Chapter 11

VERTICAL CONTRACTUAL RELATIONS

MICHAEL L. KATZ*

University of California at Berkeley

Contents

*This chapter is dedicated to one of my graduate advisors, J.F., who in 1979 told me to stop working on vertical restraints because all of the interesting theoretical questions had been definitely answered and the public policy debate was over. I am grateful to Mark Gersovitz for extensive comments on an earlier draft of this chapter.

Handbook of Industrial Organization, Volume I, Edited by R. Schmalensee and R.D. Willig

1. Introduction

This chapter is a study of the contractual relationship between two parties at successive stages in the vertical chain of production and marketing for a good. In other words, this chapter surveys the theory of the sale of intermediate goods.

There are several important respects in which intermediate good markets differ from final good markets and thus merit independent study.[1] First, intermediate good markets often involve large transactions made by sophisticated buyers. Second, the products being sold may possess very complex bundles of attributes, making problems of moral hazard more severe, or at least more complicated. Third, the buyers' demands for an intermediate good are interdependent when the buyers are product-market competitors with one another. Fourth, the buyers of an intermediate good typically are involved in a game in the downstream product market, and the sales contract for the upstream product may affect the equilibrium of this downstream game. Lastly, buyers of intermediate goods often can credibly threaten to integrate backward into supply of the intermediate good.[2]

A simple uniform posted price often is held up as the typical contract form in final good markets. Given the sophistication of buyers and the large scale of individual transactions, more complex selling schemes may be practicable intermediate good markets. Sellers may utilize sophisticated pricing mechanisms or nonprice contractual provisions. This chapter explores the private and social incentives to utilize such contracts.

In looking at contracts that go beyond a simple uniform price, attention is focused on several contractual provisions that often are seen in practice.

Quantity-dependent pricing

The price that the buyer pays per unit of the intermediate good may depend on the total amount of the good that the buyer purchases. A two-part tariff is a particularly simple form of quantity-dependent pricing. Under a two-part tariff,

[1] In that portion of Chapter 14 of this Handbook devoted to patent licensing, Jennifer Reinganum examines the terms of sale for one particular type of input: knowledge. As is evident after having read her chapter and the present one, the work on patent licensing has evolved along rather different lines than has the traditional literature on vertical contracts.

[2] Of course, these effects sometimes arise in final goods markets. For example, demands are interdependent in the cases of snob goods, bandwagon goods, and goods subject to network externalities. The threat of integration may arise in the provision of services (e.g. house cleaning and gardening).

the buyer pays a fixed fee plus a constant per-unit charge to purchase the input. Thus, the average per-unit price falls with the purchase volume. In the vertical relations literature, the fixed fee in a two-part tariff often is called a *franchise fee*. Another well-known form of quantity-dependent pricing is termed *forcing* in the vertical relations literature. Under a forcing contract, the dealer is required to purchase a given amount of the input. Such a requirement can be thought of as a quantity-dependent pricing scheme under which the buyer's total payment to the input supplier is fixed for any quantity of input below some set level. The practice of *bundling* two units of a good into a single package also is an example of quantity-dependent pricing.

Quantity-dependent pricing need not be confined to a single good. A multi-product seller may make the price paid by the buyer for any one input a function of the quantities of all of the inputs that he purchases from that seller. For example, if one firm sells both inputs A and B, the amount charged for a unit of input A may depend on the size of the buyer's purchase of input B.

At times, the term "vertical restraint" has been used to refer to any contractual provision between two levels in the supply chain that goes beyond the use of a simple uniform price. Thus, some authors consider quantity-dependent pricing of the sort just described to be a vertical restraint. In this chapter, I reserve the term "vertical restraint" for those provisions of an intermediate good contract that make the buyer's payment to a given seller depend on variables other than the quantities of the inputs purchased from that seller.

Three such provisions can be thought of as multiproduct quantity-dependent pricing scheme under which the price charged by a given seller depends, in part, on the quantities purchased from *other* input suppliers.

Ties

A *tie* occurs when a manufacturer agrees to sell input A to a buyer only if the buyer also purchases input B from that manufacturer (and no other). A tie can be thought of as a type of multi-product (across firms) quantity-dependent pricing; the seller's price for A is infinite if the quantity of B purchased from another firm is positive. In general, the price at which the tied goods are sold also may vary with the quantities purchased from the seller who imposes the tie. As a rule, however, the vertical restraints literature has focused almost exclusively on ties for which the prices are independent of these quantities. One exception to this rule is the consideration of *multi-product bundling* under which units of several inputs are sold in fixed proportions at a package price.

It is worth noting that while a tie may be enforced through a contract, it need not be. When several inputs must be used together (i.e. form a system), it may be possible for the manufacturer to create a technological tie by designing his

products so that they can work only with each other (i.e. are incompatible with components produced by other manufacturers).

Royalties

The amount paid by the buyer to the seller need not be based solely on the unit sales of the intermediate good. Under a royalty scheme, the input buyer's payment is a function of his sales in the final goods market, measured in terms of either dollar revenues or physical units. Because these sales depend on the quantities of other inputs employed, royalties are equivalent to a certain form of multiproduct quantity-dependent pricing that cuts across input suppliers.

Requirements contracts and exclusive dealing

Under a *requirements contract*, a purchaser of input A agrees to make all of his purchase of input A from the same manufacturer. A requirements contract also can be thought of as a tying arrangement under which input A is tied to itself. In those cases in which the input being sold is a branded final product for which the buyer serves as a retailer, a requirements contract is also known as an *exclusive dealing* arrangement. Under an exclusive dealing arrangement, the dealer agrees not to sell any of the brands of rival manufacturers. The exclusive dealing arrangement may limit the retailer to selling the goods of only one producer (e.g. a McDonald's franchise only sells McDonald's brand products). Alternatively, the dealer may be allowed to carry the products of several manufacturers who are not direct competitors with one another (e.g. a grocery store might be restricted to selling only one brand of light bulb but still be allowed to sell thousands of other products).

In addition to specifying pricing practices, the intermediate good sales contract may govern the terms under which the buyer can sell the final output produced using the manufacturer's input. These provisions, known as resale restraints, are most prevalent in intermediate good markets in which the buyer serves as a retailer for the manufacturer's product. Two types of resale restraint are both the empirically most important and theoretically most analyzed.

Resale customer restraints

The input sales contract may contain provisions that restrict the set of consumers to whom the purchaser may sell the final output produced with the input. These

restrictions may be framed in terms of geographic location, or in terms of some other customer attribute.

Exclusive territories, or territorial restrictions involve the manufacturer's giving each dealer the sole right to serve consumers in a specific geographic region. This geographical exclusivity may be enforced by limiting the geographic location of authorized dealers while allowing a dealer to serve any customer who comes into the store. Alternatively, the manufacturer may explicitly forbid a retailer from serving consumers who are located in the territory assigned to another dealer. This type of restriction is necessary when consumers could otherwise be induced to cross sales territories.

The second type of resale restraint is a *customer restriction*. Under customer restrictions, certain classes of consumer are reserved for the manufacturer's own sales force. A personal computer manufacturer, for example, might reserve sales to Fortune 200 companies for the manufacturer's national sales force.

Resale price restraints

In addition to limiting the customers that a dealer may serve, the contract may restrict the price at which the dealer can sell his output. This restriction, known as *resale price maintenance*, may take the form of a price floor, a price ceiling, or both simultaneously.

The bulk of this chapter concerns itself with the private incentives to utilize the six contractual provisions just identified. In Sections 2 through 4, I use increasingly complex models to examine the private motivations for intermarket contracts that go beyond a simple uniform-pricing scheme. Within each section, I first identify conditions motivating the use of simple quantity-dependent pricing. I then show that under these conditions, there exist private incentives to implement vertical restraints as well.

In many instances, a given practice may be a potential response to several different contracting problems. No attempt is made in these sections to assess which motivation or the use of a particular type of vertical restraint is the leading or most important one. This point is taken up briefly in the discussion of public policy presented in Section 5, where I compare the social and private incentives to implement different vertical contractual provisions.

Before proceeding with the analysis of the private and social incentives to use sophisticated pricing and direct vertical restraints, there are two points worth noting. The first concerns the nature of benchmarks to which to compare vertical restraints. Much of the existing literature on vertical restraints views their use as

an intermediate case falling between the two extremes of uniform pricing and vertical integration. In comparing vertical restraints with vertical integration, this literature typically views integration as giving one party unified interests and complete control of all aspects of the combined operations.[3] A better term for what is examined might be an "unnatural union" of both interest and incentives across levels.

Sophisticated pricing and vertical restraints are responses to problems of moral hazard, adverse selection, and the need to share risk. How does vertical integration solve these problems? It is far from self-evident that these problems should disappear simply because the two parties are labelled as being in a single firm. Grossman and Hart (1986) adopt the extreme view that vertical integration does nothing to alter either the set of feasible contracts between the stages of production or the goals of the self-interested decision-makers at the two stages. While extreme, it seems to me to be a more defensible view than the opposite extreme that has been taken by much of the literature on vertical restraints and vertical integration.

To date, we have no fully articulated and convincing theory of what vertical integration *is*. Accordingly, I view the subject of this chapter to be a theory of contractual arrangements between different stages in the production chain, either within a single firm or across unrelated firms. There is little in the analysis below that lets one distinguish between a vertical restraint used by two divisions within a single firm and a vertical restraint used as part of an arm's length transaction between two separate firms.

The other point to note before getting on with the analysis concerns the relationship between the vertical restraints literature and the principal–agent literature. The study of vertical contractual relationships is just a special case of the general analysis of the principal–agent problem. The fully general analysis of the principal–agent problem remains to be conducted, however. In the interim, vertical restraints constitute an interesting and important class of special cases in which the additional structure allows one to develop a more complete characterization of the equilibrium contract than might otherwise be possible. For example, in the vertical restraints literature, attention is concentrated on a relatively narrow set of contractual provisions. While artificially limiting the set of feasible contracts often leads to incorrect conclusions, at least there is an extensive body of data showing that these practices are of empirical importance. This pattern suggests that more sophisticated provisions may be unnecessary, infeasible, or too costly to implement.

By drawing attention to characteristics of actual markets, the study of vertical restraints also points out two rather large holes in the standard principal–agent

[3] For examples of this approach, see Chapter 4 by Martin Perry in this Handbook.

literature. First, vertical restraints typically involve a principal with multiple agents, a problem that only recently has been given serious attention in the principal–agent literature. Second, vertical restraints usually are imposed in markets in which the agents of several different principals compete with one another. Such game-playing agents have been almost entirely neglected in the principal–agent literature. Unfortunately, while the existence of game-playing agents has been obvious to those studying vertical restraints, these authors have, for the most part, chosen to ignore formal analysis of such problems, relying instead on unwarranted extrapolation from models in which there is a single principal with one or more agents.

2. Private incentives: A single manufacturer with a single dealer

This section and the next two examine the private incentives to utilize vertical restraints. These incentives are illustrated through the use of a simple model which is presented in stages of increasing complexity in order to show the various roles that specific contractual provisions may play. As a starting point, this section considers an upstream monopolist who sells his output to a downstream monopolist. The downstream firm's output may be a final good or may itself be an input into yet another stage of production. With only a single party at each vertical stage, there is no role for the assignment of exclusive territories or the requirement of exclusive dealing. But all of the other practices discussed in the introduction may be feasible and, as will be shown, there may be private incentives to implement each of them.

2.1. The basic model

There is a single producer of an intermediate good, x, which is used by a single downstream firm in the production of output. I will refer to the upstream firm as the "manufacturer" and the downstream firm as the "dealer", but the downstream firm need not be merely retailer for the upstream firm. For example, the "manufacturer" might be a coal mine and the "dealer" an electric utility. Moreover, although I will refer to them as separate firms, for most of the analysis the buyer and seller could equally well be two business units within a single firm.

The manufacturer's output is combined with another input, y (competitively supplied by firms in another upstream industry at price m) and with retailer effort, e, to produce downstream revenues, $R(x, y, e; \theta)$.[4] Depending on the

[4]Strictly speaking, since there are other input suppliers, this is a multi-manufacturer setting. I will use the term "single manufacturer" to refer to markets in which there is only one strategic supplier. Markets in which y is oligopolistically supplied will be treated in Section 4.

case under consideration, θ may represent a parameter of market demand or a characteristic of the dealer that is not subject to dealer choice. Initially, the realization of θ is taken to be common knowledge between the dealer and the manufacturer. Later this assumption will be dropped to allow for uncertainty and for asymmetric information.

The variable e may be a measure of promotional effort, such as how hard the salesman works at convincing potential customers of the product's worth, that negatively affects the dealer's level of utility. Or e may be some decision variable, such as the color of the box in which the good is sold, that does not directly enter the retailer's utility function. In most of the analysis, I will treat x, y, e, and θ as scalars. In general, they may all be vectors, and most of the analysis extends trivially.

The dealer's utility, $U(I, e; \theta)$, depends on his income, I, his level of effort, e, and any retailer characteristics captured by θ. Let $W(x, y, e, \theta)$ denote the payment made by the dealer to the manufacturer. Under contract W, a type $= \theta$ dealer earns a net income of $R(x, y, e; \theta) - W(x, y, e, \theta)$ when he chooses input levels of x, y, and e.

Conditional on his accepting the contract, the dealer makes input choices to maximize his expected utility. Let $x^*(W, \theta)$, $y^*(W, \theta)$, and $e^*(W, \theta)$ denote the solution to

$$\underset{x, y, e}{\text{maximize}}\ U[R(x, y, e, \theta) - W(x, y, e, \theta) - my, e; \theta].$$

Given the agent's choices, the manufacturer earns:

$$W[x^*(W, \theta), y^*(W, \theta), e^*(W, \theta), \theta] - C[x^*(W, \theta)],$$

where $C(x)$ is the total cost of producing output x. For expositional simplicity, suppose that $C(x) = K + cx$, K and c constants.

To close the model, one must specify the contract bargaining game between the manufacturer and the dealer. In the principal–agent literature, the convention (which I will follow) is to assume that the *bargaining* is efficient and generates a (constrained) Pareto-efficient contract.[5] For convenience, rather than trace out the entire Pareto set, I will give all the power to the principal (i.e. the manufacturer) by letting him make take-it-or-leave-it contract offers. As long as the principal has some bargaining power and the principal and agent are symmetri-

[5] This assumption does not imply that the private first-best outcome is attained. Rather, it implies that given the set of feasible contracts and the information constraints faced by the two parties, there is no implementable contract that is Pareto superior to the equilibrium one.

cally informed, this assumption does not affect the qualitative nature of the equilibrium contract.[6]

The manufacturer would like to implement a contract that both induces the dealer to maximize the total expected profits of the two levels and gives all of these profits to the manufacturer. Let $x^{\pi}(\theta)$, $y^{\pi}(\theta)$, and $e^{\pi}(\theta)$ denote the solution to

$$\underset{F,\,x,\,y,\,e}{\text{maximize}}\; F - cx$$

$$\text{subject to}\quad U\left[R[x, y, e; \theta] - F - my, e; \theta\right] \geq 0,$$

where F is a scalar. The constraint in the problem above is the agent's individual rationality constraint; the agent accepts the contract offer only if the contract generates a higher utility level than the agent's next-best alternative, where the agent's opportunity utility is normalized to 0.[7]

One type of contract that induces the dealer to choose $x^{\pi}(\theta)$, $y^{\pi}(\theta)$, and $e^{\pi}(\theta)$ while transferring all rents to the manufacturer is the following:

$$W[x, y, e, \theta] = \begin{cases} G(\theta) & \text{if } \langle x, y, e\rangle = \langle x^{\pi}(\theta), y^{\pi}(\theta), e^{\pi}(\theta)\rangle, \\ G_0, & \text{otherwise,} \end{cases}$$

where $G(\theta)$ satisfies $U[R[x^{\pi}(\theta), y^{\pi}(\theta), e^{\pi}(\theta), \theta] - G(\theta) - my^{\pi}(\theta), e^{\pi}(\theta); \theta] = 0$ and G_0 is some constant that yields the agent less than his reservation utility level. Under this contract, the manufacturer orders the dealer to choose $x^{\pi}(\theta)$, $y^{\pi}(\theta)$, and $e^{\pi}(\theta)$, pays the agent his reservation wage in the event of compliance, and charges the agent a penalty otherwise.[8]

Except in special cases, such a contract would require a great deal of information on the manufacturer's part. In order to enforce the contract, the manufacturer would have to calculate the actions that the dealer should have taken and then verify that the dealer took them. To obtain this information, the manufacturer typically would have to engage in extensive and expensive monitoring of the dealer. Moreover, the need to provide complete instructions to the dealer could lead to tremendous contractual complexity and administrative cost. Hence, whether the two parties are independent corporations or are divisions within a single firm, such a contract is unlikely to be feasible or optimal.

[6]Alternative bargaining institutions are considered in the later discussion of an asymmetrically informed manufacturer–dealer pair.

[7]Cases in which the uncertainty is over the dealer's opportunity utility level can be represented by a parameterization of U with respect to θ.

[8]As noted in the introduction, many authors (in my opinion, mistakenly) implicitly attribute the ability to implement such a contract to a vertically integrated firm.

2.2. The use of the price system

2.2.1. Single-product pricing

The essence of the contract design problem is to overcome the externalities between the two stages. The actions of one party affect the profits of both, but the party at each stage makes his decisions based solely on the effects on his own profits or utility. In the version of the model introduced thus far, this problem takes the form of dealer moral hazard.

There is a very simple contract that can overcome dealer moral hazard. Under this contract, the dealer pays a fixed fee to the manufacturer and then buys the intermediate good at cost. Formally, $W[x] = F(\theta) + cx$, where $F(\theta)$ solves:

$$U\left[R(x^{\pi}(\theta), y^{\pi}(\theta), e^{\pi}(\theta); \theta) - cx^{\pi}(\theta) - my^{\pi}(\theta) - F(\theta), e^{\pi}(\theta); \theta\right] = 0.$$

As noted earlier, $F(\theta)$ often is called a franchise fee in the vertical restraints literature.

At the margin, the dealer appropriates any increase in the combined profits of the dealer and manufacturer brought about by the dealer's actions. Since he is the residual claimant to them, the dealer acts to maximize those profits (including the cost of effort in these calculations). This contract yields the manufacturer the maximal feasible level of profits given the agent's opportunity utility level.[9] Under this contract, there are no incentive problems, and there is no need for the manufacturer to monitor the dealer once the contract has been signed and the franchise fee paid.

If (a) the manufacturer can implement a two-part tariff, and (b) the manufacturer and the dealer have complete information about the state of the world, then the manufacturer can fully maximize profits without relying on either multiproduct pricing or direct vertical restraints even though dealer effort may be unobservable to the manufacturer.[10] Given the extent to which a two-part tariff economizes on contractual complexity and the need to monitor the agent, one would not expect to see the other forms of pricing or resale restraints in the setting analyzed above.

Both (a) and (b) may fail to be satisfied in practice. When there are multiple dealers, antitrust authorities might declare the use of fixed fees or other forms of

[9]This result is well known in both the principal–agent [e.g. Harris and Raviv (1979)] and price discrimination [e.g. Oi (1971)] literatures.

[10]The one role for a direct vertical restraint in this simple model is to keep the manufacturer from stealing customers from the dealer (i.e. competing in the downstream market) after having signed the contract.

nonuniform pricing to be illegal price discrimination. Arbitrage among dealers might also limit the use of quantity-dependent pricing. Of course, neither of these restrictions arises in the simple single-dealer setting under examination in this section, and it is hard to see why a manufacturer could not use a two-part tariff here. It is useful to consider the uniform-pricing case in this section, however, as a foundation for the multi-dealer analysis of the next one.

