franchise fee agreement unreachable.¹³³

The solution suggested by Blair and Kaserman is for the franchisor to adopt both the franchise fee and output royalty arrangements.¹³⁴ The output royalty arrangement is employed here

in relationship, and provides firms the advantage of accessing to outside R & D activities. *See id.* at 320.

131. *See generally* Roger D. Blair & David L. Kaserman, *Optimal Franchising*, 48 S. ECON. J. 494 (1982).

132. The concept of net present value is to measure the difference in value between having a dollar at the present period and next period. Basically, to have \$1 today is always better than to have it next month. If one has \$1 today, he can invest it and have more than \$1 in the next month. Therefore, to calculate the current value of having \$1 next month, one has to discount it with the investment opportunities foregone because he does not have it today. In finance, the formula for the calculation of the net present value (NPV) is usually written as $NPV = C_0 + (C_1 / (1+r))$. Namely, with the investment of C_0 at the current period, one can have the payoff of C_1 at the next period. R is the discount rate for C_1 which is offered by comparable investment alternatives. *See* RICHARD A. BREALY & STEWART C. MEYER, *CORPORATE FINANCE* 13 (4th ed. 1991).

133. Blair & Kaserman, *supra* note 131, at 498.

134. *See id.* at 499.

to supplement the franchise fee in dealing with the problem of divergent discount rates and to reduce transaction costs associated with bounded rationality.¹³⁵ In a nutshell, the franchisor may have to accept a lower franchise fee at the outset to induce the franchisee to enter into the franchise relationship when the prospective value of the franchise is questioned. The output royalty, on the other hand, enables the franchisor to capture profits from the increase in sales if the franchise later proves to be successful. Thus, the output royalty serves as a credible signal to the franchisee that the franchisor will do his best to protect the value of the franchise. Alternatively, even if the discount rates applied by both parties are the same, Blair and Kaserman note that the franchisee will still be concerned about the probability that the franchisor might license another franchisee in the same market after it pays the net present value of the franchise.¹³⁶ Because of market saturation, the profits that could have been gained by the franchisee would be eliminated by competition from newcomers. The franchise fee together with territorial restrictions or exclusive dealing arrangements serve to deter the franchisor from engaging in hold-up behavior, and guarantee the franchisee that it will be able to collect the expected rents in the future.¹³⁷

Mathewson and Winter provide a dynamic model to illustrate how the combined use of vertical restraints could be understood within their "externality" theory.¹³⁸ Their theory takes into account downstream firms' conjectures about neighboring rivals' reactions to price changes.¹³⁹

To begin with, consider the scenario in which a distributor believes that its price change will be immediately matched by its rivals. In other words, the size of the distributor's market is invariant to its price decision.¹⁴⁰ Under this assumption, the horizontal externality will be non-existent because a distributor is unable to free-ride on its adjacent rival who raises its price in order to reflect its costs in providing the additional services or promotional efforts. Suppose that there were no informational externality either. The manufacturer's only concern would be the vertical externality; namely, the effect the higher unit price the manufacturer charges for its product would have on the distributor's incentive to provide services or promotional effort.¹⁴¹ In this situation, the manufacturer

135. See *id.*

136. See *id.*

137. See *id.* at 499-500.

138. See Mathewson & Winter, *supra* note 80.

139. For underlying assumptions see Mathewson & Winter, *supra* note 112, at 28-30.

140. These types of conjectures are referred to as "Loschian conjectures." *Id.* at 29.

141. See generally Mathewson & Winter, *supra* note 80 and accompanying text.

could sell its product at marginal cost to induce sufficient provision of required services, and then use a franchise fee arrangement or a requirement contract to collect the downstream rents.¹⁴²

What if we take the informational externality into consideration? Since the franchisee expects that its provision of product information will have the effect of increasing the demand for its rivals, it tends to provide too little advertising from the perspective of the manufacturer. Under this situation, Mathewson and Winter suggest that the franchise fee or requirement contract plus a wholesale price equal to the manufacturer's marginal cost will not be enough to induce the optimal provision of information. Rather, the manufacturer may have to lower the wholesale price below its marginal cost to enlarge retail profits and to further enhance the incentive to advertise the manufacturer's product.¹⁴³ The addition of a price floor agreement to the franchise fee or requirement contract would prevent profits from being dissipated by lower resale prices rather than providing more product information.¹⁴⁴

Next, assume that the conjecture at the downstream market is of the "Nash type."¹⁴⁵ Under Nash conjectures, each distributor takes its rivals' prices as given, and then sets its own pricing strategy accordingly. In other words, each distributor assumes that its rivals' prices are invariant to its own price changes. Given that the products for resale are homogeneous, each distributor has an incentive to lower its own price just a little bit below its rivals' in order to attract more customers from other distributors, and keep lowering it until it is equal to the distributor's marginal cost.¹⁴⁶ This incentive creates a horizontal externality, from the manufacturer's perspective, which implies the resale price may be too low to encourage the provision of required services by the distributors.

