RAND Journal of Economics
Vol. 30, No. 4, Winter 1999
pp. 736–757

Consumer strategies limiting the monopolist's power: multiple and joint purchases

Ingela Alger*

I characterize the menu of bundles (price-quantity combinations) offered by a monopolist when consumers can buy several bundles, share bundles with others, or do both, in a two-type setting. I find that although perfect arbitrage prevents any price discrimination, partial arbitrage in the form of multiple or joint purchases may actually lead to more pronounced price discrimination than when consumers can only pick one single bundle. Further, clear predictions emerge for the price pattern, contrasting with the existing literature: with multiple purchases only, the firm offers strict quantity discounts; with joint purchases only, discounts are infeasible.

1. Introduction

■ Firms often market their goods by offering a menu of choices: consumers may choose between goods with different quality, or between packages containing different quantities of the same good. It allows the firm to price discriminate when demand is heterogeneous, but any individual consumer's demand is not observable. This practice, known as second-degree price discrimination, has been studied extensively in the literature (see, for instance, Mussa and Rosen, 1978; Maskin and Riley, 1984; and Wilson, 1992). Surprisingly, however, no model unambiguously predicts such a common phenomenon as quantity discounts, which always come at the price of restrictions on the parameters. The analysis in this article reveals that this shortcoming is due to the incomplete modelling of consumer behavior.

An assumption invariably found in previous articles is that consumers may only pick one single item, or bundle, from the menu offered by the firm. In reality, consumers often have access to other forms of arbitrage: for many goods, they can buy several bundles from the menu, share a bundle with other consumers, or do both. This article characterizes the price schedule that maximizes a monopolist's expected profit when

* Boston College; ingela.alger@bc.edu.

This article is based on a chapter of my Ph.D. thesis submitted to the Université des Sciences Sociales de Toulouse. I am grateful to Jean-Jacques Laffont for guidance, and to Guillermo Alger, Thomas Mariotti, Nour Meddahi, Patrick Rey, François Salanié, Susanna Sällström, Jean Tirole, Dimitri Vayanos, Wilfried Zantman, and seminar participants at the London School of Economics and the University of Virginia for useful comments. I am particularly indebted to Régis Renault for many valuable discussions. I also thank Editor Glenn Ellison and two anonymous referees for detailed and helpful comments. Financial support from The Royal Swedish Academy of Sciences (C C Söderströms fond), the Swedish Institute, and the Sixten Gemzéus foundation is thankfully acknowledged.

consumers have access to multiple and/or joint purchases, in a two-type setting. The predictions are clear: when only multiple purchases are possible, as for consumer goods, the firm offers strict quantity discounts; when only joint purchases are relevant, as for large consumer durables sold in units of different qualities, the good is sold with a quality premium or at a linear price; finally, when both joint and multiple purchases are relevant, as for clothes, neither quantity discounts nor premia are feasible.[1]

Although very satisfying, the strict quantity discount result is surprising: indeed, the common belief is that multiple purchases only make quantity premia infeasible. Other conjectures that have been put forward in the literature, and that the model allows us to analyze, include the following. Wilson makes the conjecture that quantity discounts should be sufficient to deter multiple purchases, as shown in the following excerpt from his book (1992, p. 71): "[subadditivity, which] means that the charge for several small purchases is no less than the charge for the purchase that is the sum of the smaller amounts, . . . prevents a customer (or arbitrageur) from circumventing the tariff by dividing a large purchase into several smaller ones." In spite of its intuitive appeal, the belief that quantity discounts should be sufficient to deter multiple purchases proves to be wrong, as will be seen shortly. Finally, the absence of multiple and joint purchases is generally seen as a necessary condition for the firm to be able to price discriminate. This is partially confirmed: indeed, unless consumers have access to both multiple and joint purchases, nonlinear pricing is feasible.

As a benchmark, I first analyze the classic single-bundle model. At the solution the high-valuation consumer is offered the socially optimal quantity and obtains a rent, whereas the low-valuation consumer obtains no rent and consumes a suboptimal quantity. As mentioned above, a feature of the single-bundle solution is that it may yield either quantity discounts or quantity premia, depending on the parameter values.

As expected, the single-bundle solution is not robust to multiple and/or joint purchases. Surprisingly, however, quantity discounts are not sufficient to prevent multiple purchases in the two-type setting. The key to understanding this is that through multiple purchases, the high-valuation consumer is not restricted to "dividing a large purchase into several smaller ones" as Wilson suggests—he can also choose to consume more or less than if he chose the bundle meant for him. Thus consumers do not arbitrage only to get a lower price, but also to influence the quantity they consume. This is a consequence of the firm offering price-quantity combinations, as opposed to setting the price and letting the consumers choose the quantity; indeed, this enables the firm to offer bundles that are above the demand curve. In other words, the average unit price paid by a consumer exceeds his marginal utility; given that price, he would therefore consume less. With multiple purchases, the high-valuation consumer may consume less by buying several small bundles; this may give him a higher utility than buying the large bundle, even if he ends up paying a slightly higher unit price.

Next, I characterize the profit-maximizing price schedule when consumers may make multiple and joint purchases, analyzing each type of arbitrage separately. To avoid integer problems, I model multiple purchases as the possibility to purchase any real number above one of any bundle. Conversely, joint purchases are modelled as the ability to buy any real number between zero and one of any bundle.

When consumers can make multiple purchases only, the price schedule turns out to be qualitatively similar to the schedule in the single-bundle model: only the high-valuation consumer obtains a rent, and the quantity sold to the low-valuation consumer

[1] As these predictions indicate, the quantity and quality interpretations of the model do not apply equally well to the various settings I consider. When developing the general ideas below, for simplicity I refer only to the quantity interpretation.

is distorted downward compared to the socially efficient level. This similarity is related to the intuition developed above: in both the single-bundle and the multiple-purchase cases, only the high-valuation consumer may consume less than the bundle meant for him through arbitrage: in one case by buying the small bundle, in the other by buying several small bundles. Therefore, only the high-valuation consumer obtains a rent. And as in the single-bundle case, to limit this rent the monopolist distorts the quantity offered to the low-valuation consumer downward compared to the socially optimal level. The quantity may, however, be more or less distorted than in the single-bundle model. When it is less distorted, total welfare and the high-valuation consumer's utility is higher. In contrast, when the distortion is larger, total welfare is smaller, and the high-valuation consumer's utility may in fact be lower than in the single-bundle model. But the firm's expected profit is always smaller.

A striking result is that the amount of discrimination, measured by the difference in the unit price in the small and large bundles, may be higher than in the single-bundle model. This is at first sight counterintuitive: intuition would suggest that better consumer arbitrage reduces the firm's ability to price discriminate. Nevertheless, the result is quite natural: given that quantity discounts are not sufficient to prevent multiple purchases, it follows that a way to prevent them is to further increase the discount.

Finally, I confirm that quantity premia are not feasible when consumers can make multiple purchases. Surprisingly, however, strict quantity discounts are always optimal. It is surprising because the following argument seems natural: if it is optimal for the firm to impose a quantity premium at the single-bundle solution, then the infeasibility of quantity premia through multiple purchases should imply that the firm offers a linear price scheme, i.e., both types of consumers pay the same unit price. The intuition behind the strict discount result is as follows. Given a linear price and multiple purchases, the firm cannot force the high-valuation consumer to purchase a bundle that is not on his demand curve. Moreover, the price must exceed marginal cost (otherwise, the firm would not be a monopolist). Thus with a linear price, the high-valuation consumer would choose a socially suboptimal quantity. Clearly, then, the firm can increase its profit by increasing the quantity and decreasing the price in the large bundle: total welfare increases, whereas the high-valuation consumer's utility is unaffected.

Turning to the effects of joint purchases only, the results are qualitatively different compared to the single-bundle and the multiple-purchase cases. First, the low-valuation consumer obtains a strictly positive utility. As opposed to the single-bundle and multiple-purchase cases, the low-valuation consumer is here an active arbitrageur: he may indeed consume less than the small bundle by sharing a bundle with others. Being an active arbitrageur, he obtains a rent. Second, joint purchases prevent the monopolist from imposing an implicit unit price exceeding marginal utility, in contrast with the single-bundle and multiple-purchase cases. If the firm were to impose such a price, the consumers would simply make joint purchases in order to diminish their consumption to the point where unit price equalled marginal utility. As a consequence, the monopolist must offer bundles along the respective demand curves, in contrast with the single-bundle and multiple-purchase cases. Note that this further implies socially suboptimal quantities to both types of consumers. The firm being a monopolist, it will set the average unit prices above marginal cost. Social inefficiency then follows from the fact that the bundles are on the demand curves, i.e., that average unit price equals marginal utility for both types of consumers.