When the manufacturer is restricted to selling his output at a constant wholesale price [i.e. $W(x) \equiv wx$, w a constant], he must set the wholesale price greater than marginal cost in order to profit from the sale of his output. But faced with $w > c$, the dealer no longer acts to maximize joint profits.

In the presence of market power, the contract between the manufacturer and the dealer must both transfer income and make use of decentralized information (i.e. give the dealer incentives). When a two-part tariff is feasible, the manufacturer has two objectives and two instruments – the wholesale price can be set to induce the paper incentives and the franchise fee can be set to transfer rents. When the manufacturer is restricted to chosing $W(x) \equiv wx$, he has only one instrument – the wholesale price. In a sense, the two roles overload a single-parameter contract. If the wholesale price is either set above marginal cost to transfer income to the manufacturer or set below marginal cost to transfer income to the dealer, the dealer's incentives are distorted. Consequently, the possibility arises that utilizing other forms of pricing and vertical restraints would raise profits.

Even when two-part tariffs are feasible, the manufacturer may be unable to set a price schedule with the marginal price equal to marginal cost and the fixed fee chosen to extract all of the dealer's surplus. The manufacturer's setting such a schedule typically requires the manufacturer to make the value of the fixed fee contingent on the value of θ. In many situations, the manufacturer, the dealer, or the courts do have the information needed to write and/or enforce such contracts. In these instances, there are roles for other contractual provisions.

2.2.2. Incomplete information and pricing above cost

I will proceed by discussing several cases of informational incompleteness under which the manufacturer would find it privately optimal to set the marginal price above marginal cost, inducing distortions in the dealer's behavior. Having done this, I will explore the roles of vertical restraints in ameliorating these distortions.

To keep the exposition as simple as possible, I will assume that the wholesale pricing scheme can be no more complicated than a two-part tariff. The restriction to two-part tariffs is not essential to the qualitative results. If one were to allow a more general single-product, quantity-dependent pricing scheme, one still would

find points on the schedule at which the marginal price did not equal marginal cost and distortions would be induced.[11]

2.2.2.1. Risk sharing. In the presence of uncertainty and dealer risk aversion, the agency contract must play a third role: share risk (i.e. provide insurance for the dealer). Three roles turn out to be one too many for a single-product pricing scheme, such as two-part tariff.

Suppose that *neither* the manufacturer nor the dealer knows the value of θ at the *time of contracting*, although they have the same prior beliefs about its value. Moreover, suppose that, if the dealer accepts the contract, the dealer learns the value of θ before making any production decisions. The manufacturer and/or the courts never learn the value of θ, and thus neither F nor w can be made contingent on θ.[12]

If the manufacturer and dealer both are risk neutral, then the restriction that F not vary with the value of θ presents no problem. A simple two-part tariff with the wholesale price equal to marginal cost still can be used to decentralize the decision-making while fully transferring income. The only difference is that now the franchise fee must be set so that it yields the dealer an *expected* utility level (EU) equal to his opportunity level. When the dealer is risk averse, however, marginal cost pricing is not the manufacturer's optimal strategy. While a tariff with $w = c$ does a good job of generating dealer incentives, the retailer bears all of the risk – the manufacturer's profits are equal to F no matter what the value of θ. This pattern of risk sharing is not an efficient one.

Given his risk neutrality, if the manufacturer could observe the value of θ, he would set F contingent on that value so that the agent would bear no risk. This form of insurance completely protects the agent from bad realizations of θ while fully overcoming dealer moral hazard. When the manufacturer cannot base the vertical contract directly on the realization of θ, he has to rely on a signal of θ's value. Holmstrom (1979) and Shavell (1979) showed that the manufacturer may be able to take the dealer's choice of x as a signal on which to base insurance payments. This intuition comes most clearly when x is the sole input in the production of dealer revenues. In this case, express final market revenues as

[11] Rey and Tirole (1986) pointed out that two-part tariffs may be one form of quantity discounting that is particularly resistant to arbitrage across dealers. To assess a fixed fee, a manufacturer need monitor only whether any given dealer is using the input, not how much the dealer is *using* (which may differ from the quantity *purchased* directly from the manufacturer). Given that all dealers face the same wholesale price, there is no role for arbitrage among fee-paying dealers. Again, this is a multidealer argument that is made here in anticipation of the later analysis.

[12] In some cases, the principal could allow the agent's *report* of θ to influence the levels of w and F (i.e. the dealer would be offered a menu of contracts). I will not consider that possibility here. The analysis is similar to that of the informed manufacturer case below. In any event, if the manufacturer does use a menu of contracts to sort dealer types, equilibrium will entail $w(\theta) > c$ for some values of θ.

$R = R(x; \theta)$. Suppose that $R_\theta > 0$ for all θ and all $x > 0$, while $R_{x\theta} > 0$ for all θ and $x \geq 0$. Given any wholesale price, a low value of θ induces the dealer to choose a low level of x.

The manufacturer would like to compensate the dealer for low realizations of θ. Since θ and the dealer's choice of x are positively related, the manufacturer can do this by simultaneously lowering F and raising w in a way that raises the dealer's income for low unit-sales levels but lowers the dealer's income for high unit-sales levels. The adverse effects on the dealer's incentives limits the usefulness of this insurance, but some insurance is optimal.

To see this point formally, suppose that $w = c$. Consider the effects of the manufacturer's simultaneously raising w by $\mathrm{d}w$ and lowering F by $\mathrm{d}F$, where

$$\mathrm{d}EU = -\int_{\underline{\theta}}^{\bar{\theta}} U_I[I^*(\theta); \theta][x^*(W, \theta)\,\mathrm{d}w + \mathrm{d}F]h(\theta)\,\mathrm{d}\theta = 0,$$

$h(\cdot)$ is the density function for the distribution of θ over the support $[\underline{\theta}, \bar{\theta}]$, and $I^*(\theta) = R[x^*(W, \theta); \theta] - wx^*(W, \theta) - F$. The change in manufacturer's profits is

$$\mathrm{d}\pi = \int_{\underline{\theta}}^{\bar{\theta}} \{x^*(W, \theta)\,\mathrm{d}w + \mathrm{d}F\}h(\theta)\,\mathrm{d}\theta > 0,$$

by the concavity of $U(\cdot)$ and the monotonicity of $x^*(\cdot, \cdot)$ with respect to θ. Since $w = c$ by hypothesis, there is no first-order effect on manufacturer profits from the cutback in x induced by the rise in the wholesale price. If $w < c$, the cutback in x would raise manufacturer profits by even more.

2.2.2.2. An informed dealer. In many instances, the manufacturer and dealer are asymmetrically informed at the time of contracting. Suppose that it is common knowledge that the dealer is better informed about the value of θ than is the manufacturer. Consider the extreme case in which the dealer knows the realization of θ at the time of contracting, but the manufacturer never observes the value of θ.

The manufacturer may be concerned with the value of θ for two reasons. First, the level of compensation necessary to induce the agent not to seek alternative employment may depend on the value of θ. Second, when θ represents a dealer-specific characteristic, some types of dealer may be more valuable to the manufacturer than others.

Begin with the case of dealer-specific θ. For example, θ may be a measure of the dealer's ability or his disutility of effort (i.e. some types really hate to work, but others do not mind it much). Suppose that one is blessed with a market in which for any value of w both $x^*(W, \theta)$ and the maximal value of F such that

$$U[R(x^*, y^*, e^*; \theta) - wx^* - my^* - F, e^*; \theta] \geq 0$$

are increasing in θ. In this situation, if the manufacturer could choose the dealer's type, he would choose one with the maximal value of θ.

In situations of asymmetric information, the structure of the bargaining institutions utilized by the manufacturer and dealer can affect the nature of the equilibrium contract. Hence, I will examine the contracting process somewhat more fully here than in the rest of this chapter. Begin by considering a market in which there is a competitive supply of would-be dealers. Caves and Murphy (1976) have suggested that the manufacturer can use the franchise fee as a screening device. Since dealer surplus is increasing in θ, only those potential dealers with θ above some cut-off level would find it profitable to accept the franchise agreement. This scheme does not, however, use all of the information that is available to the manufacturer when there is more than one potential dealer.

In a different context, Demsetz (1968) has argued that an auction could serve to induce the parties (here, the potential dealers) to reveal their private information. For example, suppose that the manufacturer set up an auction of the following form. The ith would-be dealer submits a bid F_i. The manufacturer agrees to sell the input to the highest bidder at a wholesale price equal to marginal cost and a franchise fee equal to the *second*-highest bid received. Ties are broken at random.

Under this second-price auction, it is a dominant strategy for a type-$\tilde{\theta}$ dealer to bid $\tilde{F}$, where $\tilde{F}$ satisfies

$$\max_{x,\,y,\,e} U\left[R(x, y, e; \tilde{\theta}) - cx - my - F, e; \tilde{\theta}\right] = 0.$$

The unique Nash equilibrium in bids entails the manufacturer's getting all of the surplus that would be available from the agent with the second-highest value of θ in the population of potential dealers even though the manufacturer does not know the value of the franchise. Since $w = c$, the equilibrium contract overcomes any problems of dealer moral hazard. Moreover, the potential dealer with the highest realization of θ always is selected. Loosely speaking, if the number of bidders is large relative to the variance of θ, then this process tends to appropriate most of the downstream profits. If θ is a market-specific parameter, rather than a dealer-specific one, the scheme appropriates dealer profits completely.[13]

With a single potential dealer, the manufacturer's dealer selection problem is simply whether to have a dealer or not. The manufacturer could decide that it would rather shut down than have a bad dealer, but as long as F and $w - c$ are

[13] When the value of θ need not be constant across dealers, the second-price auction presented in the text may not be privately optimal form of auction to run. There is now a large literature on optimal auction design. See, for example, Milgrom and Weber (1982).

non-negative, there is no reason to do so.[14] The reason the manufacturer cares about the dealer's type is that it determines the compensation necessary to attract the dealer. One approach is to set $w = c$ and set F sufficiently low that even the worst type of dealer would receive at least his opportunity utility level. Such a scheme would, however, overcompensate high-θ dealers. On the other hand, higher values of F would risk turning away the dealer for low realizations of θ.

When there is only a single potential bidder, competitive pressure cannot be used to induce the dealer to reveal his private information about θ. This information must be elicited through either the bargaining process or the contract itself. If multiple rounds of bargaining are possible, the manufacturer might offer a contract and see whether it was rejected. In the event of a rejection, one side would make a counteroffer. When the dealer's costs of bargaining vary systematically with the value of θ, the manufacturer may be able to take the dealer's bargaining behavior as a signal of the underlying value of θ. For example, if both the value of being a dealer and the costs of bargaining are increasing functions of θ [as in Rubinstein (1985)], then the manufacturer may be able to infer that a dealer has a high value of θ from the fact that he gives in during an early round of bargaining. When the costs of bargaining do not arise from delay, however, there may be little reason to expect the needed correlation to hold.

The contract itself may provide a mechanism by which to elicit the dealer's private information. Suppose that the manufacturer makes a single take-it-or-leave-it contract offer. As in the case of symmetric uncertainty, the dealer may signal his value of θ through his choice of x. To see how such signalling might work, again suppose that x is the sole input and that for any given contract a high-θ dealer would purchase more of the input than would a low-θ dealer. It is optimal for the manufacturer to set $w > c$ to use the relationship between x^* and θ to extract rents from high-θ dealers without driving low-θ dealers from the market.

Suppose, counterfactually, that the profit-maximizing two-part tariff entails $w = c$ and $F = F_0$. There must exist some θ_0 such that $h(\theta_0) > 0$ and

$$\underset{x}{\text{maximum}}\ U\left[R(x;\theta_0) - F_0 - cx;\theta_0\right] = 0,$$

since otherwise the manufacturer could increase profits by raising F.

Now raise w and lower F such that $x^*(W_0,\theta_0)\,\mathrm{d}w = -\mathrm{d}F > 0$. For infinitesimal changes, there is no change in the set of dealer types who accept the contract. The change in the manufacturer's profits is

$$\mathrm{d}\pi = \int_{\underline{\theta}}^{\bar{\theta}} \left\{x^*(W_0,\theta)\,\mathrm{d}w + \mathrm{d}F\right\} h(\theta)\,\mathrm{d}\theta > 0.$$

[14] When there are multiple dealers and reputation effects arise, this conclusion will need to be modified. See Section 3 below.

The inequality follows from the fact that $x^*(W_0, \theta) > x^*(W_0, \theta_0)$ for all $\theta > \theta_0$. Therefore, the privately optimal contract entails a wholesale price that is greater than the input's marginal cost of production.

2.2.2.3. An informed manufacturer. Suppose that θ is a parameterization of a market characteristic (e.g. the demand for the downstream output or the unpleasantness of dealer's job) that is better known by the manufacturer than by the dealer at the time of contracting. If it is common knowledge that the manufacturer knows the value of the intermediate good better than does the dealer, then for a range of θ's the privately optimal wholesale price may be greater than the marginal cost.

Consider the following illustrative example. Both R and R_x are increasing in θ. θ takes one of two values, $\underline{\theta}$ or $\bar{\theta}$, where $\underline{\theta} < \bar{\theta}$. The manufacturer observes the realization of θ at the outset but θ never is observable to the courts, so a contract contingent upon θ is unenforceable. The dealer learns θ after signing the contract but before making production decisions. x is the sole input in the production of downstream revenues. Finally, θ does not enter the dealer's utility function directly.[15]

Consider the following outcome. A manufacturer with a low-value franchise offers $\underline{W}(x) = \underline{F} + cx$, where $\underline{F}$ satisfies

$$U\left[R^*(\underline{W}, \underline{\theta}) - \underline{F} - cx^*(\underline{W}, \underline{\theta})\right] = 0.$$

A manufacturer with a high-value franchise offers $\overline{W}(x) = \overline{F} + \overline{w}x$, where $\overline{F}$ and $\overline{w}$ are the solution to the following problem:

$$\underset{w,\,F}{\text{maximize}}\ F + (w - c)x^*(W, \bar{\theta})$$

$$\text{subject to}\quad U\left[R^*(W, \bar{\theta}) - F - wx^*(W, \bar{\theta})\right] \geq 0 \tag{1}$$

and

$$F + (w - c)x^*(W, \underline{\theta}) \leq \underline{F}. \tag{2}$$

Given that R and R_x are increasing in θ, x^* is increasing in θ and $\underline{W}$ satisfies constraint (2) with equality while (1) does not bind. Moreover, $\underline{W}$ is more profitable than any other contract with $w \leq c$. Starting at $\underline{W}$, consider the effects of raising the wholesale price by dw and changing the franchise fee by $dF =$

[15] Equilibria with $w > c$ also could arise if θ affected the dealer's utility function rather than his revenue function. For example, they would arise if dealer effort were an input to production of downstream revenues, $U_\theta \geq 0$, $U_{\theta e} > 0$, $R_{xe} > 0$, and R were independent of θ.

$-x^*(c, \underline{\theta})\,\mathrm{d}w$. The constraints still would be satisfied and manufacturer profits would rise. Therefore, the solution to this problem entails wholesale pricing above marginal cost.

The manufacturer's private information is fully revealed to the dealer under this outcome. A manufacturer with a high-value franchise signals this fact by accepting a combination of wholesale price and license fee that he would find unattractive if he actually had a low-value franchise – by setting $w > c$, the manufacturer gives himself a stake in the dealer's volume and makes credible his claim that his dealer will have a high-volume, high-revenue operation. A manufacturer with a low-value franchise "admits" this fact and offers an efficient contract (i.e. $\underline{w} = c$).

This separating outcome is an equilibrium if dealer's beliefs are as follows. If $w = c$ and $F = \underline{F}$, or if $F + (w - c)x^*(W, \underline{\theta}) > \underline{F}$, then the dealer believes that the offer has been made by a low-θ manufacturer. Otherwise, the dealer infers that the manufacturer has a high value of θ. These beliefs are confirmed along the equilibrium path.[16]

To see that this outcome is an equilibrium, consider a deviation by the manufacturer. Clearly, any deviation by a low-θ manufacturer would be unprofitable unless it "fooled" the dealer. And, by construction, a low-θ manufacturer would not find it profitable to mimic a high-θ manufacturer. Also by construction, the strategy played under the candidate equilibrium by a high-θ manufacturer maximizes his profits subject to the constraint that the dealer believes him to have a high-value franchise. Suppose, then, that the deviation induces the dealer to believe that the high-θ manufacturer actually is a low-θ manufacturer. The greatest profits that such a deviation could generate are given by the solution to

$$\begin{aligned} &\underset{w,\,F}{\text{maximize}}\ F + (w - c)x^*(W, \bar{\theta}) \\ &\text{subject to}\quad U\left[R^*(W, \underline{\theta}) - F - wx^*(W, \underline{\theta})\right] \geq 0. \end{aligned} \tag{3}$$

Constraint (3) can be written as

$$U\left\{R^*(W, \underline{\theta}) - cx^*(W, \underline{\theta}) - \left[F + (w - c)x^*(W, \underline{\theta})\right]\right\} \geq 0.$$

Recall that $\underline{F}$ satisfies:

$$U\left[R^*(\underline{W}, \underline{\theta}) - cx^*(\underline{W}, \underline{\theta}) - \underline{F}\right] = 0,$$

[16] I am not pretending to offer a complete analysis of the equilibria of this game. For a general discussion of the type of equilibrium selection arguments that support the primacy of this equilibrium see Banks and Sobel (1987).

and hence constraint (3) implies that

$$R^*(W, \underline{\theta}) - cx^*(W, \underline{\theta}) - \left[F + (w - c)x^*(W, \underline{\theta})\right]$$
$$\geq R^*(\underline{W}, \underline{\theta}) - cx^*(\underline{W}, \underline{\theta}) - \underline{F}.$$

It follows that $F + (w - c)x^*(W, \underline{\theta}) \leq \underline{F}$, with strict inequality for any $w \neq c$. That is, constraint (3) is stronger than (2). Moreover, it is evident that constraint (3) is stricter than (1) as well. Therefore, there is no profitable deviation from the candidate strategy given the dealer's beliefs.

2.2.3. Manufacturer moral hazard

Up to this point, the dealer has been the only party to make decisions after the contract has been signed. The manufacturer's sole action has been to provide as much of the intermediate good the dealer demands. Typically, there are other dimensions to manufacturer behavior. For example, the manufacturer may be able to affect the quality of his output. Or, the manufacturer may provide promotional support of sales of the *downstream* product. These activities may be such that explicit contracting is infeasible either due to problems of measurement and observability, or due to the need to vary the levels of these activities in response to market conditions that are themselves unobservable to the courts.

If the manufacturer's post-contracting effort affects downstream market revenues and explicit contracting over this effort is infeasible, then the privately optimal wholesale price may again be greater than marginal cost. If $w \leq c$, the manufacturer has no incentive to encourage final sales. Hence, setting $w > c$ is the only way to get the manufacturer to provide effort. If the level of manufacturer support is sufficiently important to downstream sales, the benefits from the increase in manufacturer incentives will outweigh the negative effects from the distortion in dealer incentives that raising the wholesale price above marginal cost induces.

2.3. The need for more sophisticated contracts

For a variety of reasons, the privately optimal wholesale price may differ from marginal cost. A divergence between the wholesale price and marginal cost levels distorts the dealer's incentives. In such cases, other forms of pricing and direct vertical restraints may be useful means of ameliorating these distortions.