Assume first there is no informational externality. For the manufacturer to internalize both the vertical and horizontal externalities, Mathewson and Winter propose two combinations of contractual restraints. First, the manufacturer could attempt to transform the downstream competitive environment into one in which any price change will immediately be matched by rivals. For example, by imposing territorial restraints on the downstream firms, the distributors will be able to assume that their market areas will be

142. Mathewson & Winter, *supra* note 112, at 34, 35.

143. *See id.* at 37.

144. *See id.*

145. We refer to the conjectures underlying the Nash equilibrium concept in game theory. In a Nash equilibrium, each player, in deciding whether to continue to play a different strategy, assumes that the other player will continue to play his current strategy. *See, e.g.,* PINDYCK & RUBINFELD, *supra* note 19, at 421.

146. *See id.* at 428-29.

invariant to their own price changes and the horizontal externality would be virtually eliminated. By adding a franchise fee arrangement or requirement contract to the exclusive territorial protection as a rent-collecting mechanism, the manufacturer can arrange incentives for the distributor that closely match his own.¹⁴⁷ The second method is the use of minimum resale price maintenance to correct the underpricing incentive, and then a franchise fee arrangement to extract the downstream profits.¹⁴⁸

A similar argument applies to the situation where the informational externality exists. Territorial restrictions reduce the likelihood that one retailer's advertising effort will spill over to the benefit of another retailer. Indeed, the manufacturer could require retailers to do only local advertising, and carry out national market advertising on its own. Minimum resale price maintenance serves to internalize both the horizontal and informational externalities. Packaged with the franchise fee arrangement or requirement contract to transfer rents up to the manufacturer, these arrangements can come close to aligning the retailer's incentives with those of the manufacturer.

C. Policy Implications

Since both the Blair and Mathewson studies are based on various assumptions, it is difficult to come up with a general summary regarding the theory of optimal franchising. With the risk of oversimplification, however, we may tentatively conclude that optimal franchising theory suggests that the combined use of various vertical restraints is frequently spurred by the manufacturer's need to balance the conflict between incentive provision and the dilution of profits. In order to solve the free-riding problem, for example, the upstream firm may need to grant some sort of protection to its distributors if the costs of vertical integration are prohibitive. In the meantime, however, these incentive-enhancing measures, such as minimum resale price maintenance, exclusive dealing, or territorial exclusivity, simultaneously increase the downstream firm's market power and also its ability to capture part of the profits that could have been garnered entirely by the manufacturer. Arrangements such as maximum resale price maintenance, the franchise fee, or the requirements contract, on the other hand, are used together with those protective devices to extract the enhanced downstream profits.

Part of this justification for the manufacturer's use of vertical restraints, as a package, was recognized but rejected by the Supreme

147. See Mathewson & Winter, *supra* note 112, at 36.

148. See *id.*

Court in the now overruled *Albrecht v. Herald Company*.¹⁴⁹ In that case, the Supreme Court refused to accept the Court of Appeals' reasoning that a price ceiling for the resale of the defendant's newspaper was necessary to protect consumers from being overcharged by distributors who gained their monopoly power through exclusive territorial protection.¹⁵⁰ Rather, the Court held that the use of one type of illegal restraint to cure the pernicious effect of another illegal restraint is impermissible.¹⁵¹

The analysis here indicates why the doctrinal formula adopted in *Albrecht* is inconsistent with a policy that aims to enhance consumer welfare. The *State Oil* Court replaced the *Albrecht* rule with a rule of reason test that looks to the functions of the maximum price-fixing restraint.¹⁵² However, it is unclear, and too early to tell, whether one can infer from *State Oil* that courts, consistent with optimal franchising theory, will start to review all combined uses of vertical restraints under the rule of reason standard, even when they include restraints that are illegal per se under the current standard.

D. Some Implications for Doctrine

We have provided a large, though not exhaustive, catalog of procompetitive theories of vertical control. Our goal, in this piece, has been to provide a road map for courts to use in thinking through the rule of reason issues generated in vertical restraint cases.

This analysis stakes out a middle ground between the current approach of the law and the more radical per se legality position suggested early on in the Chicago School analysis of vertical restraints.¹⁵³ We have not addressed anticompetitive theories of vertical control here. Although our arguments have implications for the plausibility of anticompetitive theories in certain settings, we have not attempted to provide a general refutation of anticompetitive theories. Our goal, instead, has been to present much of the "post-Chicago school" development in the law and economics of vertical control. Much of this literature draws on the economic theory of the firm and the theory of contracts.¹⁵⁴

What does this analysis imply for the law? There are many ways to answer this, and a detailed analysis of the law is probably better

149. 390 U.S. 145 (1968).

150. See *id.* at 153.

151. See *id.*

152. See *State Oil Co. v. Khan*, 522 U.S. 3 (1997).

153. See generally BORK, *supra* note 3; Director & Levi, *supra* note 3.

154. See Martin K. Perry, *Vertical Integration: Determinants and Effects*, in 1 HANDBOOK OF INDUSTRIAL ORGANIZATION 183, 185 (Richard Schmalensee & Robert D. Willig eds., 1989) ("The theory of vertical integration is situated at the intersection of the theory of the firm, the theory of contracts, and the theory of markets.").

saved for another paper. However, some general conclusions can be offered now.