As in the multiple-purchase case, the firm is worse off with joint purchases than in the single-bundle case, and the high-valuation consumer may be worse or better off. Furthermore, the intuition that joint purchases preclude quantity discounts is formally confirmed. Hence, put together, multiple and joint purchases make any price discrimination infeasible. Nevertheless, an important conclusion is that if arbitrage is not perfect in the sense that either multiple or joint purchases are not possible consumer actions, the monopolist can price discriminate and may even discriminate more than in the single-bundle case.

The related literature is small, and it deals with quite distinct problems. McManus (1998) provides an analysis of two-part tariffs, when a high-demand and a low-demand consumer can cooperate in order to pay the fixed fee only once. In contrast with my results, the firm is almost always better off when the consumers can cooperate. The reason is simple: when consumers of different types can cooperate (which is not the case in my model), the firm can set the fixed fee equal to the sum of the gross surpluses, whereas without consumer arbitrage, the fee cannot exceed the smallest of these surpluses. Further, Innes and Sexton (1993, 1994, 1996) study the issue of coalition formation among consumers, but in a very different context. The threat of the coalitions is to set up their own production units in order to compete with the monopolist. Consumers are supposed to be homogeneous, implying that nondiscrimination is optimal when there is no threat of collusion. A central result of Innes and Sexton's research is that price discrimination emerges as the optimal way to prevent collusion between consumers. Finally, Hammond (1987) shows in a general equilibrium model that goods that are exchangeable between consumers must be sold at linear prices: the exchangeability allows consumers to equalize their marginal rates of substitution. In contrast with that model, the present approach allows for a more detailed analysis by making a distinction between multiple and joint purchases.

The remainder of the article is organized as follows. Section 2 summarizes the single-bundle model and shows that it may not be robust to multiple and joint purchases. Section 3 characterizes the profit-maximizing price scheme when consumers have access to multiple purchases only, and compares it to the single-bundle solution. Section 4 first allows for joint purchases only, then for both multiple and joint purchases. Section 5 provides a conclusion and relates some empirical evidence to the results.

2. The single-bundle model: results and properties

■ Consider a monopolist who produces a good at constant marginal cost $c > 0$. The market for the good consists of a large number of consumers (normalized to a continuum with mass 1). Consumers have heterogeneous tastes. A consumer of type θ derives net surplus $U(\theta, q, t) = \theta V(q) - t$ from consuming a bundle with q units of the good at total price t. An alternative interpretation is that q stands for the quality of the good, sold in single units.[2] V is a twice continuously differentiable and strictly concave function defined on $[0, \infty)$, with $V(0) = 0$, $V' > 0$, $V'' < 0$. For each consumer, the parameter θ takes the value $\underline{\theta} > 0$ with probability α, and the value $\overline{\theta} > \underline{\theta}$ with probability $(1 - \alpha)$. There is thus a proportion α of low-valuation consumers and a proportion $1 - \alpha$ of high-valuation consumers. In equilibrium the monopolist offers two bundles, denoted $\{\underline{q}, \underline{t}\}$ and $\{\overline{q}, \overline{t}\}$, the former being meant for the low-valuation consumer and the latter for the high-valuation consumer. Further, let $\underline{p} \equiv \underline{t}/\underline{q}$ and $\overline{p} \equiv \overline{t}/\overline{q}$ denote the implicit unit prices.

The feature that renders the firm's problem nontrivial is the crucial, and plausible, assumption that an individual consumer's valuation θ is private information. Indeed,

[2] At the end of this section I give examples of goods that are relevant for the quantity and the quality interpretations in the single-bundle, multiple-purchase, and joint-purchase cases.

otherwise the monopolist would choose the quantity that maximizes the social surplus $\theta V(q) - cq$ and extract the whole surplus from the consumer by setting $t(q) = \theta V(q)$. Let $\overline{q}^*$ and $\underline{q}^*$ denote the first-best consumption levels, i.e., $\overline{\theta}V'(\overline{q}^*) = c$ and $\underline{\theta}V'(\underline{q}^*) = c$. Figure 1 shows the complete-information solution, with $\underline{Q}^*$ and $\overline{Q}^*$ denoting the respective bundles. In the figure, the indifference curves corresponding to zero utility for both types of consumers are drawn. Note that since utility is increasing in q and decreasing in t, points below an indifference curve strictly dominate those on the curve. Further, note that the slope of the line passing through the origin and a given bundle is the unit price implicit for that bundle.

Given that θ is private information, full rent extraction with socially optimal quantities is not an implementable scheme. For instance, a consumer of type $\overline{\theta}$ strictly prefers the bundle $\{q, t\} = \{\underline{q}^*, \underline{\theta}V(\underline{q}^*)\}$ to the bundle $\{\overline{q}^*, \overline{\theta}V(\overline{q}^*)\}$.

In the single-bundle model, the consumer has two options: choose one bundle or the other. Given this type of arbitrage, it is easy to show that it is optimal for the monopolist to deter the high-valuation consumer from choosing the low-valuation consumer's bundle (and vice versa). The bundles that maximize the firm's expected profit $\alpha(\underline{t} - c\underline{q}) + (1 - \alpha)(\overline{t} - c\overline{q})$ are such that the incentive constraint for the high-valuation consumer (1) and the participation constraint for the low-valuation consumer (2) are binding.

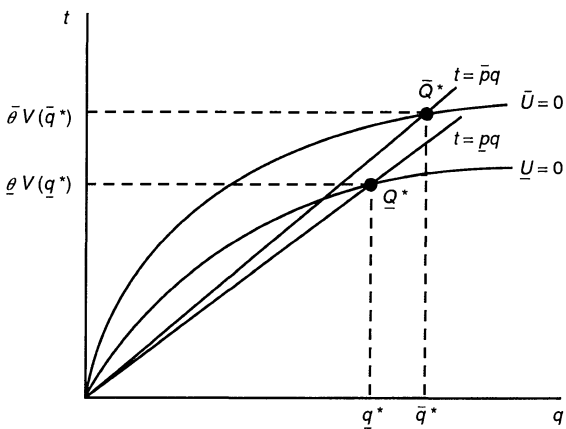
$$\overline{\theta}V(\overline{q}) - \overline{t} \geq \overline{\theta}V(\underline{q}) - \underline{t} \tag{1}$$

$$\underline{\theta}V(\underline{q}) - \underline{t} \geq 0. \tag{2}$$

The high-valuation consumer thus obtains a rent equal to $(\overline{\theta} - \underline{\theta})V(\underline{q})$. The firm makes a first-order gain by reducing this rent, at a second-order cost of reducing $\underline{q}$ compared to the first-best $\underline{q}^*$. More precisely, the quantity offered to the low-valuation consumer, $\underline{q}^s$, is defined by the following expression, where the superscript s stands for single-bundle:

FIGURE 1

THE SOLUTION UNDER COMPLETE INFORMATION



$$\underline{\theta}V'(\underline{q}^s) = c + \frac{1-\alpha}{\alpha}(\overline{\theta} - \underline{\theta})V'(\underline{q}^s).$$

To summarize, the following two bundles maximize the monopolist's expected profit in the two-type case in the single-bundle model:

$$\{\overline{q}^s, \overline{t}^s\} = \{\overline{q}^*, \overline{\theta}V(\overline{q}^*) - (\overline{\theta} - \underline{\theta})V(\underline{q}^s)\}, \qquad \{\underline{q}^s, \underline{t}^s\} = \{\underline{q}^s, \underline{\theta}V(\underline{q}^s)\}.$$

This characterization of the solution is incomplete, however. The monopolist may indeed prefer not to serve the low-valuation consumers at all; it would then be able to extract the whole surplus of the high-valuation consumer. To avoid that solution, one may assume either that $V'(0) = +\infty$ or that α is sufficiently large. I assume the latter throughout the model. Due to the implicit nature of the solution, however, I do not determine the relevant threshold but instead rely on numerical examples to ensure that I am not working on the empty set.