2.3.1. Multi-product pricing: Ties and royalties

As Burstein (1960b) observed, the manufacturer faces a taxation problem: he would like to collect as much revenue as possible from the dealer. His ability to

do so is limited by the dealer's individual rationality constraint. By tying inputs, the manufacturer can face the dealer with the following decision. Either the dealer accepts the manufacturer's set of prices (i.e. his set of commodity taxes) on the set of tied inputs, or the dealer faces the market price vector for the inputs produced by other firms and an infinite price for the manufacturer's input. As Burstein put it, the dealer is confronted with an "all-or-nothing choice" of prices. The dealer makes this choice by comparing the two resulting levels of utility.

Given the form of the dealer's individual rationality constraint, when there is no uncertainty about θ and no problem of manufacturer moral hazard, it is in the interest of the manufacturer to raise revenues through an efficient set of taxes. Hence, the earlier result that the firm would like to set $w = c$ and collect the revenue through the use of a nondistorting fixed fee. If franchise fees are illegal and the manufacturer relies on single-product pricing, he must set the wholesale price greater than marginal cost in order to extract downstream surplus. Such pricing distorts dealer behavior and limits the extent to which the manufacturer can extract downstream profits. As a general rule, the greater the number of commodities that the manufacturer can subject to taxation, the more efficiently revenues can be raised and the greater are manufacturer profits.[17] As Burstein pointed out, this relationship can hold even when the demands for the tied goods are *completely unrelated*.

It is useful to think of the distortions induced by $w > c$ in terms of derived demand. Wholesale pricing above cost induces the dealer to purchase too little of the manufacturer's output for two reasons. First, the dealer engages in input substitution away from x. Second, in response to higher costs, the dealer may contract his output and thus reduce the quantity of all inputs simultaneously. In those cases in which final output is unambiguously defined (i.e. the dealer cannot affect the final product quality), the second problem can be viewed in terms of pricing the final output – the dealer takes w as the marginal cost in computing the final price, and thus sets the final price too high when $w > c$. This practice is known in the literature as "double marginalization".[18]

To prevent distortions in relative input levels, the manufacturer would like to mandate the use of inputs in efficient proportions. Barring this power, the manufacturer may be able to induce the dealer to choose an efficient input mix through the use of a tying arrangement. If all inputs can be subject to taxation (i.e. tied), the manufacturer can mark up the prices of all inputs proportionally. Let n denote the price of input y when it is tied with x. By settiing prices such that $w/n = c/m$, the manufacturer can induce the dealer to choose an efficient

[17]This principle is stated, and its implications for business behavior are more fully drawn out, in Burstein (1960a).

[18]Spengler (1950) appears to have first identified the double marginalization effect in the context of vertical control. He pointed out that if one takes the unnatural union view of vertical integration, then the elimination of the double mark-up is a benefit of vertical integration when only single-product, uniform pricing is allowed.

mix of x and y (recall that m is the competitive supply price of input y). A tying arrangement works perfectly to overcome input mix distortions when it can be applied to all inputs. When some or all of the other input levels (e.g. dealer effort) are unobservable to either the manufacturer or the courts, it is impossible to implement a comprehensive tie-in, and input distortions may be induced. Loosely speaking, these distortions can be minimized by increasing the prices of those tied inputs that are the least substitutable for untied inputs.[19]

Turn now to the distortion in level of final output, and suppose that no input substitution is possible. Given this assumption, for any input, i, there is a function $f_i(\cdot)$ such that the firm cannot produce z units of output unless $f_i(z)$ units of input i are utilized. If all inputs must be used in proportions that vary only with the level of final output, then any multi-product pricing scheme is equivalent to some (possibly nonuniform) single-product price schedule.

Ties need not be equivalent to single-product price schedules when input substitution is feasible. Suppose, for example, that the manufacturer supplies two inputs, one fixed and the other variable. Moreover, suppose that the manufacturer has a monopoly on the variable factor but faces competition in the supply of the fixed factor. In this setting, a tie may serve as an implicit two-part tariff. The manufacturer prices the variable input at marginal cost and uses a mark-up on the fixed input as a franchise fee. As with a two-part tariff, the manufacturer is able to transfer rents in a way that does not distort the dealer's marginal cost function. Hence, the final output level is unaffected. Of course, when there is no perfectly fixed input, the manufacturer may find it optional to implement a multi-product tie.

When the incompleteness of information is the reason that $w > c$ in the first place, the manufacturer's intermediate good pricing problem is an extremely complex one. For example, suppose that the manufacturer cannot observe the dealer's value of θ and that one factor is fixed and the other variable. For incentive reasons (i.e. to minimize the input distortion) the manufacturer should raise the price of the fixed input to collect revenue. But in order to obtain a signal of θ, it may be necessary to mark up the price of the variable input.[20] In general, the manufacturer will find it profitable to tie the purchase of y to the purchase of x and sell both inputs at prices that deviate from their production cost.[21]

[19] For a more complete discussion of these issues, the reader is referred to the literature on optimal commodity taxation. See, for example, Atkinson and Stiglitz (1980).

[20] In the vertical restraints literature, this practice is known as metering. Bowman (1957) and Telser (1965) credit Aaron Director with first having suggested that businesses use tying to facilitate metering. For a brief mention of this theory, see Director and Levi (1956).

[21] If input y is not competitively supplied, there may be another role for ties. Suppose that inputs x and y are complementary in the sense that lowering the price of x raises the demand for y, and thus may raise the profits of the firms producing y. The producer of x does not count the increase in other firms' sales as a benefit when calculating its optimal price for x. When the two inputs are tied, the single seller internalizes the externality across inputs. See Bowman (1957) for an early discussion of this point.

Royalties are another form of multi-product pricing that may be available to the manufacturer. Under a revenue royalty scheme, the manufacturer receives some fraction of the dealer's final market revenues. A royalty can be thought of as a form of multi-product pricing because it is equivalent to a proportional mark-up on all inputs. Hence, royalties can be used like ties to alleviate input distortions. Formally, let the dealer's revenues net of payments to the manufacturer be denoted by $T[R]$. Suppressing θ for notational convenience, the dealer

$$\underset{x,\,y,\,e}{\text{maximizes}}\; U\{T[R(x, y, e)] - cx - my, e\},$$

which leads to an efficient input mix. Likewise, in cases in which dealer output is well defined and observable, a royalty scheme has the advantage over a tie that the manufacturer does not have to monitor the levels of other downstream inputs (e.g. dealer effort) in order to limit input distortions. Of course, the imposition of a royalty scheme still distorts the dealer's overall incentives to produce revenues.

The use of royalties also can increase the manufacturer's profits when the manufacturer and dealer are asymmetrically informed about some parameter, θ, and the final price is more sensitive to the value of θ than is the level of the dealer's input use. Suppose, for example, that the demand for final output is given by

$$D(p; \theta) = \begin{cases} D^0 & \text{if } p \leq \theta, \\ 0, & \text{if } p > \theta, \end{cases}$$

and that only the dealer observes the realization of θ. Then x alone is of no value as a signal of θ, but clearly observing R (or the final price) is. Thus, a revenue-based royalty can be one way of extracting the rents that might otherwise accrue to a dealer with a high realization of θ.

There can be important benefits from using both ties and royalties simultaneously. A contract under which all inputs are priced below cost by the same percentage as the revenue royalty rate is equivalent to a pure profits tax. In this case, there is no distortion in the dealer's behavior. Unfortunately (from the manufacturer's viewpoint), it often is impossible to subsidize dealer effort.

2.3.2. Resale restraints

As noted above, there are two reasons for which the dealer's derived demand curve for x is downward sloping: (a) he responds to an increase in his marginal costs by reducing the total amount of output sold; and (b) he responds to an increase in the cost of a single input by substituting other inputs for x in the production of final output. A resale price ceiling appears to be the obvious solution to the problem posed by (a) since the dealer restricts his output by

raising the final price of the good. If the manufacturer places the ceiling on the final price, this output restriction is constrained (as long as the ceiling does not lead to the dealer's shutting down).

Suppose that, for any fixed two-part tariff, dealer profit and output both are increasing in θ. Then, if the optimal two-part tariff involves a wholesale price that is greater than marginal cost and the dealer cannot affect the final product quality, a two-part tariff coupled with a resale price ceiling is privately superior to a two-part tariff alone. If the dealer type charging the highest final price is not one for whom the individual rationality constraint is binding, then imposing a ceiling on the final price raises unit sales (and manufacturer profit) without inducing any dealer to drop out of the market. If the dealer type charging the highest price is one for whom the individual rationality constraint is binding, then the manufacturer should set the ceiling price equal to the highest current price, raise w by $\mathrm{d}w$, and lower F by $\mathrm{d}F = -x^*(w, \underline{\theta})\,\mathrm{d}w$, where $\underline{\theta}$ denotes the lowest dealer type. This shift has no effect on either the type-$\underline{\theta}$ dealer's profit or the profit that the manufacturer derives from this type of dealer. It can be shown, however, that the manufacturer's profit derived from all other types of dealer increase.

While the use of a price ceiling can attenuate the dealer's restriction of final output, resale price maintenance does not work perfectly in cases in which there is uncertainty or asymmetric information. If θ is uncontractable, then it is impossible to set $p = p^m(\theta)$ for all θ.[22] In the single-dealer case, this failing of resale price maintenance is an important one because uncertainty or asymmetric information are the motivations for resale price maintenance when two-part tariffs are feasible.

Resale price maintenance also fails to solve the input substitution problem; if input substitution is feasible, the dealer tends to use too little of x relative to the other inputs whether or not a price ceiling is imposed. A related problem concerns the definition of a resale price ceiling. For what good is the ceiling price set? In order to establish the profitability of a price ceiling, it was assumed above that the dealer could not affect the quality of the final product. Typically, the dealer can affect product quality (e.g. the dealer can provide varying levels of service to final consumers). The dealer may respond to a price ceiling by reducing the quality of his output.[23] Essentially, this is another form of input mix

[22] This analysis is a special case of Rey and Tirole's (1986) examination of the resale price maintenance in a multi-dealer setting.

[23] Let $D(p,q)$ denote the final demand function, and let $C[D(p, q), q]$ denote the dealer's total cost function. The dealer's profit-maximizing choice of product quality is an increasing function of the price he charges under the following conditions: $D_1 < 0$, $D_2 > 0$, $D_{12} \geq 0$, $C_1 > 0$, $C_2 > 0$, $C_{11} \geq 0$, and $C_{12} \geq 0$.

distortion. To complicate matters, this effect occurs on top of the input mix distortion induced by the nonmarginal-cost pricing of the manufacturer's output.

Turning to nonprice restrictions, customer restraints may be utilized to extract rents from the dealer when the manufacturer and dealer are imperfectly informed at the time of contracting. Suppose that θ effects only the market revenue function $R(x, y, e; \theta)$. If the variation in R is due to uncertainty with respect to a particular class of customers, the manufacturer may reserve that class for his own sales force. Of course, some form of contract is needed between the manufacturer and his own sales force, and contracting problems may arise in this relationship as well. It makes sense to implement this division of customers only if there is some reason that the manufacturer is better able to monitor his own sales force (perhaps that is what is meant by the sales force being "his own"), or if his own sales force need be given less discretion, so that problems of moral hazard are less severe. This type of customer restraint is an example of what is known more generally in the literature as *tapered* vertical integration , where one firm is present at a vertical stage as both a buyer and a seller. Like all theories of vertical integration, this one needs further development.

3. Private incentives: A single manufacturer with many dealers

Most manufacturers sell their output to more than one dealer. In the case of a manufacturer with a single dealer, correcting the externalities between the manufacturer and the dealer is the essence of the contracting problem. These externalities, and the need to correct for them, remain present when the manufacturer has multiple dealers. There is a new type of externality as well – externalities across the dealers. In making his decisions, each dealer ignores the effects that his actions have on the profits of the other dealers. Dealers thus may fail to maximize joint profits, opening up a potential role for vertical restraints even in those situations in which the manufacturer is otherwise able to extract all of the existing surplus from the dealers.[24]

While the presence of multiple dealers opens up new control problems for the manufacturer, it also opens up a new control instrument. With multiple dealers, the manufacturer can implement a relative performance scheme under which a dealer's compensation depends not just on his performance, but also on the performance of other dealers. And, of course, when there are multiple dealers, the

[24] There is another new effect that can arise in the multiple dealer case. Public policy may force the manufacturer to offer the same contract to all dealers. In such cases, the manufacturer must essentially act as if he does not know the type of any given dealer. The distortions induced by two-part pricing and the possible remedies are similar to those discussed in the case of a single dealer. In fact, if the dealers serve isolated markets, the two problems are isomorphic to one another.

assignment of exclusive territories by the manufacturer no longer is a trivial policy.

3.1. Externalities across dealers

It sometimes is argued that the manufacturer's profits are greatest when his dealers are perfect competitors with one another. If no input substitution is possible and the dealers simply choose their output levels for a good of fixed quality, then this argument is correct. Under perfect competition, the difference between the final product price, p, and the wholesale price, w, is as low as is compatible with non-negative dealer profits. Hence, under downstream perfect competition, for any given level of w, manufacturer profits, $X(p)\{w - c\}$, are as large as possible [here, $X(p)$ denotes the dealers' aggregate demand for the manufacturer's output]. By competing with one another, the dealers drive down their margins and allow the manufacturer to appropriate all of the profits from production and sale of the good.

When dealers can do more than simply choose a single uniform price for their output, the manufacturer need not favor strong intrabrand competition. Intrabrand competition may inhibit certain aspects of dealer behavior that would increase manufacturer profits. In particular, dealer competition may constrain certain forms of final good pricing and limit promotional activities. In these situations, resale restraints may be imposed to obtain outcomes that are not supportable as competitive equilibria in the final market.

3.1.1. Price competition as a negative externality

Consider first the effects of dealer competition on downstream pricing. In the presence of strong intrabrand competition, dealers are forced to charge each consumer a price based on the cost of supplying that consumer. Both high price–cost margins and price discrimination against final consumers are incompatible with vigorous intrabrand competition. Yet, these may be profit-maximizing forms of pricing.

3.1.1.1. The erosion of the price–cost margin.[25] The manufacturer would like to induce the monopoly price, $p^m(c; \theta)$. Faced with wholesale price w, competitive retailers set price $p^c(w; \theta) < p^m(w; \theta)$. Relying solely on the price mechanism,

[25]As shown by the analysis in Section 2, if dealers are imperfect competitors and either franchise fees are infeasible or informational incompleteness leads to wholesale pricing above marginal cost, then the manufacturer may find that the dealers' price–cost margin is too high. In the present section, I focus solely on the new problem of an unprofitably low margin that arises when there are competing dealers.

the only way for the manufacturer to induce $p^{\pi}(c; \theta)$ is to raise the wholesale price above marginal cost. As long as x is a variable input and is essential to downstream production [i.e. $R(0, y, e; \theta) = 0$ for all y, e, and θ], there exists some $w > c$ such that $p^{c}(w; \theta) = p^{\pi}(c; \theta)$. In the absence of any input substitution possibilities, this wholesale price induces the joint-profit-maximizing downstream equilibrium, and the manufacturer can extract any dealer surplus (or cover any dealer losses) through the appropriate choice of the franchise fee.[26] When input substitution is possible, setting $w > c$ distorts the dealer's behavior. Hence, there may be private gains from the use of multi-product pricing or resale restraints.

As many authors have noted [e.g. Dixit (1983) and Mathewson and Winter (1984)], if all parties know θ at the time of contracting, the manufacturer can impose a resale price *floor* equal to $p^{\pi}(c; \theta)$ while setting $w = c$ and using F to appropriate any dealer surplus.[27] Note, however, that if the manufacturer does not know the value of θ at the time of contracting, he may be unable to set the price floor equal to $p^{\pi}(c; \theta)$ for all realizations of θ.

Another way to prevent the erosion of the dealer margin is to eliminate intrabrand competition through the assignment of exclusive territories. The analysis presented in Section 2 can be interpreted as showing that if the manufacturer and dealers all have complete information and the manufacturer assigns exclusive territories, then the manufacturer's profits are maximized by setting $w = c$ and raising F until each dealer just earns his opportunity utility level. Of course, as was discussed in Section 2, the assignment of exclusive territories alone need not be sufficient to attain the perfect extraction of dealer rents when the parties have incomplete information.

3.1.1.2. Price discrimination. Intrabrand competition may reduce overall industry profits (and thus the amount of rents that the manufacturer can extract) by blocking price discrimination in the final market. A simple example illustrates this point. Suppose that consumers are located in two identical cities, A and B. There are N_1 consumers in each city who demand either one or no unit of the good with reservation price $\bar{p}_1$. There are N_2 consumers in each city who demand either one or no unit of the good with reservation price $\bar{p}_2$, where $\bar{p}_1 < \bar{p}_2$. There are two dealers, each of whom is able to identify the type of any given customer. One dealer is located in city A, and one is located in city B. There are no transportation costs between the two cities. The dealers are Bertrand competitors; each dealer chooses his final good prices under the assumption that the prices of his rival are fixed.

[26] Recall that negative franchise fees are allowed.

[27] If lump-sum payments to dealers are infeasible, then in some instances the manufacturer may use a resale price floor to preserve the downstream margin as a means of transferring income to the dealers.

The manufacturer provides the sole input into downstream production and must choose a contract for his dealers. The manufacturer would like to have his dealers charge $\bar{p}_i$ to type-i consumers, $i = 1, 2$, and use the fixed fee charged to the dealers to appropriate all of the profits. But when the contract is simply a two-part tariff, the dealers drive the final price down to the wholesale price level for *all* consumers, and there is no price discrimination. The manufacturer must choose between setting $w = \bar{p}_1$ and earning profits of $2(\bar{p}_1 - c)(N_1 + N_2)$, or setting $w = \bar{p}_2$ and earning profits of $2(\bar{p}_2 - c)N_2$.

The manufacturer can facilitate price discrimination by placing territorial customer resale restraints on his dealers. If the manufacturer assigns exclusive territories to the two dealers, then each dealer is a downstream monopolist within his home city. Given any $w \leq \bar{p}_1$, each dealer charges $\bar{p}_1$ to the N_1 consumers with low reservation prices and $\bar{p}_2$ to the N_2 consumers with high reservation prices, yielding profits of $(\bar{p}_1 - w)N_1 + (\bar{p}_2 - w)N_2 - F$. The manufacturer can set F to drive dealer profits to zero. The resulting manufacturer profits are $(2\bar{p}_1 - c)N_1 + 2(\bar{p}_2 - c)N_2$, which are greater than the profits under nonexclusive territories.

Somewhat more generally, if two-part tariffs are available to the manufacturer, dealers are risk netural, the manufacturer and dealers are symmetrically informed ex ante, and there is no problem of manufacturer moral hazard, then the assignment of exclusive territories yields weakly greater profits than does dealer competition.[28] The reason for this result is as follows. Under either downstream monopoly or competition, the manufacturer appropriates all of the profits through his choice of the fixed fee. Under competition, the manufacturer can induce any nondiscriminatory pricing scheme, although $w > c$ results in an inefficient input mix. Under exclusive territories, the manufacturer can set $w = c$ to induce profit-maximizing discriminatory scheme without inducing an inefficient dealer input mix. The assignment of exclusive territories yields strictly higher profits than does competition whenever a final good monopolist would find it profitable to discriminate.