First, this analysis rather obviously suggests that the rule of reason should be applied across the board in the analysis of vertical control mechanisms. Antitrust law seems to be approaching this position, especially with the recent decision in *State Oil*, but the law still has a long way to go.

Second, antitrust courts should focus on the *functions* of the various vertical control methods in determining whether they violate the antitrust laws. Per se analysis is inappropriate because it permits and encourages the court to ignore the function of the vertical restraint. More important, even if rule of reason analysis were adopted with respect to a particular vertical control method, this analysis suggests that it should be applied from a functional perspective. This should be contrasted with the doctrinal perspective generally taken by courts.

These basic lessons can be illustrated with several areas of vertical restraint doctrine.

(1) *Resale Price Maintenance*

Minimum resale price maintenance remains per se unlawful.¹⁵⁵ The Court has shown a tendency lately to raise the burden of proof for plaintiffs in resale price maintenance cases. For example, in *Monsanto Company v. Spray-Rite Service Corporation*,¹⁵⁶ the Supreme Court held that in order to survive a motion for summary judgment, a terminated dealer must show something more than that the manufacturer terminated him after receiving complaints from other dealers.¹⁵⁷ In particular, under *Monsanto*, the plaintiff must show some evidence of conspiracy, "evidence that tends to exclude the possibility that the manufacturer and nonterminated distributors were acting independently."¹⁵⁸ In *Business Electronics Corporation v. Sharp Electronics Corporation*¹⁵⁹ the Court raised this burden of proof substantially higher, perhaps beyond the reach of most plaintiffs. The Court held in *Business Electronics* that an agreement between a manufacturer and a dealer to terminate another dealer is per se unlawful only if there is a concomitant agreement to maintain

155. The per se illegality rule was adopted in *Dr. Miles Medical Company v. John D. Park & Sons Co.*, 220 U.S. 373, 408-09 (1911). The rule was modified in *State Oil Company v. Khan*, 522 U.S. 3 (1997), bringing maximum resale price maintenance under the rule of reason test.

156. 465 U.S. 752 (1984).

157. *See id.* at 764.

158. *Id.* at 764.

159. 485 U.S. 717 (1988).

resale prices at some particular level.¹⁶⁰

What explains this pattern of maintaining the rule against minimum resale price maintenance while at the same time shifting the burden of proof to the detriment of plaintiffs? The Court's reasoning in both *Monsanto* and *Business Electronics* makes clear that it recognizes that the rule of reason justifications accepted in the area of territorial restraints (in *Sylvania*)¹⁶¹ apply to the analysis of resale price maintenance.¹⁶² Rather than overturn the per se rule against resale price maintenance, the Court has taken the more conservative path of raising the burden of proof for plaintiffs.

Our analysis makes clear, in economic terms, what is already implicitly the message in the Court's recent cases: rule of reason analysis should be applied both to territorial restrictions and to vertical resale price maintenance. We take the argument a step further, of course, and say that the rule of reason test should be sensitive to the economic function of the restraint.

(2) *Territorial Restrictions*

The analysis of vertical territorial restrictions was brought under the rule of reason in *Sylvania*, but even in this area of the law the analysis tends not to focus on the economic function of the restraint. *Sylvania* holds that a territorial restraint will be held reasonable if its benefits for interbrand competition outweigh its detrimental effects on intrabrand competition.¹⁶³ However, this doctrine appears too narrow when viewed in light of the theories presented in this paper.

160. See *id.* at 735-36.

161. *Continental T.V. v. GTE Sylvania*, 433 U.S. 36, 59 (1977).

162. See *Monsanto*, 465 U.S. at 762-63. In *Monsanto*, the Court noted that:

A manufacturer and its distributors have legitimate reasons to exchange information about the prices and the reception of their products in the market. Moreover, it is precisely in cases in which the manufacturer attempts to further a particular marketing strategy by means of agreements on often costly nonprice restrictions that it will have the most interest in the distributors' resale prices. The manufacturer often will want to ensure that its distributors earn sufficient profit to pay for programs such as hiring and training additional salesmen or demonstrating the technical features of the product, and will want to see that 'free-riders' do not interfere. See *Sylvania* [433 U.S. at 55]. Thus, the manufacturer's strongly felt concern about resale prices does not necessarily mean that it has done more than the *Colgate* doctrine allows.

Id. See also *Business Electronics v. Sharp Electronics Corp.*, 485 U.S. 717, 726 (1988) ("Our approach to the question presented in the present case is guided by the premises of *GTE Sylvania* and *Monsanto*: that there is a presumption in favor of a rule-of-reason standard; that departure from that standard must be justified by demonstrable economic effect, such as the facilitation of cartelizing, rather than formalistic distinctions; that interbrand competition is the primary concern of antitrust laws; and that rules in this area should be formulated with a view towards protecting the doctrine of *GTE Sylvania*.").