In Figure 2, points L and H depict a possible solution in the single-bundle model. The binding participation constraint for the low-valuation consumer implies that the indifference curve for this type of consumer through $\{\underline{q}^s, \underline{t}^s\}$ passes through the origin. Also, the indifference curve of the high-valuation consumer through $\{\overline{q}^s, \overline{t}^s\}$ passes through $\{\underline{q}^s, \underline{t}^s\}$, reflecting the fact that the incentive-compatibility constraint is binding.

Without further conditions on the parameters, there may be quantity premia or quantity discounts, i.e., $\overline{p}^s > \underline{p}^s$ or $\overline{p}^s < \underline{p}^s$.[3] For instance, for the function $V(q) = [1 - (1 - q)^2]/2$,

$$\overline{p}^s > \underline{p}^s \Leftrightarrow c > \underline{\theta}[\underline{\theta} - (1 - \alpha)\overline{\theta}]/[(2\alpha - 1)\overline{\theta} + (1 - \alpha)\underline{\theta}].$$

The solution depicted in Figure 2 is such that $\overline{p}^s < \underline{p}^s$. Now recall that points under an indifference curve are strictly preferred to those on the curve. In particular, the low-valuation consumer strictly prefers J to L, and the high-valuation consumer strictly prefers M to H. Now M is attainable by buying two small bundles, and J is accessible by buying one small bundle jointly with another consumer. Hence:

Observation 1. The single-bundle solution may not be robust to consumer actions such as multiple and joint purchases, where these are possible.

This is not very surprising per se although it indicates that modelling multiple and joint purchases is called for. Figure 2, however, allows for a more surprising conclusion:

Observation 2. A decreasing average unit price is not sufficient to prevent multiple purchases. Similarly, an increasing average unit price is not sufficient to prevent joint purchases.

Indeed, in Figure 2, the average unit price is decreasing ($\overline{p}^s < \underline{p}^s$). Nevertheless, the high-valuation consumer prefers to buy two small bundles (be at M in the figure) to buying one large bundle (be at H). Clearly, then, subadditivity is not sufficient to prevent multiple purchases of the small bundle. Further, note that the low-valuation consumer prefers to share a small bundle with another consumer (be at J) to buying a whole small bundle (be at L). Note also that this is true independently of $\overline{p}^s$, and hence even if there is a quantity premium, $\overline{p}^s > \underline{p}^s$.

[3] This is not particular to the discrete case: in the single-bundle model with a continuum of types, a decreasing average unit price is implied by a nondecreasing hazard rate for the distribution of types; although this is a common feature of many distributions, it is still a restriction.

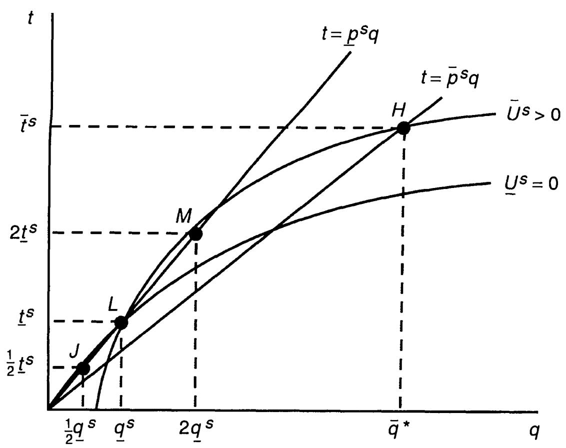
Observation 2 indicates that consumers do not arbitrage only to obtain the good at the lowest unit price. The key to understanding the above observation is that when a firm offers bundles as opposed to setting prices and letting the consumers choose the quantities, it seeks to offer bundles that are above the demand curve, i.e., such that the average unit price exceeds marginal utility. In Figure 2, at H the slope of the high-valuation consumer's indifference curve (i.e., the consumer's marginal utility) is smaller than the slope of the line passing through the origin and H (i.e., the implicit unit price $\bar{p}^s$). It follows that a high-valuation consumer would buy less than $\bar{q}^*$ given the implicit unit price. And he may actually be better off with several small bundles than with the large bundle, even if he pays a higher unit price (of course, he may ultimately prefer sharing a large bundle with other consumers, but that is not relevant here).

Let me briefly comment on the continuum-of-types case in relation to Observation 2. Indeed, under some circumstances a nonincreasing average price is sufficient to prevent multiple purchases with a continuum of types. In particular, that is true if the types are distributed on some interval $[0, \bar{\theta}]$. To see this, think of the single-bundle solution in that case (see Maskin and Riley, 1984, or Tirole, 1988): under some regularity assumptions, the menu of bundles $\{q(\theta), t(\theta)\}$ is a monotonic function t: $[0, \bar{q}^*] \rightarrow R$; further, for every type θ, the utility is maximized at $\{q(\theta), t(\theta)\}$, meaning that the indifference curve of a consumer of type θ is tangent to t at $\{q(\theta), t(\theta)\}$. In that case, a nonincreasing average unit price implies that t is concave. Clearly, this is robust to multiple purchases. However, a nonincreasing average price may not be sufficient to prevent multiple purchases if the types are distributed on some interval $[\underline{\theta}, \bar{\theta}]$, where $\underline{\theta} > 0$. Indeed, in that case a menu of bundles exhibiting quantity discounts is still a concave function, but it does not start at the origin. A consumer with a high θ, say $\tilde{\theta}$, may therefore strictly prefer buying several small bundles to buying the bundle meant for him: indeed, the line from the origin through the smallest bundle on the menu may lie strictly below consumer $\tilde{\theta}$'s indifference curve that is tangent to the menu.

Before analyzing the effects of multiple and joint purchases, it is useful to discuss the applicability of the various settings, in particular with respect to the two possible interpretations of q, namely quality and quantity. Let me start with the quantity interpretation. It is difficult to find goods for which neither multiple nor joint purchases are

FIGURE 2

THE SINGLE-BUNDLE SOLUTION



possible. One potential example is utilities, such as electricity, gas, and telephone, for which there are obvious technical difficulties with joint purchases, and multiple purchases does not make much sense. It is much easier to find goods for which multiple purchases are possible. For some of those, prohibitive transaction costs effectively rule out joint purchases. Relevant goods include most of those found in a supermarket, like chocolate bars, cereals, vegetables, soft drinks, etc. Goods for which it is relevant to consider both multiple and joint purchases include large consumer goods, for which transaction costs are relatively low. An example is clothes. Obviously if jeans are sold in bundles, it may be of interest to buy a large bundle with a friend if there is a quantity discount. Conversely, a consumer may want to buy two sweaters instead of a bundle of four sweaters. Finally, I can think of no good for which only joint purchases are possible for the quantity interpretation.

When it comes to the quality interpretation, it is easier to find applications for the single-bundle model: any good for which a consumer typically has a "true" unit demand (i.e., either he buys one unit or no unit at all) is suitable. Cars for everyday use are an example of such a good. Other examples could be computers for everyday extensive use, transportation between city x and y on a certain date by airplane or train, and seats in a theater. In contrast, it is difficult to find relevant examples for multiple purchases as modelled here. Indeed, it requires finding goods for which there is a linear tradeoff between quantity and quality (i.e., consuming two units with quality x is equivalent to consuming one unit with quality $2x$). However, such tradeoffs probably do not exist to a significant extent. Indeed, taking wine as an example, a consumer who has developed a taste for quality, and who can afford quality, probably does not even consider that several bottles of low-quality wine could replace one bottle of high-quality wine. Finally, examples of goods worth purchasing jointly, but not multiply, are of two kinds. First, for some goods consumers do not have a true unit demand, because the good is not consumed at all times. Examples include lawn mowers, snow blowers, and swimming pools. Neighbors may quite easily buy these goods jointly.[4] The joint purchase implies a saving but does not necessarily imply that a different quality level is purchased than if no joint purchase was made. For the second category, interpret θ as the income level (see Tirole, 1988). For people with a low θ, making a joint purchase could be the only way to consume the high-quality good. Taking wine (or any high-quality food) as an example, a consumer might wish to but could not afford to spend a thousand dollars alone on an exclusive wine, but he would accept buying the bottle jointly with some friends.

To summarize, for the quantity interpretation of q, both multiple purchases and joint purchases are highly relevant, although for some goods joint purchases are ruled out due to transactions costs. For the quality interpretation, joint purchases seem more relevant than multiple purchases. Further, the single-bundle model appears to apply more easily to the quality than to the quantity interpretation. Keeping these observations in mind, I nonetheless refer only to the quantity interpretation when developing the model below, in order to keep the exposition simple.