Other resale restraints also may serve to facilitate downstream price discrimination. Suppose that, in the example above, the customers with low reservation prices correspond to national accounts. By imposing a resale restraint based on customer class, the manufacturer can set $w = \bar{p}_2$ to induce a price of $\bar{p}_2$ for sales made to local accounts through his dealers, while selling at $\bar{p}_1$ to the

[28]This result does not generalize to markets in which franchise fees are infeasible. If the manufacturer is restricted to choosing a uniform wholesale price, then he weakly prefers perfect competition downstream to a set of local monopolists who act as third-degree price discriminators. The reason is as follows. The manufacturer's profits are $w - c$ times the quantity sold and, for any given uniform price, demand is greater when the dealers are perfect competitors than when they are local monopolists.

national accounts that he has reserved for himself. Of course, in the simple example considered here, the price charged by the manufacturer is lower than w, so that explicit restrictions on dealer sales to national accounts are unneeded. In the presence of input substitution possibilities or imperfect competition among the dealers serving high-value consumers, however, the manufacturer might have incentives to set $w < p_1$, and explicit customer restraints would be needed.

Bowman (1955) noted that a common argument of his day was that resale price maintenance could be used to facilitate price discrimination by preventing the dealers from making sales to low-price customers. The manufacturer's relative ability to monitor his dealers' prices versus their customer lists would be an important consideration in choosing between the two types of resale restraint to support price discrimination.

The discussion so far has not addressed the question: How do the manufacturer and his dealers identify the different customer types in order to price discriminate? In some cases, the manufacturer and his dealers may discriminate among final consumers through the use of a sorting mechanism. Bolton and Bonnano (1987) analyzed a market in which consumers differ in their tastes for some service that can be viewed as product quality. By offering different variants of the good, the firms can induce consumers to reveal their types through their purchase decisions, facilitating a greater extraction of consumer surplus. Suppose, for example, that the costs of quality and distribution of consumer tastes are such that industry profits are maximized when two variants of the final good are offered, one with high quality and one with low.

Can the manufacturer induce this outcome given that each dealer is free to choose the single quality that he offers? For simplicity, assume that there are two dealers. Bolton and Bonnano found that, for sufficiently low values of the wholesale price, the dealers do choose to produce different variants of the good. The dealers differentiate themselves vertically in order to relax price competition.[29] For these low values of the wholesale price, dealer competition leads to downstream prices that are below their joint-profit-maximizing levels. If the manufacturer raises the wholesale price, however, dealer differentiation and the resultant consumer sorting are eliminated. Intuitively, if the wholesale price is set too high, a low-quality dealer finds himself undercut by the high-quality dealer. Two-part pricing alone cannot induce the joint-profit-maximizing outcome.

One way to preserve the dealer's price–cost margins is to impose a resale price floor (which is independent of dealer quality). But when price competition is limited, the dealers turn to quality competition. Again, both dealers offer high-

[29]As Bolton and Bonanno (1987) note, this type of result is a well-known one in the literature on vertically differentiated oligopoly [see, for example, Shaked and Sutton (1982)].

quality variants and the dealer variety is eliminated. While the imposition of a resale price restraint can raise profits in this setting, it does not work perfectly.

Bolton and Bonnano went on to show that a more sophisticated contract in which any given dealer's franchise fee is a function of all of the prices charged in the downstream market can work perfectly in their simple setting. As they noted, however, such a contract might be very costly to enforce when the number of dealers is large.[30]

In some instances, the cost of selling to a given customer may depend upon characteristics of that customer. In such a market, price discrimination may take the form of charging equal prices to two consumers even though the firm incurs different costs to serve them. As Caves (1986) discussed, resale price maintenance may facilitate this form of price discrimination by forcing dealers to charge uniform prices. Hence, resale price maintenance may provide a means of discriminating against consumers who desire low levels of promotional activities or service and thus are cheaper to serve.

So far, I have discussed cases in which the manufacturer would like to induce prices that discriminate against final consumers within a single geographic region. The manufacturer also might want to practice price discrimination across geographic regions. As White (1981) observed, even if dealers are located in different regions, varying the wholesale price to dealers in each region may be insufficient to achieve price discrimination. When the dealers are perfect competitors who can costlessly arbitrage across regions, the price in all regions is driven to "just under" the second lowest value of w_i.[31] Similar effects arise even when arbitrage across regions is costly (i.e. dealers are spatially differentiated), although clearly the arbitrage problem is less severe.

The imposition of (enforceable) territorial restrictions can be used to prevent arbitrage and allow geographic price discrimination. Note that if each dealer has an exclusive territory, then in the absence of other incentive or informational problems, the manufacturer can induce optimal discrimination by setting all of the wholesale prices equal to marginal cost. Obviously, the use of a system of resale price maintenance under which the maintained price varied across regions also would achieve the desired downstream geographic discrimination.

3.1.2. Dealer free riding

3.1.2.1. Tangible presale services. As just shown, the manufacturer may implement vertical restraints when intrabrand competition limits certain forms of pricing. Similarly, vertical restraints may be profitable when the competitive

[30] In order to enforce the contract, each dealer would have to observe the prices charged by all other dealers.

[31] In the event that two or more firms tie for having the lowest price, these firms price at the common value of w.

equilibrium fails to support promotional activities at the level that would maximize joint profits.

Suppose that there are two Ford dealers in some city. If one of the dealers runs a series of television commercials promoting Ford automobiles, these ads benefit both the dealer paying for the ads and the other dealer, who is going along for a free ride. The advertising dealer's incentives to run the commercials are smaller than the joint-profit-maximizing incentives for two reasons: (1) in choosing his level of advertising, a dealer ignores any increase in profits that his rival enjoys as a result of advertising spillovers; and (2) given the rivalry between the two dealers, the spillovers may reduce the profits of the advertising firm, further diminishing incentives. This sort of argument can be applied to other types of promotional and service activities, such as operating showrooms and having salesmen provide pre-sale information to consumers.

The manufacturer could rely on the price system to restore some of the lost incentives. One way, using single-product pricing, is to set the wholesale price below marginal cost in order to make the additional sales generated by promotional activities more valuable to the dealers. But such a two-part tariff would lead to other incentive problems, and in any case, intrabrand competition might erode the dealers' margins. The manufacturer might be able to implement a multi-product pricing scheme to subsidize the inputs to promotional activities. But, again, the erosion of the dealers' margins may reduce the profitability of this strategy. Moreover, the use of such schemes may be limited by the manufacturer's inability to observe the input levels. This problem is particularly acute when unobservable dealer effort is one of the inputs to the production of promotional activities.

Turning to resale restraints, Telser (1960) and Yamey (1954) noted that resale price maintenance can be used to preserve a large dealer margin that generates downstream incentives to engage in promotional activities.[32,33] The imposition of resale price maintenance does not, however, correct the distortion that nonmarginal cost pricing of the input might cause. Telser and Yamey also identified a second, and somewhat indirect, way in which a resale price floor encourages the provision of promotional effort. A restriction on price competition makes it difficult for a low-service dealer to free ride on a high-service dealer because it limits the ways in which the low-service dealer can attract consumers away from the other dealer. In this way, resale price maintenance both protects a high-service

[32] Silcock (1938) presented an informal version of this argument.

[33] Bowman (1952, 1955) and Comment (1951) also argued that resale price maintenance was necessary to compensate dealers for providing services subject to externalities. In these articles, the problem is treated as one of adverse selection with two types of dealer (i.e. low-cost, low-service dealers and high-cost, high-service dealers) rather than one of moral hazard. These informal models correspond closely to Marvel and McCafferty's (1984) model of quality certification discussed in Subsection 3.1.2.3 below.

dealer and gives a low-service dealer an incentive to raise his service level. This type of effect is likely to be strongest when the promotional activity is something like a showroom or expert advice, where the customer has to come into the store to receive the service. Of course, if a dealer can offer particular services that are enjoyed by his customers only, then he may use these services to attract customers from other dealers and free ride on their provision of those promotional activities with a public good nature.

Recently, Mathewson and Winter (1984) and Perry and Porter (1986) examined the use of resale price maintenance to correct for service externalities in fully-specified algebraic models. Telser and Yamey, in contrast, relied solely on verbal arguments. Mathewson and Winter considered a market in which the dealers are spatially-differentiated monopolistic competitors. Mathewson and Winter found that manufacturer would indeed use a resale price floor if he could couple it with a franchise fee (or equivalent form of nonuniform pricing), Perry and Porter considered a modified version of Dixit's (1979) model of monopolistic competition with quality (service) competition. They extended Dixit's model, which is itself an extension of the well-known Dixit–Stiglitz (1976) model, to allow for service externalities. Surprisingly, under this specification of demand and market equilibrium, if the manufacturer's only two instruments are a uniform wholesale price and the ability to fix the resale price, the resulting retail margin is independent of the degree of service externality. Rather than a floor to preserve dealers' promotional incentives, the limit on resale pricing takes the form of a *ceiling* to control the double marginalization problem. This result is a provocative one, and it will be interesting to see to what extent this finding carries over to other model specifications.[34]

In addition to using resale price restraints to limit dealers' free riding indirectly, the manufacturer can use customer resale restraints to block free riding directly. One way is to impose a system of exclusive territories under which the final market is divided into nonoverlapping geographic segments with one dealer in each segment. While exclusive territories may limit dealer free riding, there are costs associated with them. The efficient size of market region in terms of transportation and production costs may differ from the efficient market size based on media coverage or other promotional considerations. And, as discussed below, intrabrand competition may serve as a valuable control instrument to the manufacturer, an instrument that is lost when exclusive territories are assigned.

Nongeographic customer restraints can work much like exclusive territories to prevent free riding. Faced with customer restrictions, a dealer is prevented from

[34]Asymmetric equilibria (i.e. full-service dealers co-existing with discounters) also deserve attention. Perry and Porter (1986) restricted their attention to symmetric equilibria.

stealing the business of a consumer in whom another dealer has invested time and money in making a sales pitch or providing other customer-specific promotional services for the manufacturer's brand. This may be one reason why large customers, who have offices in many sales regions, often are reserved for national-account sales forces.

3.1.2.2. Product reputation. Reputational externalities can arise when product quality is unobservable to consumers prior to purchase and they form expectations about the quality of one dealer's product based in part of the quality of other dealers' products. Reputational effects are likely to be particularly important in the case of a franchise, where the dealer presents himself to final consumers under the manufacturer's trade name. A consumer's prediction of the quality of a Burger King hamburger served in Berkeley, for example, is likely to be heavily influenced by the quality of hamburgers consumed in the past at other Burger Kings.

Two conditions must be satisfied in order for dealer reputations to give rise to externalities. One, customers must believe that the product quality observed by sampling (i.e. buying a unit of the good) at one dealer is a good predictor of the product quality offered by another dealer. Consumers may believe such a link exists either because the dealers all use the same input provided by the manufacturer that is a key determinant of product quality, or merely because the consumers know that all dealers face the same optimization problem in choosing their actions.[35] The second necessary condition for the existence of reputational externalities is that either individual consumers are mobile across dealers, or a consumer bases his or her expectations on the experience of other consumers who are spread across dealers.

If product quality is not influenced by dealer choices (i.e. quality is determined entirely at attributes of the manufacturer's input), then the existence of reputational externalities poses no additional problems of dealer moral hazard. Often, however, the dealer can affect consumers' valuation of the good. Suppose that there are multiple inputs to downstream production and that consumers care about the input mix. The relative input levels can be thought of as the "quality" of the service provided by the output. Moreover, suppose that consumers are unable to judge the quality prior to purchase and do rely on past experience with the good. Any one dealer's choice of quality affects the future demand (i.e. repeat purchases) both for his output and for the output of all of the other retailers. As

[35]In fact, consumers probably believe that all dealers offer the same quality of product simply because that is what consumers typically observe in product markets. The reasons given in the text really are explanations of why these beliefs tend to be correct in equilibrium.

in the case of promotional externalities, dealers fail to internalize the reputational externality in choosing their input mixes and hence tend to undersupply costly quality.

The manufacturer's potential responses to reputational externalities are similar to those in the case of promotional externalities. For example, the manufacturer could lower the wholesale price and impose a resale price floor to make repeat purchases more valuable and thus increase the incentives to provide high-quality output. But again, such pricing could induce other distortions.

Some authors argued that resale price maintenance is profitable when the price itself affects a product's reputation and consumers' perceptions of its value. Taussig (1916) suggested that a resale price floor could be used to keep prices high in a market in which final consumers are sensitive to price as an indicator of the snob value of a "prestige good" (a rose by any other price might not smell as sweet). He asserted that such effects are particularly likely in markets in which the dealer is simply a retailer for the manufacturer's branded product. Absent a resale price floor, any one dealer might have incentive to lower his price in order to attract more sales, even though the overall effect of this move would be to lower the total unit sales of the good by eroding its image. Proponents of resale price maintenance argued that this effect would be particularly strong in markets in which multi-product dealers would otherwise use the manufacturer's product as a loss leader to attract consumers.[36]

Silcock (1938) argued that resale price maintenance could increase unit sales of a good by providing uniform pricing. He asserted that in the absence of price dispersion, consumers would tend to buy out of habit. He then argued that price dispersion across dealers would encourage consumer search, which would, in turn, induce consumers to think about product quality and whether they really wanted the good. As the final leg of his argument, Silcock posited that an introspective consumer would tend to purchase less of the good than would a buyer who acts out of habit.

Returning to more narrowly conceived economic models, the manufacturer may have incentive to implement a multi-product pricing scheme in response to problems of reputational externalities. By tying the other inputs to downstream production to his product, the manufacturer may be able to choose the relative prices in such a way as to induce the collectively optimal input mix for each dealer. In fact, manufacturer-imposed standards on the quality of dealer output can be thought of as a particular tie entailing multi-product quantity-dependent pricing. As always, when unobservable dealer effort is one of the inputs, complete tying is infeasible, and it may be impossible to induce the joint-profit-maximizing quality level.

[36]Bowman (1955) discussed, and rejected, the loss leader argument.

3.1.2.3. Quality certification. Marvel and McCafferty (1984) also considered a market in which consumers value product quality or stylishness, but are unable to observe these attributes prior to purchase. Unlike the cases just considered, the dealers in their model cannot influence the product's quality (i.e. there is no input substitution). Despite this fact, dealers can affect consumers' *prepurchase* beliefs about these attributes.

Each dealer carries a variety of products and establishes a *dealer-specific* reputation for not carrying products of quality below some threshold level. For example, a stereo store might stock the products of several different manufacturers and establish a reputation for carrying only that equipment that is of the highest quality. Or, a certain overpriced New York department store might carry only those designer denims that it knew to be stylish. In either case, as long as the quality or stylishness of a given product is known to be constant across dealers, consumers can use the fact that a high-threshold dealer carries a good as a means of product certification.

Consumers raise their expectations of the manufacturer's product quality when they observe the good being sold by a store that has a reputation for carrying only high-quality products even if they do not purchase that good from the store. Consumers do not care about the cutoff level or reputation of the particular dealer from whom they purchase the good, and they purchase the product from the store with the lowest price. Thus, by stocking the good, a dealer with a high quality threshold raises consumer expectations, expands demand, and creates a positive externality for all dealers and, ultimately, for the manufacturer.

Quality certification is costly. These costs can arise from the extra testing or market research that dealers with high thresholds must conduct to ensure that their more stringent standards are met. Marvel and McCafferty took the distribution of quality thresholds among potential dealers to be exogenous.[37] In other words, a potential dealer is unwilling to alter his reputation for the sake of a single product. This assumption may be a reasonable one for a dealer who sells a large number of different products (e.g. a department store).

A key feature of the Marvel–McCafferty model is that market demand is increasing in the number of dealers selling the product. This effect arises when dealers are differentiated and consumers value dealer variety.[38] In a market in which the downstream firms are spatially-differentiated retailers, the greater the number of dealers, the lower the average distance that a consumer must travel to reach a dealer carrying the manufacturer's product. Or, consumers may value

[37]Hence, in contrast to the case of tangible presale services considered in Subsection 3.1.2.1 above, the manufacturer faces a problem of dealer adverse selection rather than dealer moral hazard.

[38]Marvel and McCafferty's assumption that retailers are differentiated appears to be somewhat unnatural in the light of their assumption that all dealers are price takers in the final good market. Their analysis could, however, be generalized to allow for some downstream market power.

dealer variety because the downstream firms use the manufacturer's output to produce differentiated goods.

Given that certification is costly and there is a benefit from increased dealer density, the manufacturer may desire a mix of dealers. The manufacturer would like to have some dealers with high thresholds in order to certify the product's quality, while having other dealers with low thresholds in order to minimize distribution costs.[39] The manufacturer's problem is to induce this pattern of dealers.

If the manufacturer can select his retailers and can set dealer-specific two-part tariffs, then he can implement the desired distribution of dealers exactly. The quality certification mechanism does not introduce any new reasons to use contractual provisions beyond single-product pricing as long as the manufacturer is allowed to charge different prices to different dealers. In effect, the manufacturer can purchase certification services by setting a low (or even negative) franchise fee for dealers with high quality thresholds. Such a pricing scheme may, however, run afoul of the Robinson–Patman Act.[40]

Suppose, as did Marvel and McCafferty, that the manufacturer can set only a uniform wholesale price (i.e. franchise fees or lump-sum payments are infeasible), and that the manufacturer's right to refuse dealers is limited to setting a marketwide standard. That is, the manufacturer can choose a single quality level such that any store with a quality threshold below that level is prohibited from carrying the product, but any store with a higher threshold must be allowed to serve as a dealer if it so desires.

The manufacturer faces an adverse selection problem. For any given wholesale price, low-threshold, low-cost dealers tend to drive high-threshold, high-cost dealers out of the market, lowering the product's image. Hence, the manufacturer may find it desirable to set a binding standard in order to facilitate the survival of stores with high quality thresholds. The problem with this approach is that by throwing out low-cost dealers, the manufacturer limits his coverage of the market.

The manufacturer may be able to increase his profits by implementing a resale price floor. By setting a final price floor, the manufacturer can increase the equilibrium price–cost margin and the profits of existing dealers. New dealers will be attracted. Given the order in which firms find it profitable to enter the market, these new dealers will have higher quality thresholds than any of the existing dealers. Hence, the product's image will be improved and sales will rise.

[39] Formally, consumers in Marvel and McCafferty's model base their expectations of product attributes by taking the highest quality cutoff level of any store that stocks the manufacturer's product. Under this extreme assumption, the manufacturer's optimal portfolio of dealer types comprises a single dealer with a threshold just equal to the true product quality and all other dealers with the lowest possible thresholds.

[40] For a further discussion of this point, see Bowman (1952) and Comment (1951).

The manufacturer's profits will rise as well. Put another way, through a resale price floor, the manufacturer can insure the survival of a high-threshold dealer, while continuing to have low-cost, low-threshold dealers serve customers.[41]

Although not considered by Marvel and McCafferty, the assignment of exclusive territories might also be used to protect high-cost, high-threshold firms. But note that, for this strategy to be profitable, the certification effects would have to carry across the boundaries of the exclusive territories, a condition that might be unlikely to hold.

3.2. A new control instrument

3.2.1. Relative performance schemes

When there is uncertainty about some parameter of the dealer's profit function and the dealer is risk averse, the agency contract must both provide the dealer with incentives and provide the dealer with insurance. In the single-dealer case, there is a conflict between these two roles for the contract. When there are multiple dealers, the manufacturer may be able to separate these two roles and offer a contract that provides both full incentives and full insurance.