163. 433 U.S. at 54, 58-59 (White, J., concurring).

It suggests that a territorial restraint should be invalidated when it provides no apparent enhancements to interbrand competition. This is too narrow, because it is possible that a territorial restraint may benefit consumers without substantially enhancing interbrand competition. For example, consider the case where there is no interbrand competition and the monopolist manufacturer uses territorial restraints to avoid or eliminate the successive monopoly problem. A functional approach to rule of reason analysis would avoid applying an excessively narrow test to vertical control methods.

(3) Vertical Integration

Vertical integration doctrine—specifically, vertical merger doctrine—remains largely dominated by the “foreclosure theory” adopted early in *du Pont*.¹⁶⁴ Of course, the Justice Department Guidelines have moved beyond the early case law, by giving greater scope to efficiency defenses.¹⁶⁵ In addition, before the premerger notification law had the effect of freezing the development of merger doctrine, the case law had moved substantially in the direction of applying a rule of reason test.¹⁶⁶ Still, the foreclosure theory remains a starting point for antitrust attacks on vertical mergers and exclusive dealing arrangements.¹⁶⁷

But foreclosure is only one of many economic effects associated with vertical integration. And it is easy to construct an example in which foreclosure may be desirable. Suppose a monopolist is selling to a distributor who himself enjoys a high degree of monopoly power. The monopolist manufacturer must worry about the “double-marginalization” problem introduced in our discussion of successive monopoly: the monopolist will apply a monopoly surcharge to the item he produces, and the distributor will apply an additional monopolistic surcharge to the same item at the downstream level.¹⁶⁸ The monopolist can eliminate this problem by vertically integrating

164. *United States v. E.I. du Pont de Nemours & Co.*, 353 U.S. 586, 592 (1957). On the continuing importance of the foreclosure theory, see, e.g., David Reiffen & Michael Vita, *Comment: Is There New Thinking on Vertical Mergers?*, 63 ANTITRUST L.J. 917, 918-19 (1995)(criticizing foreclosure theory).

165. The significant innovation in enforcement rules came in the 1984 guidelines. See generally U.S. Dept. of Justice Merger Guidelines (1984).

166. The shift toward rule of reason analysis in vertical merger doctrine occurred through an indirect route. The famous decision in *Brown Shoe Company v. United States*, set out a truncated rule of reason test for both vertical and horizontal mergers. See *Brown Shoe*, 370 U.S. 294, 328 (1962)(setting out rule of reason test for vertical mergers). Then the Supreme Court, in *United States v. General Dynamics Corp.*, 415 U.S. 486, 510-11 (1974), reinterpreted *Brown Shoe* in a way that made it difficult for enforcement authorities to evade the burdensome requirements of the rule of reason test.

167. See Reiffen & Vita, *supra* note 164, at 917-18.

168. See text accompanying notes 64-67.

forward, and transferring his own products at marginal cost to the downstream segment of the integrated unit. In this example, the monopolist gains, and so do consumers. Note also that because the downstream distributor was assumed initially to have monopoly power, this case may involve a high risk of foreclosure, because the downstream segment of the integrated unit may focus on selling only the monopolist's product.

Further, given the horizontal externalities observed in manufacturer-distributor relationships, vertical integration and exclusive dealing should be seen as potential solutions to these problems. Foreclosure of competitors may result in these cases, but consumers gain to the extent vertical control enhances the provision of services at the downstream level.

(4) *The Tying Doctrine*

Tying analysis in vertical settings is still controlled by a partial per se rule, initially set out in full in *Northern Pacific Railway Company v. United States*.¹⁶⁹

The per se rule applies "whenever a party has sufficient economic power with respect to the tying product to appreciably restrain free competition in the market for the tied product and a 'not insubstantial' amount of interstate commerce is affected."¹⁷⁰

The Court has made no effort to distinguish tying as a vertical control mechanism from other cases of tying. As a vertical control mechanism, one should feel somewhat safer in assuming that the consumer misperception issues are for the most part unimportant,¹⁷¹ because people who make a business of dealing with products are likely to be well informed about the items they buy and sell.

We have emphasized two potential functions for tying as a vertical control mechanism: (1) tying may be used by an upstream manufacturer to protect a product's reputation by preventing free-riding by distributors, or (2) tying might be used by an input monopolist to prevent a downstream firm from substituting away from the monopolized input.

In the former case, consumers are likely to benefit from the

169. 356 U.S. 1 (1958).

170. *Id.* at 6. See also *Eastman Kodak Co. v. Image Technical Servs., Inc.*, 504 U.S. 451, 461-62 (1992); *Jefferson Parish Hosp. Dist. No. 2 v. Hyde*, 466 U.S. 2, 9-10 (1984); *Fortner Enter., Inc. v. United States Steel Corp.*, 394 U.S. 495, 497-99 (1969).