3. Multiple purchases

■ In this section, consumers may make multiple but not joint purchases. If the firm could identify a consumer, it could prevent multiple purchases simply by forbidding them. I therefore assume that purchases are made anonymously, implying that the firm

[4] Note that there should be significant moral hazard problems connected with such joint purchases. All transactions costs will, however, be disregarded in the model.

cannot impose any restriction on the number of bundles bought by any consumer. I also assume that a bundle sold to a consumer cannot depend on the bundles sold to other consumers. These are realistic assumptions for consumer goods. It would be ideal to assume that a consumer can only buy discrete numbers of bundles. However, this turns out to be technically problematic. Therefore, I assume from the start that the consumer can buy any real number (exceeding one) of bundles when making multiple purchases, although I will comment on the integer case at the end of this section. In Figure 1, all price-quantity combinations on the price lines above $\underline{Q}^*$ and $\overline{Q}^*$, respectively, are attainable through multiple purchases.

□ **The profit-maximizing price schedule.** To determine the profit-maximizing price schedule, I need to know how the consumers behave in equilibrium: how many bundles will each type of consumer buy? Fortunately, the revelation principle applies here: given any two bundles $\{\underline{t}, \underline{q}\}$ and $\{\overline{t}, \overline{q}\}$, and whatever the choice of any type of consumer facing these bundles, the monopolist can redesign the bundles so as to replicate this choice. Therefore, attention can be restricted to bundles such as in equilibrium the high-valuation consumer buys one bundle $\{\overline{t}, \overline{q}\}$ and the low-valuation consumer buys one bundle $\{\underline{t}, \underline{q}\}$. The expected profit of the monopolist is thus[5]

$$\alpha(\underline{t} - c\underline{q}) + (1 - \alpha)(\overline{t} - c\overline{q}). \tag{3}$$

To ensure the above-described equilibrium behavior, the bundles must satisfy individual-rationality and incentive-compatibility constraints. For a consumer of type $\underline{\theta}$ to prefer one bundle $\{\underline{q}, \underline{t}\}$ to any number of bundles $\{\underline{q}, \underline{t}\}$ or $\{\overline{q}, \overline{t}\}$, the following two sets of incentive-compatibility constraints must be respected (in the constraints, k and k' are real numbers):

$$\underline{\theta}V(\underline{q}) - \underline{t} \geq \underline{\theta}V(k\overline{q}) - k\overline{t} \qquad \forall\, k \geq 1 \tag{4}$$

$$\underline{\theta}V(\underline{q}) - \underline{t} \geq \underline{\theta}V(k\underline{q}) - k\underline{t} \qquad \forall\, k > 1 \tag{5}$$

$$\underline{\theta}V(\underline{q}) - \underline{t} \geq \underline{\theta}V(k\underline{q} + k'\overline{q}) - k\underline{t} - k'\overline{t} \quad \forall\, k, k' \geq 1. \tag{6}$$

The first set of constraints ensures that a consumer of type $\underline{\theta}$ prefers *one* bundle $\{\underline{q}, \underline{t}\}$ to *any number* $k \geq 1$ of bundles $\{\overline{q}, \overline{t}\}$; the second set ensures that if the same consumer could choose between any other number $k \geq 1$ of bundles $\{\underline{q}, \underline{t}\}$, he chooses exactly one such bundle; and the third set ensures that the same consumer would not be better off with any combination of the two bundles. Then, for a consumer of type $\overline{\theta}$ to prefer one bundle $\{\overline{q}, \overline{t}\}$ to any other number of bundles $\{\overline{q}, \overline{t}\}$ or $\{\underline{q}, \underline{t}\}$, the incentive-compatibility constraints to satisfy are[6]

$$\overline{\theta}V(\overline{q}) - \overline{t} \geq \overline{\theta}V(k\underline{q}) - k\underline{t} \qquad \forall\, k \geq 1 \tag{7}$$

$$\overline{\theta}V(\overline{q}) - \overline{t} \geq \overline{\theta}V(k\overline{q}) - k\overline{t} \qquad \forall\, k > 1 \tag{8}$$

$$\overline{\theta}V(\overline{q}) - \overline{t} \geq \overline{\theta}V(k\underline{q} + k'\overline{q}) - k\underline{t} - k'\overline{t} \quad \forall\, k, k' \geq 1. \tag{9}$$

It might seem odd to think that the monopolist would prevent a consumer from

[5] As in the single-bundle case, it is sufficient to assume that α is sufficiently large for the firm to serve the low-valuation consumers, i.e., for $\underline{q} > 0$.

[6] Note that there is no solution if V is linear. Indeed, should a consumer enjoy a positive rent from

consuming several bundles if he makes a nonnegative profit on the sale of one bundle. Writing the constraints in this manner, however, is only an implication of the revelation principle. If the monopolist proposes a bundle $\{\bar{q}, \bar{t}\}$ such that a consumer of type $\bar{\theta}$ actually buys two such bundles, he might as well propose the bundle $\{q', t'\} = \{2\bar{q}, 2\bar{t}\}$ to $\bar{\theta}$. The incentive-compatibility constraints guarantee that this would indeed be the case. The participation constraints complete the list of constraints defining the multiple-purchase problem:

$$\bar{\theta}V(\bar{q}) - \bar{t} \geq 0 \tag{10}$$

$$\underline{\theta}V(\underline{q}) - \underline{t} \geq 0. \tag{11}$$

Although the maximization problem seems complex, many constraints can be easily eliminated. As argued in the previous section, the monopolist seeks to offer bundles that are above the demand curves. A declining marginal utility therefore implies that a consumer does not wish to buy multiples of the bundle meant for him (thus eliminating constraints (5) and (8)). Further, it is quite obvious that constraints (6) and (9) are slack: buying only bundles with the lowest average unit price is better than any combination of the two bundles. As described in the proposition below, the two binding constraints are the participation constraint for the low-valuation consumer (11) and one of the incentive-compatibility constraints (7) for the high-valuation consumer, namely the one corresponding to the high-valuation consumer's preferred number of small bundles. In the proposition the superscript m stands for multiple purchases.

Proposition 1. When consumers can make multiple purchases of bundles, the profit-maximizing price-quantity schedule is such that the low-valuation consumer gets no rent, whereas the high-valuation consumer gets a strictly positive rent. The transfers $\underline{t}^m$ and $\bar{t}^m$ are uniquely determined by the binding constraints (11) and (7) for $k^*(\underline{q})$, where $k^*(\underline{q}) \geq 1$ is the preferred number of bundles $\{\underline{q}, \underline{t}^m\}$ of the consumer of type $\bar{\theta}$:

$$\underline{t}^m(\underline{q}, \bar{q}) = \underline{\theta}V(\underline{q}) \qquad \bar{t}^m(\underline{q}, \bar{q}) = \bar{\theta}V(\bar{q}) - \bar{\theta}V(k^*(\underline{q})\underline{q}) + k^*(\underline{q})\underline{\theta}V(\underline{q}).$$

Proof. See the Appendix.

The result corresponds to the intuition developed in the previous section: only the high-valuation consumer is attracted by multiple purchases. Thus the rent pattern is the same as in the single-bundle model, although the rent to the high-valuation consumer now is determined not by the utility he would get by buying one small bundle, but by his preferred number $k^*(\underline{q}) \geq 1$ of such bundles. The high-valuation consumer's rent is therefore

$$\bar{\theta}V(k^*(\underline{q})\underline{q}) - k^*(\underline{q})\underline{\theta}V(\underline{q}) > 0.$$

This rent is increasing in $\underline{q}$. Indeed, since the high-valuation consumer is able to purchase any quantity larger than one of the small bundle, the lower the average unit price in the small bundle, the higher his utility. And the implicit unit price in the small bundle is a decreasing function of $\underline{q}$, given that the low-valuation consumer's utility is zero and that V is strictly concave. I therefore obtain the following:

consuming one bundle, which must be the case for a consumer of type $\bar{\theta}$ not to choose $\{\underline{q}, \underline{t}\}$, he would buy an infinite number of bundles, if the rent were linear in k.

Proposition 2. When the consumers can make multiple purchases of bundles, the profit-maximizing price-quantity schedule is such that

(i) the quantity offered to the high-valuation consumer is the efficient one: $\overline{q}^m = \overline{q}^*$;

(ii) the quantity offered to the low-valuation consumer is always distorted downward compared with the first best. The following expression defines $\underline{q}^m$:

$$\underline{\theta} V'(\underline{q}^m) = c + \frac{1-\alpha}{\alpha} k^*(\underline{q}^m)[\overline{\theta} V'(k^*(\underline{q}^m)\underline{q}^m) - \underline{\theta} V'(\underline{q}^m)].$$

Proof. See the Appendix.