These points are best seen within the context of a simple example. The upstream firm owns a patent that he can license to (for now) a single, risk averse, downstream firm. The downstream firm earns revenues of

$$R(x, e; \theta) = \begin{cases} 0, & \text{if } x = 0, \\ r(e + \theta), & \text{if } x = 1, \end{cases}$$

where $r(\cdot)$ is some increasing function. At the time of contracting, the dealer does not know the realization of θ, although he learns it before choosing his level of effort.[42] The dealer cares about θ solely through its influence on r; θ does not enter the dealer's utility function directly. Ex post, the manufacturer can observe the value of r (and hence the sum of e and θ), but he cannot observe the individual values of e and θ.

Define $e^{m}(\cdot)$ and $F^{m}(\cdot)$ as the solution to

$$\underset{e(\cdot),\, F(\cdot)}{\text{maximize}} \; \underset{\theta}{E}\, F(\theta)$$

$$\text{subject to} \quad \underset{\theta}{E}\, U\left[r(e(\theta) + \theta) - F(\theta), e(\theta)\right] \geq 0.$$

[41] If there is an unlimited number of dealers of each type, then in equilibrium all stores are of the lowest threshold allowed, and there is nothing that the manufacturer can do indirectly.

[42] As was shown in Section 2, when the value of θ is known to would-be dealers at the time of contracting, an auction among would-be dealers can serve as a relative (prospective) performance scheme.

$e^m(\cdot)$ and $F^m(\cdot)$ are the solution to the manufacturer's problem when contractual provisions based on the realization of θ are enforceable. This solution satisfies the first-order conditions $U_1 = \mu$ for all θ, μ some constant, and $U_1 r' + U_2 = 0$. The dealer has a constant marginal utility of income across realizations of θ, and the level of effort always is the efficient one.

This contract is not incentive compatible when θ is unobservable to the manufacturer. If the license fee were to depend on the report of θ, the dealer always would report that value that minimized the license fee. Under this outcome, the dealer would not be insured against poor realizations of θ, but he would supply effort at the efficient level.

Suppose that contracts based on r are enforceable. That is, all parties, including the courts, can observe $\theta + e$, although only the dealer can observe the components of this sum individually. When the license fee depends only on the value of r, it is a royalty scheme: $F(r)$. The dealer's payoff is $U[r - F(r), e]$. The manufacturer would like to set the royalty rate to insure the dealer against low realizations of θ while at the same time inducing the downstream firm to supply the optimal amount of effort.

To provide full incentives to supply effort, the dealer's compensation must be a sharply increasing function of his gross profit (i.e. $dF/dr = 0$). But, under this contract, the dealer bears all of the risk. A full insurance contract (i.e. $dF/dr = -1$), on the other hand, would eliminate any incentive to provide effort. Given his incomplete and imperfect information, the manufacturer is unable to separate dealer shirking from bad luck. The profit-maximizing contract represents a compromise between the provision of insurance and incentives.

Now, suppose that the manufacturer has two dealers, 1 and 2, who serve markets that are independent of one another in the sense that the actions of one dealer have no effect on the sales or profits of the other dealer; there is no direct competition between the dealers. Moreover, suppose that each market is identical to the one just considered for the case of a single dealer. It might appear that the optimal contract for the single-dealer problem is the optimal contract for each dealer in this case as well. It need not be. When the values of the unobservable parameters that affect dealer profits are correlated across markets, the manufacturer may be able to increase his profits by implementing a relative performance scheme under which a given dealer's compensation depends not just on his performance, but also on the performance of the other dealer.

To keep matters as simple as possible, suppose that the realization of θ is the same for all dealers. θ may be thought of as a measure of the quality of the innovation that is being licensed. The following contract attains the full private optimum. Define $F^n(r)$ such that

$$F^n[r(e^m(\theta) + \theta)] = F^m(\theta).$$

The royalty paid by dealer i is equal to $F^n(r_j)$, $j \neq i$, where r_j is the revenue earned by dealer j.

Faced with this contract, firm i chooses $e_i(\theta)$ to

$$\underset{e(\theta)}{\text{maximize}}\ U\left[r(e(\theta) + \theta) - F^n(r_j), e(\theta)\right].$$

There exists a symmetric equilibrium in which, for all i and θ, $e_i(\theta) = e^m(\theta)$, and thus $F^n[r(e_i(\theta) + \theta)] = F^m(\theta)$.[43]

The essential benefit of the relative performance mechanism is that it allows the manufacturer to separate insurance and incentives in the contract. Dealer incentives are realized by making each dealer the residual claimant to the revenues that he generates. The dealer is simultaneously insured by making the fixed fee depend upon the performance of the *other* dealer. If the other dealer does poorly, then the original dealer pays a lower "royalty" than he otherwise would.

When θ is not constant across dealers and the realizations are not perfectly correlated, this type of contract does not fully solve the incentive and insurance problems. It is clear, however, that relative performance schemes often can increase manufacturer profits in more general environments.

3.2.2. The market mechanism as an incentive and insurance scheme

The sort of scheme sketched above does not resemble any of the contract types typically considered in the vertical restraints literature. Rey and Tirole (1986) applied this type of analysis to more traditional contract forms. They showed, inter alia, that direct competition between the dealers can serve as one form of relative performance scheme.

The following example illustrates their analysis. Output is produced according to a constant proportions downstream production technology where one unit of the input provided by the monopolistic manufacturer must be combined with one unit of a competitively supplied unit, with cost θ_2, to produce one unit of output. There are two dealers, one located in city A and the other in city B, where each city has a demand curve $P = \theta_1 - X$. There are no transportation costs between the two cities. The dealers are Bertrand competitors and face the same values of θ_1 and θ_2 as one another.

At the time of contracting, all parties know the underlying parameters of the dealer's profit function except for θ_1 and θ_2. After the contracts have been signed, but before any price or output decisions are made, the dealers learn the realizations of θ_1 and θ_2.

[43]Shleifer (1985) considers a similar sort of relative performance scheme in the context of a regulated industry.

The joint-profit-maximizing price is $p^m(c; \theta_1, \theta_2) = (\theta_1 + \theta_2 + c)/2$. The manufacturer would like to induce the dealers to set this price. Since θ_1 and θ_2 are unobservable to the manufacturer, the contract cannot be contingent upon their realizations. Suppose that the manufacturer relies solely on single-product pricing and allows unfettered competition between the dealers. If the manufacturer set $W(x) = F + wx$ and imposed no other restrictions on dealer behavior, the dealers would drive the final price down to $w + \theta_2$. Dealer competition is relative performance scheme in the following sense: the residual demand curve that a dealer faces depends on the cost level of his downstream rival.

Unlike the monopoly price, the competitive price shows no response to shifts in demand as measured by θ_1. The equilibrium price fully adjusts to changes in cost θ_2, however. In fact, the competitive price overadjusts. The use of a two-part tariff fails to solve the problem of dealer moral hazard in price setting. Note, though, that the two-part tariff does provide the dealers with full insurance; each dealer earns $-F$ no matter what the values of θ_1 and θ_2.

Now, suppose that the manufacturer imposes exclusive territories and restricts each dealer to selling in his home city only. Each dealer is a final good monopolist in his home city and sets price equal to $p^m(w; \theta_1, \theta_2)$. The manufacturer's setting $w = c$ fully solves the dealer moral hazard problem, but the dealers bear all of the risk – the manufacturer's profits are equal to $2F$, independent of the realizations of θ_1 and θ_2. Exclusive dealing ameliorates one problem and exacerbates another.

Lastly, consider resale price maintenance in the form on a fixed price (i.e. a simultaneous floor and ceiling). Clearly, this price does not adjust to either changes in costs or demand.

Which contractual provisions will the manufacturer find privately optimal to implement? In the absence of uncertainty, the policies of unfettered dealer competition, resale price maintaince, and exclusive territories coupled with a franchise fee all lead to the same outcome in the final good market and to the same levels of manufacturer and dealer profits. A major finding of Rey and Tirole's work is that these policies are not equivalent in the presence of uncertainty about the optimal final good price. In the presence of uncertainty, the manufacturer has a real choice to make.

The manufacturer–dealer contract has three potential roles to play in this setting: incentive provision, income transfer, and insurance. If the dealers are risk neutral, then there is no need for insurance provision. The privately optimal two-part tariff coupled with exclusive territories yields the maximal level of manufacturer profits; the wholesale price gives dealers the correct incentives and the franchise fee transfers income. Both dealer competition and resale price maintenance do relatively poorly in terms of incentive provision. Under dealer competition prices fail to respond to changes in demand,[44] although they do

[44]This result does not carry over to more general cost structure or to non-Bertrand firms.

(suboptimally) respond to changes in cost under competition. And, of course, prices do not respond at all under resale price maintenance. Therefore, with risk neutral dealers and either cost or demand uncertainty, the manufacturer strictly prefers the assignment of exclusive territories to either unfettered intrabrand competition or resale price maintenance. Rey and Tirole showed that resale price maintenance and competition do equally well from the point of view of expected manufacturer profits.[45]

Rey and Tirole assumed that the dealers' downstream prices were the only dimension of dealer moral hazard or adverse selection. More generally, there are several types of incentive problem that can arise as a result of asymmetric information. As discussed in Section 2 above, the manufacturer may have incentives to set $w > c$ in response to these problems when his dealers are assigned exclusive territories. Pricing the intermediate good above cost leads to overly-high final good prices when the dealers have exclusive territories and there are no resale price restraints. Hence, in a general setting, the manufacturer may find it optimal to impose resale price maintenance or to allow dealer competition. A similar result obtains when fixed fees are infeasible.

What if the dealers are risk averse? Following Rey and Tirole, take the case of infinitely risk averse dealers. That is, consider dealers who care only about the minimal level of profits that they might receive. Under competition, the dealers are fully insured against both cost and demand shocks. If there are no cost shocks, full insurance with respect to demand shocks is feasible under resale price maintenance (by imposing a downstream price–cost margin equal to 0). When dealers face solely demand uncertainty, resale price maintenance and competition are equally profitable for the manufacturer.[46] Because of their superior insurance properties, both policies are preferred to the assignment of exclusive territories.[47] Turning to cost shocks, resale price maintenance does poorly in providing insurance, and it is the least profitable of the three policies. Competition is the most profitable, while the assignment of exclusive territories falls in between the two. In summary, dealer competition can sometimes serve as an effective relative performance scheme.

3.3. The number of dealers

The manufacturer's profits typically vary with the number of dealers. As shown above, an increase in the number of dealers may have negative effects due to the

[45] Rey and Tirole noted that this equivalence is a general one when the dealers face solely demand uncertainty. In the presence of cost uncertainty, however, this result is dependent on the linearity of demand.

[46] Rey and Tirole also showed that when dealers are slightly differentiated, resale price maintenance is superior to dealer competition because it avoids the problems associated with double marginalization.

[47] Recall from Section 2 that the manufacturer insures a dealer with an exclusive territory by setting $w > c$, which distorts dealer incentives.

increased dealer pricing and promotional externalities. In his study of the tuna industry, Gallick (1984) identified another reason that having more than one dealer may lower the manufacturer's profits. Suppose that the quality of the manufacturer's output varies across units in a way that cannot be controlled by the manufacturer, but can be observed by him. If the manufacturer can sell his output to a variety of dealers, then each dealer must be concerned about getting an adverse selection of output. In response, each dealer may have to examine the quality of the input that he purchases. This inspection is costly and is of no collective value if all dealers value any given quality of output equally. Having the manufacturer sell all of his output to a single dealer avoides this adverse selection problem and economizes on inspection costs.

Of course, having multiple dealers can give rise to several types of benefits. First, having multiple dealers may be a valuable element of the vertical control scheme. The resulting increase in intrabrand competition serves to limit the markup of the final price over marginal cost, which in some cases is beneficial to the manufacturer. Moreover, his having multiple dealers allows the manufacturer to develop a benchmark against which to measure any one dealer's performance. A second benefit of multiple dealers arises when dealer variety is valuable to consumers: the derived demand for the manufacturer's output is an increasing function of the number of dealers, ceteris paribus. Finally, increased dealer variety may allow greater discrimination against final consumers by increasing the fineness of the partition of consumer types that is created by consumers' self-selecting across dealers.

There are several different ways in which the manufacturer may influence the number of dealers. The number of dealers may be chosen directly by the manufacturer (he simply refuses to deal with more than a set number of downstream firms). Alternatively, the manufacturer may indirectly control the number of dealers through his pricing policy or the nonterritorial resale restraint provisions of his contract offer. Suppose, for instance, that there is free entry into the downstream industry and the manufacturer affects the number of dealers indirectly through his single-product price schedule. Changes in either the franchise fee or the wholesale price change the equilibrium number of dealers. In the case of the fixed fee, the effect of pricing on the number of dealers is straightforward. The higher the fixed fee, the lower the number of retailers in a zero-profit equilibrium.[48]

Katz and Rosen (1985) and Seade (1985) showed that raising the wholesale price may *raise* or lower the number of retailers. The initially surprising result that a higher input price can increase the number of firms has the following intuition. Raising the price to any one firm, lowers that firm's equilibrium output

[48]To get the desired number of retailers, the manufacturer might have to set a negative franchise fee.

and profit levels. This decrease in output increases the profits of rival firms. The gain to rivals may dominate the loss to the higher-cost firm, so that an industry-wide price shift may increase industry profits and make new entry profitable. In effect, the increase in dealer marginal cost induces a cartel-like cutback in downstream output whose benefits may dominate the increased costs.

Turning to resale restraints, Gould and Preston (1965) examined the manufacturer's incentives to impose a resale price floor in a downstream market with a symmetric, perfectly competitive equilibrium.[49] Gould and Preston argued that, when final consumers value dealer variety, a manufacturer would have incentives to implement a resale price floor in order to increase the number of dealers by preserving their price–cost margin.[50] Early proponents of resale price maintenance claimed that resale price floors would be even more strongly needed in markets with asymmetric equilibria in which some (multi-product) dealers would use a product as a loss leader, discouraging other dealers from carrying the good.[51]

Bittlingmayer (1983), Mathewson and Winter (1983b), and Perry and Porter (1986) examined the use of resale price maintenance by a manufacturer with monopolistically competitive dealers.[52] Contrary to Gould and Preston, these authors all found that a manufacturer who cannot set a nonzero franchise fee will use a resale price *ceiling* to keep down both the dealers' price–cost margin and, hence, the number of dealers. In these models, resale price restraints are used to ameliorate the double marginalization problem rather than to increase the number of dealers. Since the dealers in Gould and Preston's model are perfect competitors, there is no problem of double marginalization, and the manufacturer imposes a resale price floor. Perry and Porter did, however, find that a manufacturer who can set positive franchise fees will impose a resale price floor which either increases or leaves unchanged the number of dealers in comparison with the equilibrium absent resale price maintenance.

One reason for the manufacturer in the Perry and Porter model to impose a resale price floor is to give dealers incentives to provide promotional activities that are subject to externalities. Gould and Preston pointed out that when the manufacturer sets a resale price floor for this reason, he may find it desirable to limit the number of dealers directly. One way to restrict the number of dealers is to assign a limited number of exclusive territories.

[49]Gould and Preston themselves noted that the assumption of perfect competition among dealers appears to be inconsistent with consumers' valuing dealer variety. Gould and Preston hypothesized that their analysis would carry over to markets with monopolistically competitive dealers.

[50]This point also was made informally by Bowman (1955) and Silcock (1938).

[51]See Bowman (1955) for a discussion of, and attack on, the loss leader argument. See Yamey (1952) for an interesting history of the use of this argument by supporters of resale price maintenance.

[52]Bittlingmayer (1983) and Mathewson and Winter (1983b) considered spatially-differentiated dealers facing either linear or exponential point demand functions in a market with linear transportation costs and no promotional spillovers. Perry and Porter (1986) is described in Subsection 3.1.2.2 above.

As Preston (1965) observed, the imposition of exclusive territories also can be used to raise profits per dealer and increase the number of dealers that the downstream market can support. The reason is that, for any given wholesale price and set of dealers, the assignment of exclusive territories leads to higher final prices than does competition among dealers. In the face of dealer moral hazard or adverse selection, the manufacturer may find it preferable to assign exclusive territories rather than to rely on negative franchise fees.

4. Private incentives: Multiple manufacturers

The situation most often encountered in actual markets, but least seen in economics journals, is that of multiple manufacturers, each of whom has many dealers. As in the cases examined in earlier sections, there may be externalities across these dealers. There also is a new type of externality that may arise when there is more than one upstream firm: externalities across manufacturers.

In this section I will focus on the new issues that these externalities raise, including: (1) the possibility of service externalities across *brands*; (2) the effects of vertical contracts on interbrand competition; and (3) the effects of vertical contracts on the conditions of entry into the manufacturing stage. I will discuss each of these areas in turn. In this discussion, I will consider, inter alia, the use of the final type of vertical restraint identified in the introduction: when there is more than one upstream producer, a manufacturer may require his dealers to agree not to carry the products of rival suppliers.

4.1. Free riding across manufacturers

Consider the following variant of the free riding story that was told for the case of a single manufacturer with multiple dealers. There is a stereo dealer who carries several brands of hi-fi equipment. Suppose that the manufacturers rely solely on uniform pricing and that one manufacturer engages in heavy advertising that attracts consumers to the dealer. Once consumers have come into the store, the dealer has incentives to persuade the consumers to purchase whichever brand offers the largest profit margin. That brand need not be the one that conducted the advertising responsible for attracting the consumers to the store. As with free riding across dealers, free riding across manufacturers may lead to levels of promotional activity below the profit-maximizing ones.[53]

[53] In those cases in which competition among manufacturers would otherwise lead to equilibrium advertising intensities above their collusive levels, the manufacturers may benefit from the reduction in advertising incentives due to free riding.

There are several means by which to deal with the problem of free riding across brands. One is to adjust the single-product pricing policy. The advertising manufacturer could lower his wholesale price and compensate for this price reduction by increasing the franchise fee. If each manufacturer set his wholesale price equal to his marginal cost, then there would be no distortions in the dealer's marginal incentives. As was seen in Sections 2 and 3, however, a two-part tariff with a low wholesale price may suffer several disadvantages. It may exacerbate problems of manufacturer moral hazard and adverse selection. And a low wholesale price may lead to unprofitably low prices for downstream output due to inter-dealer competition.

The imposition of a resale price floor could eliminate this second problem. But the price floor would do nothing to mitigate the problems of manufacturer moral hazard and adverse selection. Resale price maintenance also would fail to mitigate the problems of wholesale price cutting by rival manufacturers and inefficient risk-bearing by the dealers.

In principle, multiproduct pricing could be used in response to free riding across brands. The advertising manufacturer could demand a royalty payment based on the dealer's total revenues or unit sales across all brands. Under such a scheme, the advertising manufacturer would appropriate some of the promotional benefits that spill over to rival brands.[54] In many respects, this form of multiproduct pricing is like a cooperative advertising agreement.

Marvel (1982) identified the imposition of exclusive dealerships as another way for a manufacturer to respond to the problem of within-store brand switching. Of course, any gains from exclusive dealing due to greater appropriation of promotional efforts must be balanced against the increase in distribution costs due to lost economies of scope in downstream selling. Exclusive dealing also might reduce manufacturer profits through its effects on the pattern of consumer sampling. Consumers engaged in sequential search may have incentives to sample multi-brand retailers first in order to minimize search costs. Consequently, single-brand retailers might suffer diminished sales.