171. On consumer misperception as a basis for applying antitrust law to tying, see Richard Craswell, *Tying Requirements in Competitive Markets: The Consumer Protection Issues*, 62 B.U. L. REV. 661 (1982). Craswell's point is that tying may be both harmful to consumers and profitable for the firm in a competitive market in which consumers lack sufficient information to tell whether the tied (or tying) product's price exceeds the competitive level.

prevention of free-riding at the downstream level.¹⁷² This function should be taken into account in tying analysis.¹⁷³

Consider the case in which tying is used by the input monopolist to prevent substitution at the downstream level. It is fair to describe this as a case in which tying is used to enhance monopoly power.¹⁷⁴ In the extreme case in which the input monopolist ties the monopolized input to all of the potential input substitutes, the input monopolist can replicate the outcome under vertical integration.¹⁷⁵ However, this extreme is unlikely to be observed. The more likely result is that the input monopolist will be able to reduce without completely eliminating the downstream firm's power to substitute away from the monopolized input. Society loses to the extent tying increases the monopoly power of the input monopolist. However, social welfare improves to the extent the allocation of productive inputs remains closer to the cost-minimizing combination. Thus, even in the case in which tying is used in a vertical setting solely to enhance monopoly power, there may be offsetting welfare gains (as the sum of consumers and producers surplus) that should be taken into account.¹⁷⁶

As this brief discussion shows, antitrust case law has not developed the appropriate degree of sensitivity to the functions of vertical control mechanisms. Scanning across vertical merger and vertical restraint doctrine, as we have here, one sees a variety of *per se* and rule of reason doctrines, with few if any of them encouraging a careful review of the function of the relevant vertical control

172. It is not certain, of course, that consumers will always benefit. Whether tying for the purpose of preventing free-riding benefits consumers depends on the extent to which consumers benefit by the provision of better services. In some settings, the marginal benefits of enhanced service-provision may be greater than its costs. See F. M. SCHERER & DAVID ROSS, *INDUSTRIAL MARKET STRUCTURE AND ECONOMIC PERFORMANCE* 542-48 (3d ed. 1990). However, competition should lead firms to offer enhanced service-provision in those markets in which the consumer benefits are large.

173. For an argument for applying rule of reason analysis to tying that is consistent with this theory, see *Jefferson Parish Hospital District No. 2 v. Hyde*, 466 U.S. 2, 40-42 (1984) (O'Connor, J., concurring). Justice O'Connor's concurrence argues that coordination problems and the need to prevent free-riding justify the hospital's decision to close its anesthesiology department.

174. On the general question whether certain vertical control methods can actually enhance monopoly power, see Louis Kaplow, *Extension of Monopoly Power Through Leverage*, 85 COLUM. L. REV. 515 (1985) (contending that tying can enhance monopoly power), and Ward S. Bowman Jr., *Tying Arrangements and the Leverage Problem*, 67 YALE L.J. 19 (1957) (contending that tying cannot enhance monopoly power).

175. See generally Blair & Kaserman, *supra* note 116.

176. Williamson has argued that efficiency gains should be taken into account in merger cases. See Oliver E. Williamson, *Economies as an Antitrust Defense: The Welfare Tradeoffs*, 58 AM. ECON. REV. 18 (1968). A similar welfare-tradeoff analysis can be applied to tying in the monopoly-enhancement case.

mechanism. These rules must be replaced with a functionally oriented, rule of reason analysis if there is to be any hope for the courts giving careful consideration to the arguments presented in this article.

Conclusion

In this article, we have surveyed procompetitive theories of vertical control, under a taxonomy of market structures. We hope this approach will provide courts and enforcement agencies with a clear analytical path to follow in evaluating the competitive impact of various vertical control mechanisms, whether vertical integration or vertical restraints. In particular, by using, as we have, market structure as a starting point, we hope courts will find it easier to prioritize relevant factors.

We have also devoted a substantial part of this paper to demonstrating the functional equivalence of vertical integration and various vertical restraints in reducing costs associated with the principal-agent and successive monopoly problems, and we have demonstrated the complementary roles that vertical control mechanisms play in supplementing each other in an optimal distribution scheme. These arguments suggest that courts should reconsider the validity of applying different legal standards to different types of vertical restraints, and replace rigid doctrinal categories with a functionally oriented, rule of reason analysis.

Appendix

I. Vertical Integration and Variable-Proportions Productions

In this section, we elaborate on the argument of Part III.D of the text. As Figure 4 and the accompanying analysis will show, the use of vertical integration under variable proportions production could increase the integrated firm's profits but could be welfare reducing for consumers.¹⁷⁷