Figure 3 depicts a possible solution in the multiple-purchase case: the high-valuation consumer is indifferent between the large bundle and his preferred number of small bundles. Note that $\overline{q}^m > k^*(\underline{q}^m)\underline{q}^m$; this is a general property of the solution (see the proof of Proposition 2).

Although the qualitative features are the same as in the single-bundle model,[7] further analysis reveals some interesting differences between the multiple-purchase case and the single-bundle model.

□ **Comparing the multiple-purchase and the single-bundle solutions.** To start with, a striking general qualitative difference exists between the two solutions:

Proposition 3. If the high-valuation consumer can buy any real number $k \in [1, \infty)$ of bundles $\{\underline{q}, \underline{t}\}$, the profit-maximizing bundles $\{\underline{t}^m, \underline{q}^m\}$ and $\{\overline{t}^m, \overline{q}^m\}$ are such that the implicit unit price is strictly decreasing: $\underline{p}^m > \overline{p}^m$.

Proof. When $k^*(\underline{q}^m)$ is an interior solution, the result follows from the tangency of the indifference curve of the high-valuation consumer to the line defined by the equation $t = \underline{p}^m q$, together with the fact that $\overline{q}^m > k^*(\underline{q}^m)\underline{q}^m$ (see the proof of Proposition 2).

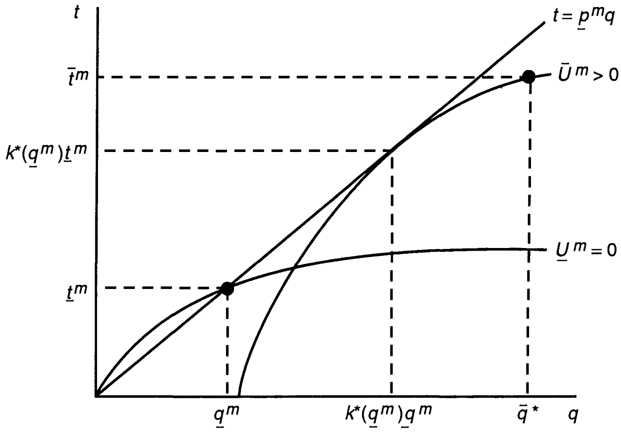
If $k^*(\underline{q}^m)$ is not an interior solution, then $k^*(\underline{q}^m) = 1$, and $\overline{p} < \underline{p}$ is implied by $\overline{q}^m > \underline{q}^m$ together with the fact that the line defined by the equation $t = \underline{p}^m q$ to the right of $\underline{q}^m$ must lie above the indifference curve of the high-valuation consumer for $k^*(\underline{q}) = 1$. *Q.E.D.*

When consumers can choose their preferred real number of bundles, the quantity in the large bundle can be exactly replicated by a certain number of small bundles, in which case the firm would not sell any large bundles if there were a quantity premium $\underline{p}^m < \overline{p}^m$. Thus, a nonincreasing average unit price is necessary. But the above proposition further indicates that the solution always exhibits strict quantity discounts. This is quite surprising. Indeed, consider parameter values for which the firm imposes quantity premia in the single-bundle model ($\underline{p}^s < \overline{p}^s$); intuition would suggest that for the same parameter values, the firm would simply sell the good at a linear price $\underline{p}^m = \overline{p}^m$ when the threat of multiple purchases makes premia infeasible. But as Proposition 3 shows, this is not true. The reason behind this result is that if the firm sells the good at a linear unit price, then the quantity offered to the high-valuation consumer cannot exceed the quantity demanded by that consumer at that price for the multiple-purchase constraints to be satisfied. Therefore, marginal utility exceeds marginal cost, implying that total surplus can be increased by increasing $\overline{q}$: by also decreasing $\overline{p}$, the consumer's utility can be kept constant, implying that the firm's profit increases.

[7] Note, moreover, that the single-bundle model appears as a special case; just set $k^*(\underline{q}) \equiv 1$.

FIGURE 3

THE SOLUTION WITH MULTIPLE PURCHASES ONLY



The firm's expected profit is nevertheless inferior compared to the single-bundle model. Let Π^s and Π^m denote the firm's expected profit in the single-bundle and multiple-purchase models, respectively.

Proposition 4. The firm's profit is larger if consumers can only pick a single bundle than if they can make multiple purchases of bundles: $\Pi^s \geq \Pi^m$.

Proof. The result follows from a revealed preference argument: at the multiple-purchase solution described in Propositions 1 and 2, all the constraints of the single-bundle model are satisfied. Hence, the multiple-purchase solution can be chosen in the single-bundle model; if it is not chosen, it must yield a lower expected profit than the single-bundle solution. *Q.E.D.*

This is very intuitive: given that the consumers have access to a more efficient arbitrage technology, the firm is worse off. Letting $\overline{U}^s$ and $\overline{U}^m$ denote the high-valuation consumer's utility in the single-bundle and multiple-purchase models, respectively, the following is implied by Proposition 4.

Corollary 1. If $\underline{q}^m > \underline{q}^s$, social surplus is higher when consumers can make multiple purchases as opposed to when they can only pick a single bundle, implying that $\overline{U}^m > \overline{U}^s$.

However, it is not generally true, nor false, that $\underline{q}^m > \underline{q}^s$, as shown by the examples in Table 1. The examples also indicate that it may be the case that $\overline{U}^m < \overline{U}^s$ if $\underline{q}^m < \underline{q}^s$.[8] The possibility to make multiple purchases, as opposed to being able to arbitrage only between two given bundles, therefore does not necessarily increase the high-valuation consumer's utility.

Finally, let me compare the amount of discrimination itself, naturally defined as the difference in average unit price between the two bundles, $\underline{p} - \overline{p}$. Recall from Observation 2 that quantity discounts are not sufficient to prevent multiple purchases. This intuitively suggests that if there is a quantity discount at the single-bundle solution ($\underline{p}^s - \overline{p}^s > 0$), and the single-bundle solution is not robust to multiple purchases, then

[8] The examples show only the values that are relevant for a particular result or observation. For each example, the values of all variables have been calculated; all examples are such that $\underline{q} > 0$, $\overline{q} > 0$, and $k^*(\underline{q}) > 1$. All numerical examples are available upon request from the author.

TABLE 1 **Examples of the Single-Bundle and Multiple-Purchase Solutions: $\underline{q}$ and $\overline{U}$**

$V(q)$	$\underline{\theta}$	$\overline{\theta}$	c	α	a	$\underline{q}^s$	$\underline{q}^m$	$\overline{U}^s$	$\overline{U}^m$
$\frac{q(2-q)}{2}$	8.0	14	5.0	.75	—	.17	.29	.92	1.83
$\frac{q^{1-a}}{1-a}$	40.0	116	10.0	.8	.4	6.4	2.1	385.5	274.6

the discrimination should be even more pronounced for the multiple-purchase solution: $\underline{p}^m - \overline{p}^m > \underline{p}^s - \overline{p}^s > 0$. This may be the case, as in the first example in Table 2. However, it appears not to be generally true, as the second example shows.

□ **The integer case.** I have partially analyzed the case when consumers can buy only discrete numbers of bundles, and the following summarizes the findings.[9] Some of the above results apply. In particular, Proposition 1 is valid as it stands, the only difference being that the preferred number of bundles $k^*(\underline{q})$ must be an integer. Further, the single-bundle solution may be robust to multiple purchases. Related to this, quantity premia are feasible under some circumstances. To see this, assume that there is a quantity premium and that $\overline{\theta} - \underline{\theta}$ is small. Then, making a multiple purchase of the small bundle could imply a much larger consumption than with one large bundle. If $\overline{p} - \underline{p}$ is small enough, the high-valuation consumer prefers the large bundle to two or more small bundles. Turning to the quantities, there is no distortion at the top ($\overline{q} = \overline{q}^*$); the determination of $\underline{q}$ is problematic, since it involves differentiating $k^*(\underline{q})$. Interestingly, intuition suggests that $\underline{q}$ may be distorted upward. To see this, set the quantity in the small bundle to $\underline{q}^*$ and assume that the high-valuation consumer then prefers two small bundles to one large, but that a small increase of $\underline{q}$ implies that he prefers the large bundle to two small ones. This remains to be formally confirmed, however.