4.2. Lock-in

A complete contract typically is unprofitable (due to the tremendous costs of calculation and administration) or entirely infeasible (due to the incompleteness of information). Hence, the agreement between the two parties rarely, if ever, fully specifies what should happen in all possible contingencies. This incompleteness of contracts gives rise to possible opportunistic behavior by one or both of

[54]Antitrust authorities probably would object to this type of arrangement on the grounds that it is an illegal attempt to monopolize the final good market.

the parties to the contract. Opportunistic behavior presents a particular problem when one side of the manufacturer–dealer relationship makes expenditures on assets that have value only in that relationship. Once the relationship-specific costs have been sunk by one party, the other party may reopen bargaining over the terms of the contract.

For concreteness, assume that the manufacturer is the one who must make significant relationship-specific investments. Once the manufacturer has made the sunk investments, he is locked-in to the relationship to some extent. The dealer may be able to take advantage of this fact to renegotiate the contract in a way that is more favorable to him. The dealer may induce renegotiation by threatening to break off the relationship or to engage in behavior that takes advantage of uncontracted contingencies that have arisen since the initial signing of the contract.

Both parties may suffer from the possibility of opportunistic behavior by the dealer. Faced with the prospect of the dealer's renegotiating the contract, the manufacturer may choose an inefficiently low level of investment in order to limit the extent to which he is locked-in to the relationship. This reduction in the investment level may reduce the value of the relationship to both parties. Hence, both parties may have ex ante incentives to limit the dealer's ability and incentive to engage in opportunistic behavior.

There are two types of force that can limit the dealer's opportunistic behavior and, thus, can help restore efficiency in investment and production choices. One force arises from the fact that the relationship is an ongoing one. If the dealer takes advantage of the manufacturer today, he may be forgoing future profits that could be enjoyed from the relationship.[55] As recognized by Schelling (1956, 1960) and developed more fully by Williamson (1983) and others, the parties may take actions designed to increase the value of maintaining the relationship to both sides. One way is to have the dealer offer a hostage. That is, to have the dealer invest in a relationship-specific asset of his own. This hostage serves two roles. First, it attenuates the dealer's threat to terminate the relationship. Second, the existence of a hostage makes the *manufacturer's* terminating the relationship a stronger punishment that can be imposed on the dealer in response to his behaving opportunistically.

In some instances, the pricing and resale restraint provisions of the manufacturer–dealer contract can serve to create hostages. One means of creating a hostage, for example, is to have a low wholesale price coupled with a large franchise fee; the right to the low wholesale price becomes a relationship-specific asset owned by the dealer. In this case, an artificial asset is created. Alternatively,

[55]See Williamson (1983) and the references cited therein for a more complete discussion of self-enforcing long-term contracts.

the contract may be used to create a relationship-specific asset out of an asset that would exist in any event but would not otherwise be a relationship-specific one. Dealer reputation may be a candidate for such an asset. A multi-brand dealer may be able to develop a reputation as a good dealer per se. When the dealer carries only one brand, his reputation may become inextricably linked with that product. Hence, an exclusive dealing arrangement may lead to the dealer's reputation becoming a relationship-specific asset.

Another way to deal with opportunistic recontracting is to limit the threats that the dealer can make.[56] A requirements contract, for instance, eliminates the buyer's ability to threaten to take his business to another manufacturer. Suppose that the manufacturer offers the dealer insurance in the form of a long-term contract that specifies a constant wholesale price in a market in which the spot prices charged by other manufacturers vary over time. If two-part tariffs are feasible, the manufacturer can sell the option to buy the input at a given price for a lump-sum payment made in advance (or spread over time according to a path that is independent of the dealer's purchase levels). If franchise fees are infeasible, however, the manufacturer must rely on the dealer's "paying" for this insurance by making purchases when the price set in the contract is greater than the current spot market price. But the dealer has incentives to make purchases from other manufacturers during such periods. As Blair and Kaserman (1984) noted, tying in the form of a requirements contract can prevent the dealer from diverting purchases to other manufacturers in periods of low spot prices.

While the emphasis here has been on opportunistic behavior by the dealer, the manufacturer could engage in such actions as well. In the case just considered, if the dealer pays for insurance in advance, then the contract has to set the long-term price and terms of delivery to ensure that manufacturer will, in fact, deliver the product in times of high market prices. This example brings up the important, if obvious, point that the parties need to make sure that their attempts to correct for dealer misincentives do not themselves generate manufacturer misincentives. The parties must seek a balance in creating a bilateral hostage exchange to facilitate efficient trade of the intermediate good.

4.3. Effects on interbrand competition

The nature of manufacturer–dealer contracts may affect competition among manufacturers or among the dealers of the different manufacturers. The manu-

[56]As Farrell and Shapiro (1987) showed, it is not always the case that efficiency is improved by limiting the one party's ability to take inefficient actions. Such a prohibition may drive the opportunistic party to take even more inefficient actions to achieve the same end.

facturers would like to use the pricing and restraint provisions of these contracts as means of diminishing interbrand competition. There are several mechanisms through which the manufacturers' contracts with their dealers may at least partially attain this goal.

4.3.1. Lower cross-price elasticities

In many models of oligopoly, the lower are the cross-price elasticities of demand, the weaker is interbrand competition and the higher are producer profits. There are two ways in which provisions of the manufacturer–dealer contract may be used to lower the cross-price elasticities of demand. First, as discussed earlier in this section and in Subsection 3.1.2, the imposition of vertical restraints may increase the levels of promotional activities and other forms of nonprice competition, which may in turn strengthen product differentiation and lower cross-brand elasticities. Thus, as Comanor (1968) argued, restraints such as resale price maintenance or exclusive territories may weaken *interbrand* competition.

The second way in which vertical restraints may lower cross-price elasticities is through their effects on consumer search costs. When all manufacturers demand that their dealers agree to exclusive arrangements and dealers are spatially differentiated, consumer search costs are higher than they would be if a single dealer were to carry several different brands of the product. Under exclusive dealing arrangements, consumers who want to comparison shop are forced to visit several stores. The higher costs of comparison shopping tend to discourage consumers from undertaking this search activity, limiting the extent of interbrand competition and raising industry profits.

Will manufacturers all demand exclusive dealing arrangements? The answer depends in part on the contracting institutions. Suppose that each manufacturer independently chooses the price of his output and whether to demand an exclusive arrangement. If all other manufacturers demand exclusive dealing, then a manufacturer has no real choice to make and may as well require an exclusive arrangement himself; industrywide exclusive dealing always is one equilibrium. What if manufacturers are allowed to coordinate their contract offers? In this case, industrywide exclusive dealing is an equilibrium outcome if and only if the value that consumers place on comparison shopping is sufficiently low that it does not pay any subset of manufacturers to allow their dealers to carry each other's products.

4.3.2. Facilitate collusion among manufacturers

The manufacturers would collectively gain from successful collusion. Typically, individual firms have incentives to cheat on any collusive agreement, and the firms must be able to monitor one another in order to collude successfully.

Contract provisions that improve monitoring may thus facilitate collusion.[57] Telser (1960) examined the role of resale price maintenance in facilitating collusion. Absent resale price maintenance, it may be difficult to tell if the retail price has dropped because the manufacturer has cheated on the collusive agreement or because one of the dealers has chosen to lower his final price on his own initiative. Under resale price maintenance, each brand has a single final price which is set by its manufacturer. Thus, it is relatively simple for the upstream firms to monitor whether their rivals are adhering to the cartel agreement.[58] Collusion also may be easier to sustain because the number of participants in any given agreement is reduced.

Similarly, territorial restraints can facilitate collusions by limiting the number of parties to the agreement. Here, the collusion across brands takes place among the dealers. If each brand assigns exclusive territories to his dealers and territories across brands largely match one another, then only those dealers in any given territory need to reach a collusive agreement.

Telser also noted that exclusive dealing arrangements might be used to limit manufacturers' incentive to make secret wholesale price cuts. If a dealer carries several different brands, any given manufacturer may lower his wholesale price in an attempt to get the dealer to promote that brand at the expense of its rivals. If each dealer carries a single brand, then obviously no such effect arises. Of course, there still would be incentives to lower the wholesale price in order to encourage increased promotional competition among dealers carrying different brands.[59] This type of promotional activity may, however, be easier to monitor than would be in-store promotional activity aimed at inducing consumers to switch brands.

4.3.3. Common agency: Exclusive dealing in reverse

Interbrand competition can be viewed as a negative externality across manufacturers. An obvious way to internalize this externality is to vest all of the decision-making power in a single economic actor who has been made the residual claimant to industry profits. If he is risk neutral, this residual claimant will maximize industry profits. Brennan and Kimmel (1983) and Bernheim and Whinston (1985) identified a common dealer as a natural party to serve as the

[57]As shown by Abreu, Milgrom and Pearce (1987), improvements in monitoring may have the opposite effect in some cases.

[58]If a manufacturer cannot discriminate among his dealers, then the need for resale price maintenance to improve monitoring may be limited. Cheating by the manufacturer would be comparatively easy to detect – rival manufacturers would look for instances in which several dealers for a given brand lowered their retail prices simultaneously.

[59]In his examination of the use of resale price maintenance to facilitate collusion by improving the manufacturers' monitoring of each other, Bowman (1955) concluded that this wholesale price cutting and the resulting promotional competition would undermine the use of resale price maintenance to sustain collusion.

internalizing agent. If all manufacturers sold their output to a single downstream firm and each manufacturer set his wholesale price equal to his marginal cost, the common dealer would act to maximize the collective profits of the two stages.[60]

A big question is: Will the manufacturers all in fact choose to sell to a single dealer? In effect, Brennan and Kimmel examined an industry in which there is a single potential downstream dealer.[61] But more generally, one must consider the possibility that some manufacturer could hire an independent dealer who could free ride on the output restrictions of the other, cartelized brands. As yet, economic theory does not offer compelling answers to questions concerning coalition formation. It is, however, useful to consider a simple and rather artificial coalition formation process.[62]

There are two manufacturers playing the following five-stage game. In stage one, the firms simultaneously announce the contracts that each one is willing to offer a common dealer. In stage two, each firm announces whether it is willing to go with the common dealer under these terms. If they agree to do so, then at stage three the manufacturers simultaneously offer these two-part tariffs to a common dealer. If either manufacturer refuses to seek a common dealer, then in stage three the manufacturers simultaneously offer two-part tariffs (which need not be equal to the contracts proposed in stage one) to different dealers. In the penultimate stage, any dealer to whom an offer has been made decides whether to accept it. The fifth stage consists of some product-market game played among those dealers who accept contracts.

It is convenient to summarize the fifth-stage outcome with a reduced form profit function. Let $\pi^{m}(w_1, w_2)$ denote downstream industry profits when there is a multi-product monopolist with input costs w_1 and w_2. Let $\pi_1^{d}(w_1, w_2)$ denote the profits of dealer i when there is a duopoly in the final good market and dealers 1 and 2 face input prices w_1 and w_2, respectively. A natural assumption is that

$$\pi^{m}(w_1, w_2) \geq \pi_1^{d}(w_1, w_2) + \pi_2^{d}(w_1, w_2),$$

for all w_1 and w_2, with strict inequality whenever the duopolists are not able to achieve the cartel outcome.

[60]Alternatively, one could consider a market in which there are many risk neutral, common dealers, but each one is assinged an exclusive territory.

[61]Brennan and Kimmel (1983) examined a setting in which upstream firms produce joint products, A and B. There is a single potential dealer for product A, while there is a perfectly competitive industry willing to distribute product B. Brennan and Kimmel showed that the sole potential dealer for good A might want to refuse to distribute a manufacturer's output of A unless that firm also distributed his output of B through that distributor. In this way, the distributor could serve as a common agent for all upstream firms for both goods.

[62]This is not the process studied by Bernheim and Whinston (1985). They used a model with a more natural contracting process to reach conclusions similar to those presented here.

In the fourth stage, any dealer who has been offered a contract yielding non-negative profits accepts that contract.

If the manufacturers have chosen to make bids to a common dealer in the second stage, then in the third stage the first-stage offers are put forth. If the manufacturers have chosen to make bids to separate dealers, then in the third stage manufacturer i chooses w_1 and F_i to

$$\begin{aligned} &\text{maximize} && F_i + \{w_i - c\} x_i(w_1, w_2) \\ &\text{subject to} && \pi_i^{\mathrm{d}}(w_1, w_2) - F_i \geq 0. \end{aligned}$$

Let w_i^{n} (for Nash) denote the equilibrium value of w_i given that the manufacturers choose separate agents. For convenience, assume that the values of w_1^{n} and w_2^{n} are unique.

In the second stage, the manufacturer i decides whether to have a common dealer by comparing $\pi_i^{\mathrm{d}}(w_1^{\mathrm{n}}, w_2^{\mathrm{n}})$ with the profits that he would earn under the proposals made in the first stage.

Now, consider the first stage. Since he later can veto the use of a common agent, manufacturer i chooses his bid to maximize his profits under common agency conditional on inducing the other parties to accept the agreement. Hence, if manufacturer j ($j \neq i$) has bid (F_j, w_j), manufacturer i's best reply is the solution to

$$\begin{aligned} &\text{maximize} && F_i + \{w_i - c\} x_i(w_1, w_2) \\ &\text{subject to} && \pi^{\mathrm{m}}(w_1, w_2) - F_1 - F_2 \geq 0 \\ & && \text{and} \\ & && F_j + \{w_j - c\} x_j(w_1, w_2) \geq \pi_i^{\mathrm{d}}(w_1^{\mathrm{n}}, w_2^{\mathrm{n}}). \end{aligned}$$

The form of the agent's individual rationality constraint, the first constraint, reflects the institutional structure that he must either accept both offers or reject both offers. As noted several times previously, the optimal wholesale price is $w_i = c$. The fixed fee is set to drive the dealer's net profits to 0.

It is trivial to verify that there is a continuum of subgame perfect equilibria under each of which the manfacturers hire a common agent, $F_1 + F_2 = \pi^{\mathrm{m}}(c, c)$, $F_i \geq \pi_i^{\mathrm{d}}(w_1^{\mathrm{n}}, w_2^{\mathrm{n}})$, and $w_i = c$, for $= 1, 2$. Under all of these equilibria, the externality across manufacturers is fully internalized by the common agent and industry profits are maximized. The equilibria differ only in the division of profits between the two manufacturers.

Given the (theoretical) benefits of common agency, why would manufacturers ever choose to have exclusive dealers? One possible explanation is that the result above is sensitive to the institutional structure. Another reason for choosing separate dealers may follow from the fact that, when $w_i > c$, a common agent does not maximize industry profits. As shown throughout this chapter, there are

forces limiting the profitability of marginal cost (wholesale) pricing even when there is a common dealer. An important area for further exploration is whether independent dealers can be more profitable than a common one in such settings. For example, it would be interesting to see whether the use of separate dealers might be more profitable than a common agent because downstream competition better overcomes the double marginalization problem.

4.3.4. Observable contracts as precommitments

The provisions of the contracts between manufacturers and their dealers may influence the nature of competition among dealers of different brands. Suppose that each dealer carries a single brand and that the dealers carrying brand A can observe the sales contracts or franchise agreements under which the dealers of brand B operate. Brand-A dealers will condition their product-market behavior on the contracts of their rivals. As several authors [i.e. Fershtman and Judd (1986, 1987), Rey and Stiglitz (1986), and Ross (1987)] have pointed out, a profit-maximizing manufacturer takes these effects into account in choosing the contract to offer his dealers.[63]

The following two-stage game illustrates how the presence of observable agency contracts can affect the product-market equilibrium. In the first stage, each manufacturer offers a contract to his dealer, which the dealer then accepts or rejects. The manufacturers in this example are restricted to offering two-part tariffs. After the first stage is complete, the dealers' contracts become common knowledge among all parties. Then, in the second stage, each dealer chooses a level of final output. x is the only input needed to produce final output, and one unit of x is required per unit of final output.

In any pure strategy equilibrium, manufacturer i must set the franchise fee charged to his dealer equal to $\pi_1^{\mathrm{d}}(w_1, w_2)$. Thus, given w_j, manufacturer i's profits when it chooses w_i (and adjusts the fixed fee appropriately) are

$$\pi_i^{\mathrm{d}}(w_1, w_2) + \{w_i - c\}x_i(w_1, w_2).$$

Suppose that firm i sets $w_i = c$. Differentiating manufacturer i's profits with respect to w_i, one obtains:

$$\partial\pi_i^{\mathrm{d}}/\partial w_i + x_i = x_i P' \partial x_j/\partial w_i.$$

For a Cournot dealer, $x_i P' \partial x_j/\partial w_i$ is negative. Thus, the manufacturer has

[63]Schelling (1960) first suggested the strategic use of contracts as precommitments. Spencer and Brander (1983) and Brander and Spencer (1985) presented a formal model of this type of effect in the context of international trade policy.

incentives to set w_i below c. Under the condition that an increase in x_j shifts down the residual marginal revenue curve for dealer i, the two-stage perfect Nash equilibrium entails each manufacturer's setting his wholesale price below marginal cost.[64]

The intuition behind this result is clear. By setting his wholesale price below marginal cost, the manufacturer makes his dealer a more aggressive competitor in the product market. Faced with a more aggressive rival, the other dealer reduces his output, raising the profits of the first dealer. That dealer's manufacturer then appropriates the profits through the fixed fee.

The set of feasible contracts is greatly restricted in the example just considered. Fershtman, Judd and Kalai (1987) and Katz (1988b) demonstrated that the form of the contract restriction is central to the character of the equilibrium. In fact, once one allows for a relatively general set of contracts, essentially any individually rational outcome in the product market can be attained as the perfect equilibrium in the two-stage game. Thus, if this line of analysis is to be developed further, effort must be concentrated on finding natural restrictions to place upon the nature of contracts.

4.3.5. The effects of unobservable contracts

What if the contract between a manufacturer and his dealer is not observable to the other manufacturers and their dealers? Can unobservable contracts affect the behavior of rival firms? The answer is: "Yes, but not for strategic reasons." Because the contract is unobservable to his rivals, in choosing his contract, the manufacturer ignores any effects that his choice might have on the behavior of other manufacturer's or their dealers. As shown in Katz (1988a), however, this is not to say the vertical contracts have no effect on the product–market outcome.

The effects of unobservable manufacturer–dealer contracts are most easily seen under the extreme assumption that the vertical contracts must take the form of a single uniform price. Again, suppose that there are two manufacturers, each with one dealer. In equilibrium, a manufacturer must set $w > c$ in order to earn positive profits. Even though dealer 1 cannot observe dealer 2's contract, dealer 1 knows that dealer 2 must face a wholesale price that is greater than marginal cost. In fact, under the extreme assumption that he knows all of the parameters of manufacturer 2's contracting problem, dealer 1 can deduce the equilibrium value of w_2 and act accordingly. Dealer 2 can engage in similar reasoning, as can the two manufacturers. Hence, equilibrium entails a lower level of final output and higher final price than would equilibrium if each manufacturer–dealer pair were

[64] This type of two-stage game arises in many other contexts and is discussed more generally in Chapter 6 by Carl Shapiro in this Handbook. He shows, inter alia, that equilibrium may entail $w > c$ for other types of product–market competition.

a single decision-maker with a complete unity of incentives. Similar effects may arise with more general contracts in settings where a manufacturer–dealer contract serves to share risk or governs a situation characterized by asymmetric information.