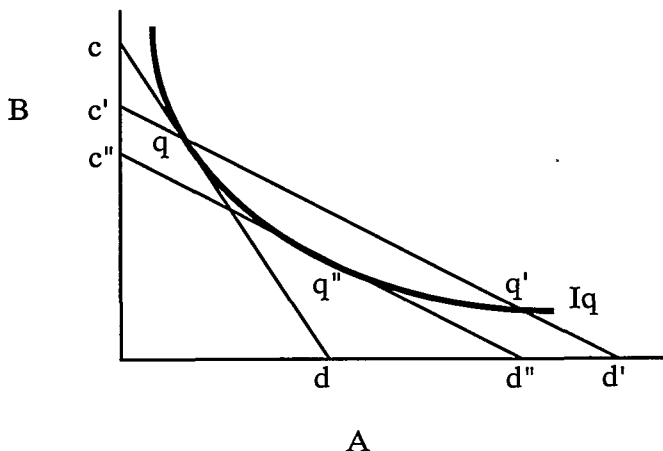


Figure 4

Suppose a competitive final market in the final product is composed of two inputs, A and B. Assume also that input market A is monopolized and input B is competitively supplied. The curve I_q in Figure 4 shows all the possible combinations of input A and B with their given prices that will yield the same industry output, Q . In a similar vein, the line cd represents the various combinations of both inputs with their given prices that will generate the same total production costs, and, therefore, could be understood as the budget constraint on the production of Q .¹⁷⁸ The other aspect of the segment cd is that point c on the vertical axis represents the maximum quantity

177. The diagram and analysis is drawn from John M. Vernon & Daniel A. Graham, *Profitability of Monopolization by Vertical Integration*, 79 J. POL. ECON. 924 (1971). See also SHUGHART, *supra* note 57, at 321-24; BLAIR & KASERMAN, *supra* note 117, at 48-50; SCHERER & ROSS, *supra* note 69, at 522-27.

178. In economic parlance, I_q and cd are called respectively the "isoquant" and "isocost" curves. See PARKIN, *supra* note 17, at 260, 263-64.

of B that could be purchased if the downstream firm spends all of its budget on B. Hence, the total costs of production can also be expressed as the marginal cost of B times c .

With I_q and the budget restraint of cd , the cost-efficient way of using these two inputs will occur at the point q , where cd is tangential to I_q . To see the result, compare point q' on I_q with q .¹⁷⁹ It is obvious from Figure 4 that q' lies above the line segment cd . It means that the cost of using the input combination q' , represented by the cost line $c'd'$, to produce Q is higher than that of q . Indeed, given the shape of I_q , this observation holds for all points other than q on I_q .

In order to examine how vertical integration will affect the production mode, let us start hypothetically with the case where both A and B are competitively supplied. To avoid unnecessary complications, suppose that $c'd'$ is the line of budget constraint under this circumstance and, similarly, could be expressed as the marginal cost of B times c' . As the price of B remains unchanged, the difference between c and c' is solely an outcome resulting from the change in the pricing pattern at market A. That is, it represents the increase of profits that could have been appropriated by the monopolist had it sold A at its marginal cost. This provides an incentive for the input monopolist to integrate forward to the downstream market. Suppose that the monopolist acquires all downstream firms and then transfers input A internally to its downstream division at the competitive price. If the integrated firm decides to maintain the output level at Q , what input combination will it select? Although point q is on $c'd'$ and, therefore, is within the firm's budget constraint, it is not the least costly input combination for the integrated firm. Rather, holding constant the price ratio between the now competitively supplied A and B, one can find a new curve $c''d''$ that is tangential to I_q at q'' in Figure 4 and represents a lower cost of production. Also, with the cost curve of $c''d''$, vertical integration enlarges the monopolist's profits by the amount of $c-c''$. In this example, vertical integration increases the monopolist's power, harming consumers. However, total welfare increases because the integrated unit adopts the cost-minimizing input combination.

II. Vertical Control in Non-Competitive Markets: The Case of Monopsony

This analysis has focused on the assumption of an upstream monopoly. In a *monopsonistic* market, in which the downstream firm is the sole buyer of the upstream intermediate input, there is a strong economic incentive for the monopsonist to integrate backward. To

179. See *id.* at 266.

demonstrate this potential, Figure 5 graphically illustrates a simple monopsonistic pricing model.¹⁸⁰

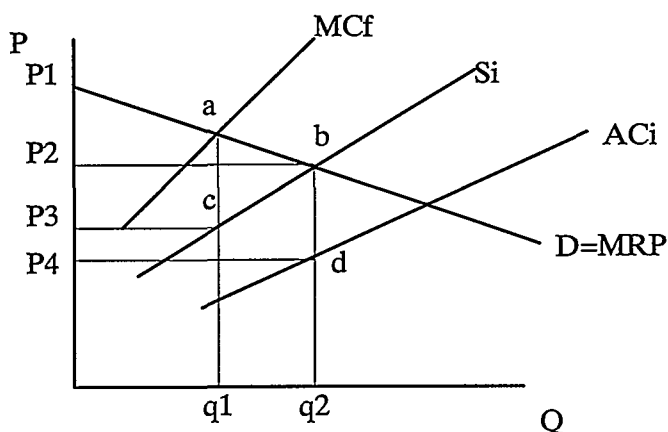


Figure 5

Suppose that the market for input supply is competitive and the demand for the input is controlled by a single buyer with a demand schedule D . For the monopsonist, D reflects the extra revenue from selling the final good produced by the last unit of input purchased, and is technically called the marginal revenue product curve ("MRP").¹⁸¹ With an upward-sloping input supply curve S_i , the marginal cost for the monopsonist to use one extra unit of the input, represented by MC_f in Figure 5, must lie above S_i . The reason is that the monopsonist has to pay a higher price for *all* units of input rather than only one more unit if it wishes to purchase that one additional unit. To further clarify this point, consider a hypothetical example in which a tobacco producer in an isolated small town hires all the residents who wish to work for the company. Assume that the producer currently hires 10 workers at the wage of \$10 per day. If the wage information is available to every employee, the producer will not be able to hire another worker and pay him \$11 without increasing the other employees' wages. In other words, in order to hire an extra worker, the producer has to pay \$11 for the eleventh worker plus a one-dollar increase for each of the workers previously hired. The marginal cost for using the extra unit of input in this case