4. Multiple and joint purchases

■ I start this section with an analysis of the effects of joint purchases alone, assuming that multiple purchases are not possible. Then I proceed to the case in which both joint and multiple purchases are possible consumer actions.

□ **Joint purchases only.** A joint purchase occurs when a group of consumers buys one or more bundles to divide among themselves. In reality, consumers' access to joint purchases depends on a number of factors, e.g., the information consumers have about each other, the number of consumers, transactions costs, etc. For simplicity I disregard these factors. First, I assume that every consumer knows that the other consumers exist, as well as their types. Second, I disregard transactions costs.[10] Third, I impose one restriction on the joint-purchase possibilities: I assume that only consumers of the same type make joint purchases. This simplifies the analysis, since I need not define how consumers of different types share bundles. Moreover, the assumption is realistic in many circumstances; neighbors buying a good jointly are more likely than not to have approximately the same income level (as noted above, the different types can indeed be interpreted as representing different income levels). Together with the assumption that there is a continuum of consumers, these assumptions imply that a consumer may

[9] The details of these are available upon request from the author.

[10] The Internet may prove to make these assumptions realistic in the near future.

TABLE 2 Comparing Amount of Price Discrimination in Single-Bundle and Multiple-Purchase Solutions: Two Examples

$V(q)$	$\underline{\theta}$	$\overline{\theta}$	c	α	$\underline{p}^s - \overline{p}^s$	$\underline{p}^m - \overline{p}^m$
$\frac{q(2-q)}{2}$	4.0	6.9	1.0	.50	.16	.36
$\frac{q(2-q)}{2}$	4.0	6.0	1.0	.42	.53	.46

effectively consume any fraction of any bundle by making a joint purchase with the other consumers of the same type. In Figure 1, with joint purchases a consumer has access to any price-quantity combination on the lines between the origin and $\underline{Q}^*$, and between the origin and $\overline{Q}^*$, respectively.

As in the multiple-purchase case, it is easy to verify that the revelation principle holds, implying that standard techniques may be used to characterize the profit-maximizing price schedule.[11] The monopolist can therefore restrict attention to bundles $\{\underline{t}, \underline{q}\}$ and $\{\overline{t}, \overline{q}\}$ such that in equilibrium a low-valuation consumer buys the former bundle and the high-valuation consumer buys the latter bundle. For this to be the equilibrium behavior, each type of consumer must prefer buying the bundle meant for his type to any fraction of any bundle. Thus, the following incentive-compatibility constraints must be satisfied.

$$\underline{\theta}V(\underline{q}) - \underline{t} \geq \underline{\theta}V\left(\frac{1}{n}\underline{q}\right) - \frac{1}{n}\underline{t} \qquad \forall\, n > 1 \tag{12}$$

$$\underline{\theta}V(\underline{q}) - \underline{t} \geq \underline{\theta}V\left(\frac{1}{n}\overline{q}\right) - \frac{1}{n}\overline{t} \qquad \forall\, n \geq 1 \tag{13}$$

$$\underline{\theta}V(\underline{q}) - \underline{t} \geq \underline{\theta}V\left(\frac{1}{n}\overline{q} + \frac{1}{n'}\underline{q}\right) - \frac{1}{n}\overline{t} - \frac{1}{n'}\underline{t} \qquad \forall\, n, n' \geq 1 \tag{14}$$

$$\overline{\theta}V(\overline{q}) - \overline{t} \geq \overline{\theta}V\left(\frac{1}{n}\underline{q}\right) - \frac{1}{n}\underline{t} \qquad \forall\, n \geq 1 \tag{15}$$

$$\overline{\theta}V(\overline{q}) - \overline{t} \geq \overline{\theta}V\left(\frac{1}{n}\overline{q}\right) - \frac{1}{n}\overline{t} \qquad \forall\, n > 1. \tag{16}$$

$$\overline{\theta}V(\overline{q}) - \overline{t} \geq \overline{\theta}V\left(\frac{1}{n}\overline{q} + \frac{1}{n'}\underline{q}\right) - \frac{1}{n}\overline{t} - \frac{1}{n'}\underline{t} \qquad \forall\, n, n' \geq 1. \tag{17}$$

Note that multiple purchases are ruled out since $n, n' \geq 1$. Together with the participation constraints (10) and (11), the incentive constraints define the set of feasible bundles. The firm's objective is to find the pair of bundles in that set which maximizes the expected profit:

$$\alpha(\underline{t} - c\underline{q}) + (1 - \alpha)(\overline{t} - c\overline{q}).$$

The following result, which is qualitatively different from the single-bundle solution, is immediate:

[11] This would not generally be true with a finite number of consumers. See the discussion about the finite case at the end of this section.

Proposition 5. When consumers can make joint purchases, the monopolist leaves a strictly positive utility to both types of consumers.

Proof. Take the constraint (12) for some $n > 1$, multiply it by n, and rewrite it to obtain

$$(n - 1)[\underline{\theta}V(\underline{q}) - \underline{t}] \geq n\underline{\theta}V\left(\frac{1}{n}\underline{q}\right) - \underline{\theta}V(\underline{q}).$$

By strict concavity of V, the right-hand side is strictly positive given that $\underline{q} > 0$. Thus the utility of the low-valuation consumer, $\underline{\theta}V(\underline{q}) - \underline{t}$, is strictly positive. Then (15) for $n = 1$ implies that the high-valuation consumer also obtains a strictly positive utility. *Q.E.D.*

This result is in stark contrast with both the single-bundle and the multiple-purchase results. The reason is simple: as noted above, at the single-bundle solution consumers pay an implicit unit price exceeding their marginal utility. Therefore, they would consume less if they could. Access to joint purchases enables the low-valuation consumer to consume less, implying that he must obtain a strictly positive rent.

Now, intuition suggests that joint purchases make quantity discounts infeasible. This intuition is confirmed in the following proposition, where $\underline{\eta}$ and $\overline{\eta}$ are the elasticities of demand for low-valuation and high-valuation consumers, respectively, and the superscript j stands for joint purchases.

Proposition 6. With a continuum of consumers, access to joint purchases implies that the average unit price is nondecreasing: $\underline{p}^j \leq \overline{p}^j$. Furthermore:

(i) both types of consumers pay an average unit price equal to their marginal utility: $\underline{\theta}V'(\underline{q}^j) = \underline{p}^j$ and $\overline{\theta}V'(\overline{q}^j) = \overline{p}^j$;

(ii) both quantities are distorted downward compared to the first-best: $\underline{q}^j < \underline{q}^*$ and $\overline{q}^j < \overline{q}^*$;

(iii) the constraint (15) for $n = 1$ may or may not be binding, depending on the parameter values. When it is not binding, the monopolist offers the menu of bundles $\{\underline{q}^j, \underline{t}^j\} = \{\underline{q}^j, \underline{p}^j\,\underline{q}^j\}$ and $\{\overline{q}^j, \overline{t}^j\} = \{\overline{q}^j, \overline{p}^j\,\overline{q}^j\}$, where $\underline{p}^j$ and $\overline{p}^j$ are defined by the standard formulas

$$\frac{\underline{p}^j - c}{\underline{p}^j} = -\frac{1}{\underline{\eta}} \qquad \frac{\overline{p}^j - c}{\overline{p}^j} = -\frac{1}{\overline{\eta}},$$

and $\underline{q}^j$ and $\overline{q}^j$ are given by $\underline{\theta}V'(\underline{q}^j) = \underline{p}^j$ and $\overline{\theta}V'(\overline{q}^j) = \overline{p}^j$.

Proof. See the Appendix.

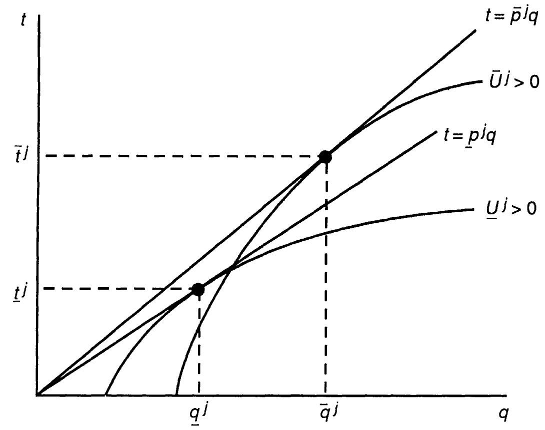
Figure 4 shows the joint-purchase solution when constraint (15) for $n = 1$ is not binding.