4.3.6. Raise rivals' costs

A manufacturer's profits typically are an increasing function of his rivals' costs. Thus, a manufacturer is willing to take costly actions that serve to raise his rivals' costs.[65] As discussed in Section 3, Katz and Rosen (1985) and Seade (1985) showed that when marginal costs are increased by some action, even a *symmetric* (across all firms) cost increase may raise a manufacturer's profits. And, if there are asymmetries among manufacturers, some of the manufacturers may be able to exploit them.[66] A standard argument in the vertical restraints literature points to the role of foreclosure through exclusive dealing arrangements as a means of raising rivals' costs. Suppose that there are economies of scale and scope in distribution. In this case, a system of exclusive dealers raises the distribution costs of smaller firms by more than it raises the distribution costs borne by larger firms. Thus, smaller firms are put at a disadvantage by this industry configuration. The net effect may be to raise the profits of larger firms, even though their costs of distribution are raised as well.

In a similar vein, intuition supports the idea that by demanding an exclusive dealing arrangement, a manufacturer with a large market share may be able to tie up the good dealers (e.g. retailers with locational advantages or who enjoy the greatest economies of scope) – if a dealer has to go with a single brand, he will tend to choose the market leader ceteris paribus.

Mathewson and Winter (1987) examined a model constructed to capture this intuition formally.[67] In their model, two manufacturers of substitute goods compete for the sales services of a single potential dealer.[68] There is a two-stage

[65] Raising the costs of both actual and potential rivals is discussed more generally in Chapter 8 by Richard Gilbert in this Handbook.

[66] In addition to coining the term, "raising rivals' costs", Salop and Scheffman (1983) provide a general analysis of a dominant firm's incentives to increase asymmetrically the costs of firms in a competitive fringe.

[67] Comanor and Frech (1985) also developed a model that attempted to embody this intuition. In their model, there is one firm that has a product differentiation advantage (in terms of the desirability of its brand to final consumers) over a competitive fringe of rival manufacturers. There are two types of retailer, one with low costs, one with high. Comanor and Frech claimed to show that the manufacturer with a differentiation advantage would impose exclusive dealing on the low-cost dealers. Schwartz (1987) argued convincingly that Comanor and Frech failed to incorporate fully the constraints on this manufacturer imposed by dealer rationality. Given the nature of their assumptions (in particular the claim that dealer margin should be taken as fixed), it is difficult to assess what the results of their model are.

[68] In terms of the intuition above, there is a second dealer with arbitrarily high costs.

game. In the first stage, each manufacturer states whether he will demand exclusive dealing. In the second stage, each manufacturer submits a bid to the dealer in the form of a uniform wholesale price (i.e. franchise fees are not allowed). If neither manufacturer has demanded an exclusive relationship, then the dealer purchases both intermediate products. If either manufacturer does demand an exclusive arrangement, then the dealer must choose from which of the two manufacturers to buy. Mathewson and Winter showed by example that if the demand for one of the intermediate goods is larger than for the other, then the manufacturer of the preferred product may demand an exclusive arrangement.

4.4. Entry deterrence

Raising *potential* rivals' costs also can be profitable. Incumbent manufacturers typically will see their profits depressed by entry and, thus, will seek to minimize the threat of entry by maximizing its cost and difficulty. In many ways, raising the costs of actual and potential rivals are similar. The main reason for treating entry deterrence separately is that there is a temporal asymmetry among the rival firms – the incumbents are in the market and the potential entrants are not. Thus, the incumbents may possess first-mover advantages. There also may be systematic differences in the sizes of entrants and incumbents. One typically would expect an entrant initially to have a lower market share than the incumbents.

Consider a single incumbent facing a single potential entrant in a market in which a dealer must have units of two inputs in order to produce final output. If the incumbent ties the sales of the two products, then the potential entrant will find it to be unprofitable to come into the market and sell only one of the two intermediate goods. The entrant will be forced to enter both markets. If there are substantial fixed and sunk costs of entering the two markets simultaneously, then the risk of entry is increased. Bork (1954) argued that the need to raise capital for entry into two markets simultaneously would not make entry more difficult since the potential reward from entry would be commensurately increased. Williamson (1979) argued that, when the profitability of entry depends upon specialized knowledge or skills that are not readily observable, a potential entrant's lack of experience in a given market may increase the firm's cost of raising capital. The threat of entry would thus be reduced if, in order to come into the target market, the entrant also had to come into a market with which he was unfamiliar.

Similar effects may arise when the "tying" occurs over time. Suppose that a firm has a monopoly or near monopoly today, but that it fears entry in the future. The firm may have incentives to reach long-term requirements contracts with consumers in order to foreclose the market to potential entrants. If all of the contracts came up for renewal on the same date, entry might be delayed, but not

prevented. The potential entrant could come in on the common contract expiration date without being at a disadvantage due to the contracts. Of course, the entrant would have to wait for the contracts to expire. When the dates are staggered and there are large fixed costs of entry, requirements contracts may block entry forever. In the presence of the staggered contracts, the entrant is able to compete for only a small portion of the total business at any one time. Given the large fixed costs of entry, it may not be profitable to go after demand in bits and pieces.

There also may be a foreclosure motivation behind exclusive dealing. The incumbent manufacturer may be able to take advantage of the temporal asymmetry by tying-up all of the top-notch retailers. Or, as discussed above, when there are economies of scale and scope in distribution, a system of exclusive dealers will raise the distribution costs of smaller firms (e.g. recent entrants) by more than it will raise the distribution costs borne by larger firms. By making small scale entry unprofitable, the incumbent (1) raises the financial risk (i.e. sunk costs) of entry; (2) makes it more credible that the incumbent will not accommodate an entrant (the entrant has to get a large market share in order to survive and thus is a greater threat to the incumbent); and (3) makes it costlier to enter since growth typically takes time.

A key objection raised by many analysts [e.g. Bork (1978) and Posner (1976)] is this: Why would a dealer sign a contract that lowers the probability of entry and lessens competition among suppliers? If there are many dealers, the answer is clear: each one may think that his individual signing decision has no effect on the likelihood of entry, and that actions by other dealers will block entry completely. Aghion and Bolton (1987) showed that a rational dealer may sign an entry-limiting contract even when the number of dealers is small. The contract that they examined has the following features. It is a long-term requirements contract that specifies the penalty to be levied on the dealer for breach of contract. This penalty is set so that an efficient entrant would find it profitable to offer the dealer such a low price that the dealer would find it profitable to breach the contract. In effect, the manufacturer–dealer contract forces the entrant to pay an entry fee that is then split between the manufacturer and dealer, giving the dealer incentives to sign the entry-limiting contract.[69]

In addition to blocking entry into a market that it already monopolizes, a firm may use its market power in one input market to block entry into another one. There is a long tradition in the vertical restraints literature of asking whether a manufacturer would use a tying contract to "leverage" a monopoly position in one input market into a second monopoly position in another input market. The prevailing view, due in large part to the efforts of members of the "Chicago

[69]This is another example of the use of the manufacturer–dealer contract as a form of precommitment.

school" of antitrust, is that a monopolist would not find it profitable to engage in leveraging.[70]

In a recent paper, Whinston (1987) showed that an incumbent could, in fact, find it profitable to use bundling to engage in what Whinston calls "strategic foreclosure". The basic idea can be seen in the following example. There are two inputs to downstream production, x and y. Initially, both are supplied by an upstream multi-product monopolist. Entry into production of x is completely blocked, say by patent protection. There is a single potential entrant into production of y, however. If the potential entrant choose to come in, the two firms play a differentiated products pricing game. Absent a bundling arrangement, there may be no way for the incumbent to deter entry into this market; the entrant may rationally predict that the incumbent would accommodate entry into the production of y. Suppose, however, that the incumbent can engage in bundling and set a package price for one unit of x and one unit of y that is lower than the sum of the price of x and the price of y. Given the proper demand conditions, this bundling has the effect of shifting downward the incumbent's (price–space) reaction curve in the y market to the point that entry becomes unprofitable. Intuitively, bundling can make the incumbent more aggressive in the y market because every additional unit of y sold as part of a bundle leads to greater profits from the sale of x as well. The ability to bundle may thus give the incumbent a credible threat to use against a potential entrant. Hence, a manufacturer may, in fact, use bundling to "leverage" his monopoly in one market into a monopoly in another one when the second market would otherwise have had a less-than-perfectly-competitive structure.

In closing this section, it should be noted that while incumbents would like to use vertical restraints to restrict entry, vertical restraints also may have indirect effects that encourage it. In particular, as noted several times earlier, some contract provisions may serve to increase product differentiation. This increased differentiation may make it easier for the entrant to find a market niche. This type of incentive is an example of what Fudenberg and Tirole (1984) amusingly have labelled the "fat-cat effect".

5. Welfare analysis

5.1. The two wedges

Heretofore, the analysis has focused on the profitability of vertical restraints. What about the welfare effects? An approach to welfare analysis taken by much of the literature is to focus on a specific practice and construct a list of reasons

[70]See Bowman (1957) for an early expression of this opinion. Whinston (1987) provides a discussion of the recent dominance of the no-leverage view.

why it may raise or lower welfare. I will take a different tack to frame the survey. Rather than looking directly at the welfare effects of a given practice, I will compare the private incentives to implement a vertical contract provision with the social ones.

Taking total surplus as the social maximand, there are two sources of divergence, or wedges, between a manufacturer's contract design incentives and the social incentives.[71,72] First, a manufacturer ignores any effects that his actions have on the level of consumer surplus. Viewed another way, activities that transfer surplus from consumers to the firm generate a private benefit for which there is no corresponding social benefit.

The second wedge is the difference between firm and industry profits. A manufacturer ignores the effects that his contract choice has on the profits of other firms, including his dealers and rival manufacturers and their dealers. The presence of these wedges may distort the manufacturer's incentives with respect to both the price and nonprice (e.g. promotional effort and product quality) decisions that he induces his dealers to make.

5.2. *Price distortions*

Consider a market in which all of the customers of any given manufacturer must be served at the same uniform price (i.e. no price discrimination is feasible). Since lower prices raise consumer surplus, there is a tendency for the manufacturers to do too little to keep downstream prices low. The standard monopolistic output restriction is one consequence of this fact. Turning to the issues at hand, the consumer surplus wedge also affects the use of vertical restraints. Suppose, for example, that the manufacturer must incur administrative costs to impose a resale price ceiling (e.g. the cost of monitoring the dealers' prices). To the extent that the price ceiling benefits consumers, the manufacturer's private incentives to impose the ceiling are lower than the social incentives.[73] Consequently, the manufacturer may fail to implement such a scheme even when imposition of the profit-maximizing price ceiling would raise welfare. When a given vertical restraint leads to higher final prices and lower consumer surplus, the bias runs in the opposite direction. For example, when input substitution is possible, the

[71] The welfare analysis of vertical restraints is, like all welfare analysis in the field of industrial organization, merely an application of the Fundamental Theorem of Industrial Organization. For a more general treatment of the Fundamental Theorem, see Katz and Shapiro (1987).

[72] Although the discussion is framed in terms of the manufacaturer's contract design incentives, it is clear that analogous effects arise in cases in which the dealer proposes the contract.

[73] Interestingly, Mathewson and Winter (1983a) showed that a resale price *floor* also could lower the prices charged by some dealers. This effect arises when the equilibrium absent the price floor would entail price dispersion and the price floor causes this asymmetric equilibrium to collapse. In this case, there would be a similar bias against a price floor.

assignment of exclusive territories may allow the manufacturer profitably to raise the price charged to final consumers.

This line of argument can be applied to the analysis of price discrimination as well.[74] To the extent that price discrimination lowers consumer surplus, there tends to be too much of it. Moreover, the manufacturer's incentives to impose vertical restraints (such as exclusive territories) that facilitate discrimination are socially excessive. There are other cases, however, in which profit-maximizing price discrimination raises consumer surplus in comparison with profit-maximizing uniform pricing. In these cases, the bias runs in the opposite direction.[75]

5.3. Nonprice distortions

In many instances, the effect of vertical restraints is to stimulate promotional activities and other forms of nonprice competition. Changes in the level of nonprice activities may affect welfare both directly and indirectly: directly because consumers may value the activities, and indirectly because changes in nonprice activities may affect the nature of price competition.

5.3.1. Promotional efforts

The effects of promotional activities have been viewed in several different lights by the literature. Much of the traditional debate centers on the question of whether promotional activities should be viewed as the provision of information (assumed to be good), or as a way to create false image differentiation (assumed to be bad). Given the unsettled nature of this field, the reader will not be surprised to find that the welfare analysis of the effects of vertical restraints on the provision of promotional activities is itself something of a mess. Rather than debate the merits of various labels, here I will try to catalogue (at least some of) the effects of promotional activities.

First, suppose that the advertising conveys false information that has the effect of shifting out the demand curve faced by the advertising firm. This case was examined by Dixit and Norman (1978) for a final good market in which all inputs are supplied at constant prices by perfectly competitive firms. For the reason already discussed, Dixit and Norman found that to the extent that advertising leads to increased prices, manufacturers tend to have overly high incentives to induce advertising. It does not, however, follow that all false

[74] Romer (1986) applied this reasoning to the case of a final good monopoly under the presumption that discrimination lowers consumer surplus.

[75] As Hal Varian makes clear in Chapter 10 of this Handbook, models of discrimination under oligopoly and monopolistic competition still are in their infancy, and more analysis is needed before reaching even tentative conclusions.

advertising lowers welfare. The increased demand due to the false advertising may counteract the usual monopolistic or oligopolistic output restriction, leading to an increase in allocative efficiency.

Comanor (1968) examined markets in which there are multiple brands and promotional efforts are aimed at convincing consumers that different brands are not good substitutes for one another, even though they in fact are. This "image differentiation" lowers between-brand cross-elasticities and reduces price competition. The net effect is to lower consumer surplus. Thus, the firms' private benefits may exceed the social benefits, and firms may devote excessive resources to advertising. Moreover, the price increases may lead to reductions in the quantities sold, further lowering total surplus.

Promotions need not be misleading to have adverse effects on the degree of interbrand price competition. Dixit and Norman's results carry over to advertising that provides truthful information. Similarly, promotional activities that reveal true product differences to consumers may have some of the same adverse effects as activities that create false image differentiation. As Salop and Stiglitz (1978) pointed out, the provision of product information need not benefit consumers. To the extent that this information reveals true product differentiation, it may reduce interbrand price competition and lower consumer surplus. Of course, the information may allow consumers to make better matches, or product choices. The net effect of information provision may be to raise or lower consumer surplus.

Information provision by one firm also has ambiguous effects on the profits of rival firms. These firms may benefit from the increased product differentiation, even if the differentiation is vertical (i.e. the promoting firm reveals that his product is of higher quality). It follows that information may be over- or under-supplied by the market. Hence, vertical restraints that lead to greater promotional activities may ameliorate or exacerbate the problem.

Fisher and McGowan (1979) argued that some promotional efforts do more than inform consumers about existence – they actually make consumers better able to derive benefits from consumption. For example, the dealer may provide information as to how to use the product. When the promotional activities are of direct value to consumers, one can then think of the degree of promotion as an aspect of product quality, to which the discussion now turns.

5.3.2. The input mix: Product quality and productive efficiency

The dealer's choice of input mix may affect consumers in two ways. First, the input mix affects costs. Consumers care about this effect only to the extent that it influences product pricing. Second, when the input mix is a dimension of product quality, consumers care about the input mix directly.

Spence (1975) showed that a monopolist may over- or underprovide quality. The reason is a simple one. When the marginal consumer values quality by more than the average inframarginal consumer, the monopolist can raise price and quality (holding unit sales constant) in a way that lowers consumer surplus. Hence, in this case the monopolist tends to oversupply quality given his output level. The direction of the bias is reversed when the marginal consumer values quality by less than does the average inframarginal consumer. Not surprisingly, given that the direction of distortion in quality is ambiguous, it is difficult to say whether vertical restraints that promote the provision of product quality are socially beneficial.

In cases in which consumers are indifferent between a family of input mixes (i.e. there are many ways to produce output of a given quality level) sophisticated pricing such as quantity-dependent pricing, ties, and royalties tend to have beneficial welfare effects by correcting inefficient input choices by dealers for a given quality level.

5.3.3. The number of dealers and product variety

The number of dealers affects welfare both as a component of service itself and through its indirect effects on the provision of other dimensions of service. Unfortunately, there are two reasons why it is difficult to assess the welfare effects that vertical contracts have through their effects on the number of dealers.

First, in some instances, a given vertical restraint may be used by a manufacturer to either increase or decrease the number of dealers in comparison with the outcome under which that provision is proscribed.[76] The second problem is that even if one knows the effect of a given vertical practice on the equilibrium number of dealers, the sign of the welfare change is not immediate. There is a well-developed literature analyzing the product selection problem in a final good industry with competitively supplied inputs. This literature demonstrates that, in general, there may be excessive or insufficient variety under either monopoly [e.g. Katz (1980)] or monopolistic competition [e.g. Spence (1976), Dixit and Stiglitz (1977), and Salop (1979)]. In the light of the fact that vertical restraints can raise or lower variety, there is a strong need for research as to whether the biases tend to move together systematically.

5.4. What does all of this have to do with public policy?

Much of the literature on vertical restraints has been conducted with the express aim of deriving policy conclusions. But in many, if not most, instances there is no

[76] Recall the discussion of exclusive territories and franchise fees in Subsection 3.3.

widespread agreement on whether a particular practice is socially beneficial or harmful. This unhappy state of affairs is due, in part, to the fact that all of the practices can be beneficial in some instances and harmful in others, and it may be extremely difficult to distinguish between the two cases.

To date, much of the policy debate has been conducted in terms of asking whether a given practice is pro- or anticompetitive, and thus whether a given practice should be banned or not. I believe that this approach is misguided on at least two counts. First, welfare effects may have nothing to do with "competition", per se. Some vertical restraints (e.g. tying) can arise, and have welfare consequences, even when both levels of the supply chain are monopolized. Second, the type of answer sought is the wrong one. On theoretical grounds, at least, any given restraint may be good or bad. Theory alone (or theory coupled with a handful of examples) is not going to answer a question of the form: Should resale price maintenance be per se illegal?

Confronted with ambiguous welfare effects, the courts and many economists have called for a rule of reason approach. But a fully general rule of reason is likely to be more than the courts can handle.[77] The key to policy design is to develop workable rules by which to identify observable market conditions under which given practices are socially desirable. In attempting to accomplish this goal, the courts and many economists [e.g. Posner (1977) and Williamson (1979)] typically have focused on the degree of market concentration as the critical condition. This aspect of market structure is concentrated on in the belief that a vertical agreement is unlikely to have "anticompetitive" effects when it is among parties having a low combined market share either upstream or down.

While this work is a beginning, it is essential that the analysis be expanded to include factors such as the market's information structure, risk characteristics (e.g. size of sunk investments), and the degree to which parties become locked-in to one another (e.g. the amount of transactions-specific capital). It also is essential that this work identify means of isolating the effects of vertical restraints in comparison with alternative institutional and contractual arrangements.

5.4.1. The proper benchmark

The measurement of both the social benefits and the social costs of a vertical practice depend critically on the choice of benchmark, as can be illustrated by the analysis of resale price maintenance. Consider, first, the social benefits of resale price maintenance. As discussed in Subsection 3.1.2.1 earlier, it has been argued that resale price maintenance is socially beneficial because it leads to increased equilibrium levels of service and promotional activities. One problem with this

[77]See Posner (1981) for a discussion of some of the problems with the rule of reason approach, both as developed by the courts and more generally.

argument is that it takes as given the fact that increased promotion is socially beneficial. Abstracting from that problem, there is another difficulty. As Comanor (1968) and Yamey (1966) argued, before concluding that vertical restraints are essential to ensure the provision of promotional activities, one must examine whether there are alternative institutional arrangements by which the manufacturer could ensure the supply of activities subject to externalities across dealers.