180. See ROGER D. BLAIR & JEFFREY L. HARRISON, *MONOPSONY* 36-42 (1993); SHUGHART, *supra* note 57, at 325-27; CARLTON & PERLOFF, *supra* note 21, at 152. Linear input supply curve and marginal cost curve are assumed in this model.

181. See PARKIN, *supra* note 17, at 413; BLAIR & HARRISON, *supra* note 180, at 39 n.7.

will be \$21 which is greater than the price for that unit \$11.¹⁸²

With similar profit-maximizing objectives as that of a monopolist, the monopsonist chooses to purchase q_1 units of the input where its MRP is equal to MCf. However, at q_1 , the monopsonist actually pays for the input according to S_i , which yields a price for the input of P3. The profits that could be earned from selling the output generated by q_1 unit of input could be measured by the area of P1P3ca. Comparing the price and output levels when the input is competitively demanded, P2 and q_2 , the monopsonist purchases less and, therefore, sells less. From the standpoint of society, the monopsonist's exercise of its buying power creates a deadweight loss that is equivalent to the area abc.

From this setup, we can also observe that there is an incentive for the monopsonist to integrate *backward* to the upstream industry. Suppose that the monopsonist acquires the upstream industry and, accordingly, all resource allocation regarding the production of the final good is now determined internally. The monopsonist under this circumstance no longer has to worry about the increase in marginal cost in connection with the increased use of the input. Rather, it could price discriminate against the input suppliers. Namely, it could pay each additional unit of input according to the price schedule on S_i without the need to pay all previous units the same higher prices. Hence, MCf becomes irrelevant for the monopsonist's purchase decision. The monopsonist now will expand its use of the input up to the new equilibrium level q_2 , where its new marginal cost curve S_i is equal to MRP. Given that the *average* cost of using q_2 is P4, which is the corresponding point on the monopsonist's average cost curve AC_i ,¹⁸³ the monopsonist could earn a total profit of P1P4db, which is obviously larger than that before vertical integration. It also eliminates the deadweight loss by raising the final output level to the level attained when no monopsony had existed. In turn, consumers stand to gain by paying lower prices for the final goods.¹⁸⁴

182. More formally, let TC_l be the monopsonist's total cost of labor. $W(l)$ is the labor supply function. By definition, we have

$$TC_l = W(l) * l.$$

Differentiate TC_l with respect to l yields the monopsonist's marginal cost of labor which is greater than $W(l)$. Namely,

$$\frac{dTC_l}{dl} = W(l) + W'(l) * l > 0 \text{ if } W'(l) > 0 \text{ if } W'(l) > 0.$$

183. To see how AC_i lies below S_i intuitively, imagine a class of 50 students with the average test score of 50. In order to raise the average test score to 51, the fifty-first student has to score more than 51 to achieve that goal.

184. See SHUGHART, *supra* note 57, at 327; BLAIR & KASERMAN, *supra* note 9, at 309-11; see also Perry, *Vertical Integration: The Monopsony Case*, 68 AM. ECON. REV. 561 (1978).

III. The Successive Monopoly Problem

Here we present a simple mathematical description of the successive monopoly problem, and potential solutions. Suppose an upstream manufacturer sells all of its output through downstream retailers. Let the upstream cost function be $C_u = c_u q_u$, and the downstream cost function be $C_d = c_d q_d$. Downstream inverse demand is $p_d = a_d - q_d$; upstream, it is $p_u = a_u - q_u$; $a_d > a_u$.

The downstream firm maximizes profit, given the wholesale price, p_u , charged by the upstream firm. This generates a "derived demand" for the upstream firm's product. Suppose the upstream market is competitive and the downstream market is monopolized. The profit maximizing quantity choice for the downstream firm will be $q_d^m = (a_d - c_d)/2 = (a_d - p_u)/2$. (To see this, set marginal revenue, $(a_d - q_d) - q_d$, equal to marginal cost, c_d , and solve for q_d .) And the downstream price will be $p_d^m = (a_d + p_u)/2$.

The upstream firm, which is competitive, will face a market price $p_u = c_u$, and will therefore sell to meet the downstream demand $(a_d - p_u)/2$, and this implies $p_u^c = c_u$, $q_u^c = (a_d - c_u)/2$.