Proposition 6 contains several insights worth commenting on. First, the intuition that joint purchases preclude quantity discounts is formally confirmed. It recalls the multiple-purchase case, although the reverse is happening: if the small bundle can be replicated by a fraction of the large bundle, then quantity discounts are not feasible. However, strict quantity premia are not necessary.[12]

[12] As an example, constant elasticity demand implies $\underline{p} = \overline{p}$, whereas linear demand implies $\underline{p} < \overline{p}$.

FIGURE 4

THE SOLUTION WITH JOINT PURCHASES ONLY



Interestingly, and in contrast with the single-bundle and multiple-purchase cases, joint purchases prevent the monopolist from offering bundles that are not on the demand curve. This is easily understood, however: given that any fraction of any bundle is attainable through joint purchases, if the implicit unit price exceeds marginal utility, the consumers will simply make joint purchases to reduce their consumption to the point where unit price equals marginal utility. Joint purchases thus take us back to the usual monopoly-pricing model, the specificity here being that there are two demand curves.[13] Therefore I find the regular monopoly-pricing solution if the incentive constraint, i.e., constraint (15) $n = 1$, is not violated at that solution: the unit price charged to a consumer is defined by the Lerner index, and the consumer gets the quantity at which his marginal utility equals the unit price.

Finally, as in the usual monopoly-pricing model, the production levels are suboptimal (both quantities are distorted downward compared to the first best). It should be noted, however, that the distortion at the top, $\overline{q} < \overline{q}^*$, is a very unusual feature for this type of hidden-information model. The distortion is explained by the usual rent-efficiency tradeoff: the rent obtained by the high-valuation consumer is increasing in $\overline{q}$; distorting $\overline{q}$ slightly induces a second-order loss for the monopolist but allows a first-order gain, namely to increase the extracted surplus. The fact that the rent is increasing in $\overline{q}$ follows from the tangency of the indifference curve with the price line, together with the strict concavity of V.

As in the multiple-purchase model, one can use a simple revealed preference argument to show that the firm's profit is higher given the single-bundle solution than with the joint-purchase solution. However, it is not clear whether total welfare is higher or lower than in the single-bundle case. There is a welfare loss in that the large quantity $\overline{q}$ is socially suboptimal when joint purchases are possible, but the low-valuation consumer's quantity may be less distorted than in the single-bundle case. It is not excluded that this lesser distortion may offset the welfare loss due to the distortion of the large quantity, although I have not found a numerical example corroborating this. In the

[13] Note that if the high-valuation consumer strictly prefers the large to the small bundle (i.e., if the "classical" incentive constraint is not binding), the solution is actually the third-degree price discrimination solution, the difference being that there is no external distinction between the different types of consumers here.

numerical examples in Table 3, total expected welfare (*TEW*) is therefore lower in the joint-purchase than in the single-bundle case. Interestingly, the high-valuation consumer's utility $\overline{U}$ may be higher or lower. Also, the amount of price discrimination here is higher with than without joint purchases.

That the firm is worse off with than without joint purchases contrasts with McManus's (1998) analysis: he finds that the firm may be better off when consumers can make joint purchases by agreeing to pay the fixed fee only once in a two-part tariff setting. The reason is that McManus allows for joint purchases between consumers of different types, whereas I do not; in my model, the bundles therefore need to satisfy participation constraints for each type of consumer, whereas in the framework of McManus, there is only a participation constraint for the coalition of consumers. In McManus's model the firm can therefore extract a larger surplus.

□ **The finite case.** The following discussion is intended to provide some insights into the modelling of joint purchases with a finite number of consumers. A modeller must consider the following three issues, each of them with a number of alternate routes.

First, one may or may not allow for joint purchases made by consumers of different types. Although allowing for them could add realism to the model, I believe this would make it more complex without yielding essential insights. Second, one may or may not take into account transactions costs. Third, and most important, the monopolist's information about realized demand structure may or may not be complete, where complete information means that the monopolist knows the total number of consumers of each type. Since the monopolist does not serve all consumers at once, there is still incomplete information about the type of each consumer. If information is incomplete, the model must settle on one particular form of uncertainty: either the monopolist knows the total number of consumers and the expected distribution of types but not the realized distribution,[14] or there is uncertainty about the total number of consumers as well as about the realized distribution of types.

It is tempting to choose one of the incomplete-information settings, for the sake of realism. The analysis of these, however, appears quite complex. I indeed conjecture that the revelation principle does not always apply in an incomplete-information setting. In other words, it may be optimal for the firm to allow for joint purchases in equilibrium. The intuition for this is the following: there exist states in which joint purchases are not relevant[15] or not very costly to prevent.[16] If these states are sufficiently likely, it is probably not profitable to deter joint purchases in all possible states. In contrast, it is easy to verify that it is optimal for the monopolist to deter any joint purchases in the complete-information case.

Nevertheless, the complete-information case is not simple either. I have analyzed the case in which the monopolist knows there are two consumers of each type, assuming that only consumers of the same type make joint purchases. In that setting, any consumer has four alternatives: one small bundle, one large bundle, half a small bundle, and half a large bundle. The main insights from the analysis are the following. First, the low-valuation consumer obtains a strictly positive rent (the proof of Proposition 5

[14] In this case, one may actually further assume either that all consumers are of the same type, i.e., there is *ex post* homogeneity of demand, or that there is *ex post* heterogeneity of demand.

[15] With two consumers only, and assuming that only consumers of the same type make joint purchases, if the realized state is one consumer of each type, the consumers cannot make any joint purchases.

[16] Assume uncertainty about the total number of consumers and *ex post* demand homogeneity. Then the number of possible joint purchases increases with the number of consumers. It should therefore be more costly for the monopolist to deter joint purchases the larger the consumer population.

TABLE 3 **Comparing Single-Bundle and Joint-Purchase Solutions: Two Examples**

$V(q)$	$\underline{\theta}$	$\overline{\theta}$	c	α	TEW^s	TEW^j	$\overline{U}^s$	$\overline{U}^j$	$\overline{p}^s - \underline{p}^s$	$\overline{p}^j - \underline{p}^j$
$\frac{q(2-q)}{2}$	5.0	15.0	4.0	.99	.14	.10	1.67	1.0	2.7	5.0
$\frac{q(2-q)}{2}$	3.0	9.0	2.5	.94	.16	.14	.26	.59	2.45	3.0

applies). Actually, this result holds as soon as there are two low-valuation consumers. Second, $\overline{q}$ is distorted downward compared to the efficient $\overline{q}^*$ if any coalition is attracted by the bundle $\{\overline{q}, \overline{t}\}$ (i.e., if at the optimum any consumer is indifferent between the bundle meant for him and buying half of the large bundle). The explanation is the same as in the infinite case. Third, there may be quantity premia or discounts. Indeed, as in the multiple-purchase integer case, here the consumers have access to only a limited number of points, and it may therefore be the case that a quantity discount is feasible with joint purchases. For instance, if there is a large difference between the quantities in the two bundles, the low-valuation consumers may not be interested in buying half a large bundle even if they end up paying a lower unit price. Fourth, and related to the third point, the profit-maximizing bundles are off the demand curves. When consumers cannot purchase any fraction of any bundle, the monopolist can charge an implicit unit price exceeding marginal utility.

□ **Multiple and joint purchases.** Coming back to the continuum-of-consumers assumption, it comes as no surprise that taken together, multiple and joint purchases imply that the firm cannot price discriminate. This follows from Propositions 3 and 6.

Corollary 2. When there is an infinite number of consumers and all the consumers can buy their preferred real number of bundles, either through multiple or joint purchases, the monopolist cannot price discriminate and offers the good according to a linear price schedule. The price is defined by the Lerner index:

$$\frac{p^\ell - c}{p^\ell} = -\frac{1}{\eta},$$

where $\eta \equiv (d\tilde{q}/dp)(p/\tilde{q})$ is the elasticity of demand, $\tilde{q}$ denoting the (average) demand $\alpha\underline{q} + (1 - \alpha)\overline{q}$ per consumer.

The model thus provides a formalization of the old conjecture that perfect arbitrage induces linear pricing (see Tirole, 1988). Indeed, as modelled here, multiple and joint purchases taken together implies perfect arbitrage possibilities for the consumers.