One approach is to have the manufacturer supply the services himself. But asymmetric information about downstream market conditions (e.g. the dealer may be better informed about the downstream market than is the manufacturer) may limit the profitability of this approach. Moreover, even if the manufacturer can observe downstream conditions, so-called antidiscrimination policy [e.g. Sections 2(d) and 2(e) of the Robinson–Patman Act] limits the manufacturer's ability to vary the level of promotional support across dealers in response to differing local market conditions.[78] Finally, when effort is one of the inputs to downstream promotion, the upstream decision-maker may not have the time to conduct downstream promotional activities in addition to carrying out his other tasks (i.e. the manufacturer may have a higher opportunity cost of his time than does the dealer).

The manufacturer could finance downstream promotional activities directly, while letting the dealers choose the direction and intensity of these efforts. For example, the manufacturer could pay dealers for maintaining showrooms. There may be problems with this sort of payment, however. First, if the dealer does not bear the costs, he may engage in excessive levels of these activities. Second, the dealer may defraud the manufacturer when the level of promotional activities is unobservable to the manufacturer. I know of one case in which a radio station sold false advertising receipts to dealers who in turn submitted the bills to the manufacturers for reimbursement. Another reason we do not observe such schemes more often in the United States may again be the Robinson–Patman Act, which prevents differential payments to retailers based on the promotional services (such as having good reputations) that the retailers provide the dealers.[79]

Comanor and other economists have suggested another form of financing at least some promotional activities – charge consumers for the services (such as technical advice from salespeople) that they receive. In this way, consumers who do not want the services do not have to pay for them, and those that want them do pay. Of course, this approach does not work when the promotional activities have a nonexcludable public good nature such as in the case of media advertising.

Where antitrust authorities allow it, the dealers themselves might attempt to internalize the externalities by forming cooperative ventures. For example, the

[78]Schwartz (1986) discusses the inefficiencies in distribution induced by Robinson–Patman Act more fully.

[79]This defect of the Robinson–Patman Act was noted by Bowman (1952, 1955) in his discussion of the private incentives to implement resale price maintenance. See also, Comment (1951).

Ford dealers in a single region might jointly advertise on television. An important issue is whether some dealers will choose to stay out of the cooperative venture and free ride. In response to such problems, one could imagine the manufacturer *requiring* his dealers to form cooperative ventures as one of the terms of the contract.

Lastly, the manufacturer might be able to implement a contract that made any one dealer's compensation a function of the collective performance of the manufacturer and all of his dealers. But to give each dealer full incentives, the manufacturer would have to give *each* dealer 100 percent of the *collective* increase in joint profits (of the manufacturer and all of his dealers) at the margin.[80]

This laundry list of alternatives demonstrates that it is not at all clear that resale price maintenance is essential to the provision of promotional services and other activities subject to free riding. There are many possible substitutes, although each has limitations. Theory alone does not tell us whether policy-makers' allowing resale price maintenance would lead to a great increase in service and promotional activities.

Comparative analysis also can be used to assess the potential social harm from a given practice. A standard argument put forth by proponents of resale price maintenance [e.g. Bork (1978)], for example, is the following one: resale price maintenance cannot have adverse welfare effects because the manufacturer could achieve the same effect by raising the wholesale price of the good. The usefulness of this approach is limited by that fact that this claim is correct only if: (a) downstream input substitution is infeasible; (b) the costs for all retailers to serve all customers are the same; and (c) all final consumers are equally sensitive to a dealer's price. When condition (a) is not satisfied, raising the wholesale price induces distortions in the downstream input mix. When conditions (b) and (c) are not satisfied, absent resale price maintenance different consumers would face different prices. There is no way that the manufacturer can force retailers to charge all customers the same price solely through raising the wholesale price.

5.4.2. Consistency of public policy

As the analysis of Sections 2 through 4 makes clear, various vertical restraints often are (partial) substitutes for one another. For example, in some settings, resale price maintenance and territorial restrictions are both ways to maintain high dealer margins. The fact that two practices may have many of the same

[80] Many of the issues that arise in the design and implementation of such schemes are discussed in the principal–agent literature. See, for example, Holmstrom (1982).

effects as one another suggests that the public policy treatment of them should be similar in many cases.[81]

There also is a need for consistency between the treatment of vertical restraints and refusals to deal. A refusal to deal occurs when a manufacturer refuses to allow a retailer to act as a dealer for the manufacturer's product on the same terms as are granted other retailers. There may be cases in which a refusal to deal is merely the result of the manufacturer's recognizing that a particular dealer is unscrupulous or is unlikely to meet his responsibilities to the manufacturer. But a refusal to deal also can serve as a means of enforcing an implicit vertical restraint. For example, de facto resale price maintenance could be imposed by refusing to deal with any discounters. Public policy towards refusal to deal should follow from the treatment of the other vertical restraints. Unfortunately, it may be very difficult to distinguish between a refusal to deal where the manufacturer is keeping out a dealer who would otherwise behave in inefficient ways and one where the manufacturer is keeping out a dealer who would act in economically efficient ways but contrary to the private interests of the manufacturer.

In addition to being internally consistent, public policy toward vertical restraints also should be consistent with other antitrust policy. For example, to the extent that resale price maintenance is an attempt to make the monitoring of prices easier, resale price maintenance and price information sharing should be subject to the same treatment.

At present, there is one particularly striking incongruity in U.S. antitrust policy. Integrated firms are allowed to implement almost any contract internally. Thus, antitrust laws that restrict independent firms' writing of arm's length contracts may have the effect of encouraging vertical integration even in cases in which this form of business organization is not the most efficient one.[82] For this reason, a public policy that prevents pernicious vertical practices by unintegrated firms may actually be worse than a policy that allows these practices – to the extent that vertical integration is possible, these practices (or their equivalent) may take place in any event, and the costs of production and distribution under vertical integration may be higher due to the use of an inefficient organizational form.

The need for consistency also arises in the case of price discrimination. A principle of consistency would argue for vertical restraints that are motivated solely by discrimination to be treated similarly to an act of price discrimination itself. One way to ensure consistency is to treat all instances of price discrimina-

[81] One does not want to make too much of the similarity of effects of different practices. As shown in earlier sections, the identity between two practices never is exact, and theory does not support the view that two different practices should be treated identically in all markets.

[82] Although they subscribed to the "unnatural union" view of vertical integration, Blair and Kaserman (1978) made a similar point for the particular case of a tying contract.

tion under price discrimination laws, rather than relying on proscriptions against certain vertical practices.

This last point brings up another important one. In addition to recognizing the need for consistency in public policy, policy makers should recognize that in some instances, other antitrust laws may be substitutes for strictures against vertical restraints. Laws covering price discrimination, for instance, may be sufficient to deal with that problem directly, so there is no need to prohibit exclusive territories for fear that they facilitate price discrimination.

Finally, it should be noted that antitrust laws may themselves motivate the use of vertical restraints. For example, as discussed in the previous subsection, resale price maintenance may be a response to the fact that a manufacturer cannot make differential payments to dealers even though some dealers provide valuable (and costly) services to the manufacturer that others do not. These effects should be taken into account when setting antitrust policy.

References

Abreu, D., Milgrom, P. and Pearce, D. (1987) 'Information and timing in repeated partnerships', unpublished draft, Harvard University.

Aghion, P. and Bolton, P. (1987) 'Contracts as a barrier to entry', *American Economic Review*, 77:388–401.

Atkinson, A.B. and Stiglitz, J.E. (1980) *Lectures in public economics*. New York: McGraw-Hill.

Banks, J.S. and Sobel, J. (1987) 'Equilibrium selection in signaling games', *Econometrica*, 55:647–661.

Bernheim, B.D. and Whinston, M.D. (1985) 'Common marketing agency as a device for facilitating collusion', *Rand Journal of Economics*, 16:269–281.

Bittlingmayer, G.A. (1983) 'A model of vertical restriction and equilibrium in retailing', *Journal of Business*, 50:477–496.

Blair, R.D. and Kaserman, D.L. (1978) 'Vertical intergration, tying, and antitrust policy', *American Economic Review*, 68:397–402.

Blair, R.D. and Kaserman, D.L. (1984) 'Tying arrangements and uncertainty', *Research in finance*, Supplement I, *Management under government intervention: A view from Mount Scopus*.

Bolton, P. and Bonanno, G. (1987) 'Vertical restraints in a model of vertical differentiation', unpublished manuscript, Harvard University.

Bork, R.H. (1954) 'Vertical integration and the Sherman Act: The legal history of an economic misconception', *University of Chicago Law Review*, 22:157–201.

Bork, R.H. (1978) *The antitrust paradox: A policy at war with itself*. New York: Basic Books.

Bowman, Jr., W.S. (1952) 'Resale price maintenance – A monopoly problem', *Journal of Business*, 25:141–155.

Bowman, Jr., W.S. (1955) 'Prerequisites and effects of resale price maintenance', *University of Chicago Law Review*, 22:825–873.

Bowman, Jr., W.S. (1957) 'Tying arrangements and the leverage problem', *Yale Law Journal*, 67:19–36.

Brander, J.A. and Spencer, B.J. (1985) 'Export subsidies and international market share rivalry', *Journal of International Economics*, 18:83–100.

Brennan, T.J. and Sheldon, K. (1983) 'Joint production and monopoly extension through tying', discussion paper 84-1, Economic Policy Office, U.S. Department of Justice.

Burstein, M.L. (1960a) 'A theory of full-line forcing', *Northwestern University Law Review*, 55:62–95.

Burstein, M.L. (1960b) 'The economics of tie-in sales', *Review of Economics and Statistics*, 42:68–73.
Caves, R.E. (1986) 'Vertical restraints in manufacaturer–distributor relations: Incidence and economic effects', in : R.E. Grieson, ed., *Antitrust and regulation*. Lexington: Lexington Books.
Caves, R.E. and Murphy, W.F. (1976) 'Franchising: Firms, markets and intangible assets', *Southern Economic Journal*, 42:572–586.
Comanor, Jr., W.S. (1968) 'Vertical territoriality and customer restrictions: White Motor and its aftermath', *Harvard Law Review*, 81:1419–1438.
Comanor, Jr. W.S. and Frech, III, H.E. (1985) 'The competitive effects of vertical agreements?', *American Economic Review*, 75:539–546.
Comment (1951) 'Resale price maintenance and the antitrust laws', *University of Chicago Law Review*, 18:369–380.
Demsetz, H. (1968) 'Why regulate utilities?', *Journal of Law and Economics*, 11:55–65.
Director, A. and Levi, E.H. (1956) 'Law and the future: Trade regulation', *Northwestern University Law Review*, 61:281–296.
Dixit, A.K. (1979) 'Quality and quantity competition', *Review of Economic Studies*, 46:587–599.
Dixit, A.K. (1983) 'Vertical integration in a monopolistically competitive industry', *International Journal of Industrial Organization*, 1:63–78.
Dixit, A.K. and Norman, V. (1978) 'Advertising and welfare', *Bell Journal of Economics*, 9:1–17.
Dixit, A.K. and Stiglitz, J.E. (1977) 'Monopolistic competition and optimal product diversity', *American Economic Review*, 67:297–308.
Farrell, J. and Shapiro, C. (1987) 'Optimal contracts with lock-in', Department of Economics, University of California, Berkeley, working paper 87-58.
Fershtman, C. and Judd, K.L. (1986) 'Strategic incentive manipulation and the principal–agent problem', unpublished manuscript, Northwestern University.
Fershtman, C. and Judd, K.L. (1987) 'Equilibrium incentives in oligopoly', *American Economic Review*, 77:927–940.
Fershtman, C., Judd, K.L. and Kalai, E. (1987) 'Cooperation through delegation', research report no. 163, Hebrew University of Jerusalem.
Fisher, F.M. and McGowan, J.J. (1979) 'Advertising and welfare: Comment', *Bell Journal of Economics*, 10:726–727.
Fudenberg, D. and Tirole, J. (1984) 'The fat-cat effect, the puppy-dog ploy, and the lean and hungry look', *American Economic Review*, 74:361–366.
Gallick, E.C. (1984) 'Exclusive dealing and vertical integration: The efficiency of contracts in the tuna industry', Bureau of Economics staff report to the Federal Trade Commission.
Gould, J.R. and Preston, L.E. (1965) 'Resale price maintenance and retail outlets', *Economica*, 32:302–312.
Grossman, S. and Hart, O.D. (1986) 'The costs and benefits of ownership: A theory of vertical and lateral integration', *Journal of Political Economy*, 94:691–719.
Harris, M. and Raviv, A. (1979) 'Optimal incentive contracts with imperfect information', *Journal of Economic Theory*, 20:231–259.
Holmstrom, B. (1979) 'Moral hazard and observability', *Bell Journal of Economics*, 10:74–91.
Holmstrom, B. (1982) 'Moral hazard in teams', *Bell Journal of Economics*, 13:324–340.
Katz, M.L. (1980) 'Multiplant monopoly in a spatial market', *Bell Journal of Economics*, 11:519–535.
Katz, M.L. (1988a) 'Game-playing agents: Unobservable contracts as precommitments', unpublished draft, Berkeley, University of California.
Katz, M.L. and Rosen, H.S. (1985) 'Tax Analysis in a oligopoly model', *Public Finance Quarterly*, 13:3–20.
Katz, M.L. and Shapiro, C. (1987) 'The two wedges: The fundamental theorem of industrial organization', unpublished manuscript, Princeton University.
Marvel, H.P. (1982) 'Exclusive dealing', *Journal of Law and Economics*, 25:1–25.
Marvel, H.P. and McCafferty, S. (1984) 'Resale price maintenance and quality certification', *Rand Journal of Economics*, 15:346–359.
Mathewson, G.F. and Winter, R.A. (1983a) 'The incentives for resale price maintenance under imperfect information', *Economic Inquiry*, 21:337–348.

Mathewson, G.F. and Winter, R.A. (1983b) 'Vertical intergration by contractual restraints in spatial markets', *Journal of Business*, 50:497–517.
Mathewson, G.F. and Winter, R.A. (1984) 'An economic theory of vertical restraints', *Rand Journal of Economics*, 15:27–38.
Mathewson, G.F. and Winter, R.A. (1987) 'The competitive effects of vertical agreements: Comment', *American Economic Review*, 77:1057–1068.
Milgrom, P.R. and Weber, R.J. (1982) 'A theory of auctions and competitive bidding', *Econometrica*, 50:1089–1122.
Myerson, R.B. (1983) 'Mechanism design by an informed principal', *Econometrica*, 51:1767–1797.
Oi, Walter, (1971) 'A Disneyland dilemma: Two-part tariffs for a Mickey Mouse monopoly', *Quarterly Journal of Economics*, 85:77–96.
Perry, M.K. and Porter, R.H. (1986) 'Resale price maintenance and exclusive territories in the presence of retail service externalities', unpublished manuscript, Bell Communications Research, Inc.
Posner, R.A. (1976) *Antitrust law: An economic perspective*. Chicago: University of Chicago Press.
Posner, R.A. (1977) 'The rule of reason and the economic approach: Reflections on the *Sylvania* decision', *University of Chicago Law Review*, 45:1–20.
Posner, R.A. (1981) 'The next step in antitrust treatment of restricted distribution: Per se legality', *University of Chicago Law Review*, 48:6–26.
Preston, L.E. (1965) 'Restrictive distribution arrangements: Economic analysis and public policy standards', *Law and Contemporary Problems*, 30:506–529.
Rey, P. and Stiglitz, J.E. (1986) 'The role of exclusive territories in producers' competition', unpublished draft, Princeton University.
Rey, P. and Tirole, J. (1986) 'The logic of vertical restraints', *American Economic Review*, 76:921–939.
Romer, D. (1986) 'Costly price discrimination', unpublished draft, Princeton University.
Ross, T.W. (1987) 'When sales maximization is profit-maximizing: A two-stage game', working paper CIORU 87-01, Carleton University.
Rubinstein, A. (1985) 'A bargaining model with incomplete information about time preferences', *Econometrica*, 53:1151–1172.
Salop, S.C. (1979) 'Monopolistic competition with outside goods', *Bell Journal of Economics*, 10:141–156.
Salop, S.C. and Scheffman, D.T. (1983) 'Raising rivals' costs', *American Economic Review*, 73:267–271.
Salop, S.C. and Stiglitz, J.E. (1978) 'Information, welfare, and product diversity', unpublished manuscript, Princeton University.
Schelling, T.C. (1956) 'An essay on bargaining', *American Economic Review*, 46:281–306.
Schelling, T.C. (1960) *The strategy of conflict*. New York: Oxford University Press.
Schwartz, M. (1986) 'The perverse effects of the Robinson–Patman Act', *The Antitrust Bulletin*, 31:733–757.
Schwartz, M. (1987) 'The competitive effects of vertical agreements: Comment', *American Economic Review*, 77:1063–1068.
Seade, J. (1985) 'Profitable cost increases and the shifting of taxation: Equilibrium responses of markets in oligopoly', unpublished manuscript, Warwick University.
Shaked, A. and Sutton, J. (1982) 'Relaxing price competition through product differentiation', *Review of Economic Studies*, 40:3–14.
Shavell, S. (1979) 'Risk sharing and incentives in the principal and agent relationship', *Bell Journal of Economics*, 10:55–73.
Shleifer, A. (1985) 'A theory of yardstick competition', *Rand Journal of Economics*, 16:319–327.
Silcock, T.H. (1938) 'Some problems of price maintenance', *Economic Journal*, 48:42–51.
Spence, A.M. (1975) 'Monopoly, quality, and regulation', *Bell Journal of Economics*, 6:417–429.
Spence, A.M. (1976) 'Product selection, fixed costs, and monopolistic competition', *Review of Economic Studies*, 43:217–235.
Spencer, B.J. and Brander, J.A. (1983) 'International R & D rivalry and industrial strategy, *Review of Economic Studies*, 50:707–722.
Spengler, J.J. (1950) 'Vertical integration and antitrust policy', *Journal of Political Economy*, 58:347–352.

Taussig (1916) 'Price maintenance', *American Economic Review Papers and Proceedings*, 6:170–184.
Telser, L.G. (1960) 'Why should manufacturers want fair trade?', *Journal of Law and Economics*, 3:86–105.
Telser, L.G. (1965) 'Abusive trade practices: An economic analysis', *Law and Contemporary Problems*, 30:488–505.
Whinston, M.D. (1987) 'Tying, foreclosure, and exclusion', discussion paper number 1343, Harvard Institute of Economic Research.
White, L.J. (1981) 'Vertical restraints in antitrust law: A coherent model', *The Antitrust Bulletin*, 26:327–345.
Williamson, O.E. (1979) 'Assessing vertical market restrictions: Antitrust ramifications of the transactions cost approach', *University of Pennsylvania Law Review*, 127:953–993.
Williamson, O.E. (1983) 'Credible commitments: Using hostages to support exchange', *American Economic Review*, 73:519–540.
Yamey, B.S. (1952) 'The origins of resale price maintenance: A study of three branches of retail trade', *The Economic Journal*, 62:522–545.
Yamey, B.S. (1954) *The economics of resale price maintenance*. London: Sir Isaac Pitman.
Yamey, B.S., ed. (1966) *Resale price maintenance*. Chicago: Aldine Publishing Company.