Now suppose the upstream market is monopolized, as well as the downstream market. The downstream firm chooses $q_d^m = (a_d - c_d)/2 = (a_d - p_u)/2$, and $p_d^m = (a_d + c_d)/2 = (a_d + p_u)/2$. The upstream firm now chooses a price and quantity combination that maximizes $\pi_u = (p_u - c_u)((a_d - p_u)/2)$. This yields the combination $p_u^m = (a_d + c_u)/2$, $q_u^m = (a_d - c_u)/4$.

These results show that the "double-marginalization" that occurs in the "monopoly-monopoly" scenario leads to considerably higher prices and lower quantity for consumers. Under the "competitive upstream-monopoly downstream" case, the ultimate price-quantity combination given to end-use consumers is $p_d^m = (a_d + c_u)/2$, $q_d^m = (a_d - c_u)/2$. This is the combination that would be chosen by an integrated monopolist. Under the monopoly-monopoly setting, the price-quantity combination facing end-use consumers is $p_d^m = (3a_d + c_u)/4$, $q_d^m = (a_d - c_u)/4$. The total welfare loss to consumers in moving from the "competition-monopoly" to the "monopoly-monopoly" case is given by $(1/2)(\Delta p)(\Delta q)$ where Δp is the change in price, Δq is the change in quantity. The loss in consumer welfare is therefore $(1/8)(a_d - c_u)^2$.

The upstream firm would improve its position by selling its good at marginal cost to the downstream retailer, and charging a fixed fee for the service. Under monopoly-monopoly, the upstream firm's profit is equal to $(1/8)(a_d - c_u)^2$. If the upstream firm were to charge a fixed fee to the downstream retailer, it could expropriate the downstream firm's profit. Thus, the upstream firm can charge any fixed (say, per-year) fee, F , such that $F \leq (p_d - c_d)((a_d - c_u)/2) = (1/4)(a_d$

$-c_u)^2$. Thus, consumer welfare can be enhanced and the profits of the upstream firm enhanced through vertical integration, or through a relationship in which the upstream firm sells at marginal cost and charges a fixed franchise fee to the downstream retailer.

Let us consider the related issue of rent-extraction by an upstream monopolist. Suppose the upstream market is monopolized and the downstream market is competitive. The quantity produced by downstream firms is therefore $q_d^c = a_d - c_d = a_d - p_u$. The downstream price is $p_d^c = c_d = p_u$.

The upstream monopolist will maximize its profits, taking into consideration the actions of the downstream firms. Thus, the upstream monopolist will set p_u in order to maximize $\pi_u = (p_u - c_u)(a_d - p_u)$. Setting marginal revenue (in price), $a_d - 2p_u$, equal to marginal cost, c_u , yields the solution $p_u^m = (a_d + c_u)/2$, $q_u^m = (a_d - c_u)/2$. This shows that the upstream monopolist can extract the full monopoly profit from end-use consumers by charging appropriately high prices to the downstream retailers.

IV. Externalities in the Manufacture-Distribution Relationship

Now suppose the downstream firm's sales depend to some extent on effort invested into promotion and advertising. If the upstream firm integrated forward, it would try to maximize $(p_d - e - c_u)q_i(e)$ where i denotes the sales of a particular retailer, and e represents effort. In the non-integrated setting, however, the retailer maximizes $(p_d - e - p_u)q_i(e)$. Clearly, if $p_u > c_u$, the retailer will choose a different level of e than that chosen by the integrated manufacturer.

Assume for now that there is only one retailer. The marginal benefit from effort for the retailer is $(p_d - e - p_u)q'(e)$, where $q'(e)$ represents the derivative of q with respect to e . The marginal benefit for the integrated manufacturer is $(p_d - e - c_u)q'(e)$. Since $(p_d - e - c_u)q'(e) > (p_d - e - p_u)q'(e)$, and since marginal cost of effort is the same and equal to one for both, the retailer chooses *too little effort* from the manufacturer's perspective. This is the *vertical externality*, described by Mathewson and Winter.

Now suppose these are two retailers. For the individual retailer, the marginal benefit of effort is $(p_d - e - p_u)(\partial q_1 / \partial e_1)$. For the manufacturer, the marginal benefit is $(p_d - e - c_u)(\partial q_1 / \partial e_1 + \partial q_2 / \partial e_1)$. As long as $\partial q_2 / \partial e_1 > 0$, the manufacturer will prefer greater effort than will the retailer. The captures the *spillover externality*.

The third externality results from sales by discounters. The manufacturer gains even when a discounter sells his product, but the retailer gets nothing. Hence, it should be clear, the retailer's incentive to invest effort is pushed, by this effect, further below that of the manufacturer.

Now consider the possibility that the retailer could increase his individual sales by reducing p_d . The manufacturer determines the level of effort that maximizes $(p_d - e - c_u)q(e, p_d)$. Let that level be e^* . However, the retailer may choose to increase his own sales by reducing p_d and setting $e < e^*$. Resale price maintenance allows the manufacturer to control this incentive. By setting a price floor $p_d^f = e^* + p_u$, the manufacturer can induce the retailer to choose $e = e^*$, and competition will lead the retailer to choose $p_d = p_d^f$.