5. Conclusion

■ In this article I have characterized the menu of bundles offered by a monopolist to a population with two types of consumers, given that consumers can purchase several bundles and/or share bundles with others. I find that although the absence of perfect arbitrage is a necessary condition for the firm to be able to price discriminate, partial arbitrage in the form of either multiple or joint purchases does not preclude price discrimination. In fact, partial arbitrage can even lead to a more pronounced discrimination than if the consumers could pick only one single bundle, as was assumed in the previous literature. Interestingly, a more pronounced discrimination does not necessarily imply a lower utility for consumers. Nevertheless, the firm is worse off when consumers can make multiple or joint purchases than when they can pick only one single bundle.

Furthermore, and in contrast with earlier literature, the model yields clear predictions for when one should observe quantity and quality discounts or premia. As the following discussion suggests, casual observations and some empirical studies are consistent with the predictions. I first focus on the quantity interpretation and then turn to the quality interpretation of the model.

As argued in Section 2, multiple purchases are always relevant for the quantity interpretation, except when there are physical restrictions. For many consumer goods, joint purchases are ruled out by transactions costs that are high relative to the value of the good. For such goods, the model predicts that quantity discounts should be observed, and that is clearly the case. Further, for goods of higher value, joint purchases are also relevant. Then neither quantity discounts nor quantity premia should be feasible, a prediction that is confirmed for at least some goods: clothes, for instance, are only rarely sold in bundles of more than one unit (those sold in bundles are typically small items like socks). Finally, Wilson (1992) gives several examples of the complex tariffication applied to different utilities, which are goods fitting into the single-bundle framework. They indicate that quantity discounts are observed. It should be noticed, however, that firms mostly offer two-part tariffs instead of menus of bundles for these goods.

When it comes to the quality interpretation, it seems justifiable to believe that a larger quantity of low-quality goods is rarely seen as the equivalent of a high-quality good, making multiple purchases irrelevant. In contrast, joint purchases seem worth contemplating for many goods, such as large consumer durables or high-quality food or drinks. For these goods, the model predicts quality premia. Further, there are many goods for which the single-bundle model applies. These goods may be sold with either quality discounts or quality premia, according to the model. The only empirical studies I am aware of indicate quality premia: Kwoka (1992) analyzes car prices in the United States, and Sällström (1991) investigates jam prices in Sweden. Further, let me cite two examples that are meant to support my belief that quality premia are not uncommon: the price differential for airline tickets in first class and economy class seems to be larger than the objective quality differential; a dress sold by some known designer as part of a *haute couture* line may be 10 or 30 times more expensive than a dress made by the same designer for the *prêt-à-porter* line, whereas the objective quality difference may be negligeable. There are, however, obvious problems when it comes to measuring quality objectively, and further empirical studies are called for to establish a body of stylized facts.

The model in this article was set in a world with two types of consumers. It would of course be interesting to check whether the results hold with a continuum of types. Further, as suggested in the previous section, the results under joint purchases are expected to be altered if the coalition formation among consumers is modelled differently. That issue deserves further investigation.

Appendix

■ The proofs of Propositions 1, 2, and 6 follow.

Proof of Proposition 1. The aim of this proof is to identify the binding constraints, which define the transfers $\underline{t}$ and $\overline{t}$ as functions of $\underline{q}$ and $\overline{q}$. To begin with, note that as usual (10) cannot be binding, as implied by (7), $k = 1$, together with (11). Next, the constraints (6) and (9) are slack, as shown by a simple argument: if the average unit prices in the two bundles are different, then buying only bundles with the lowest average unit price is strictly better than any combination of the two bundles; and if the average unit prices are the same, any combination of the two bundles can be replicated by a certain number of only one type of bundle. Further, constraints (4), $k > 1$, do not bind either. Indeed, given that (4), $k = 1$, is satisfied, strict concavity of V implies that these constraints are satisfied with slack. To summarize, the following constraints remain: (11), (5), (7), and (8) for $k > 1$, and (4) for $k = 1$.

Suppose that the constraints (4) for $k = 1$, (5) for $k > 1$, and (8) for $k > 1$ are slack. This will be verified below. Then (11) is binding and there remains only one set of incentive-compatibility constraints, namely, (7), $k \geq 1$, which can now be written

$$\overline{\theta}V(\overline{q}) - \overline{t} \geq \overline{\theta}V(k\underline{q}) - k\underline{\theta}V(\underline{q}) \qquad k \geq 1.$$

This is a continuum of constraints; for a given value of $\underline{q}$, the right-hand sides of the constraints together constitute a function of k. I now show that for every $\underline{q}$, there exists a unique $k^*(\underline{q})$ that maximizes that function $\Phi(k, \underline{q}) \equiv \overline{\theta}V(k\underline{q}) - k\underline{\theta}V(\underline{q})$. The first-order condition for an interior solution is (note that $k^*(\underline{q}) = 1$ if no interior solution exists):

$$\frac{\partial \Phi(k, \underline{q})}{\partial k} = 0 \Leftrightarrow \underline{q}\,\overline{\theta}V'(k(\underline{q})\underline{q}) - \underline{\theta}V(\underline{q}) = 0.$$

The solution to this equation is a maximum, since the second derivative is negative:

$$[\partial^2\Phi(k, \underline{q})]/\partial k^2 = \underline{q}^2\overline{\theta}V''(k\underline{q}) < 0, \quad \forall\, k,$$

by strict concavity of V. The second derivative being strictly negative, $[\partial\Phi(k, \underline{q})]/\partial k$ is a strictly monotone function $\underline{q}$, so there is a unique solution to the above equation for every $\underline{q}$.

Returning to the constraints, (7), $k \neq k^*(\underline{q})$, are implied by (7) for $k = k^*(\underline{q})$. Thus there is just one incentive constraint left, namely, (7) for $k = k^*(\underline{q})$. The monopolist binds this constraint at the optimum, otherwise $\overline{t}$ can be increased without jeopardizing any constraint, contradicting optimality. Together with constraint (11), this constraint uniquely defines the transfers $\underline{t}$ and $\overline{t}$, given any vector of quantities $(\underline{q}, \overline{q})$, yielding the expressions in the proposition.

Next verify that (5) $k > 1$ and (8) $k > 1$ are slack. A sufficient condition is that the implicit unit price exceeds marginal utility at the respective bundles. This is true for the bundle $\{\underline{q}, \underline{t}\} = \{\underline{q}, \underline{\theta}V(\underline{q})\}$ by strict concavity of V. For the bundle $\{\overline{q}, \overline{t}\}$, note that the definition of $k^*(\underline{q})$ and the fact that (7) for $k = k^*(\underline{q})$ is binding (together with (11) binding) imply that the indifference curve of the high-valuation consumer is tangent to the line defined by the equation $t = \underline{p}q$ at any interior solution (see Figure 3). It follows that for the implicit unit price to exceed marginal utility, it is sufficient that $\overline{q} - k^*(\underline{q})\underline{q} > 0$. This will be verified in the proof of Proposition 2.

Last, verify that (4) for $k = 1$ is slack. Using $\underline{t}$ and $\overline{t}$ as defined by the binding constraints as shown above, rewrite it

$$\overline{\theta}[V(\overline{q}) - V(k^*(\underline{q})\underline{q})] \geq \underline{\theta}[V(\overline{q}) - k^*(\underline{q})V(\underline{q})].$$

A sufficient condition is that $V(\overline{q}) - V(k^*(\underline{q})\underline{q}) > 0$, since $\overline{\theta} > \underline{\theta}$, and $V(k^*(\underline{q})\underline{q}) < k^*(\underline{q})V(\underline{q})$ by concavity of V. As mentioned above, it will be verified in the proof of Proposition 2 that $\overline{q} - k^*(\underline{q})\underline{q} > 0$. *Q.E.D.*

Proof of Proposition 2. Replace $\underline{t}$ and $\overline{t}$ by their optimal values as given in Proposition 1 in the objective function to obtain

$$\alpha[\underline{\theta}V(\underline{q}) - c\underline{q}] + (1 - \alpha)[\overline{\theta}V(\overline{q}) - \overline{\theta}V(k^*(\underline{q})\underline{q}) + k^*(\underline{q})\underline{\theta}V(\underline{q}) - c\overline{q}].$$

This is to be maximized with respect to $\underline{q}$ and $\overline{q}$. By Berge's maximum theorem, k^* is a continuous and differentiable function of $\underline{q}$. Assuming that the objective function is concave,[17] the following two first-order conditions define the solution:

$$(1 - \alpha)[\overline{\theta}V'(\overline{q}) - c] = 0$$

$$\alpha[\underline{\theta}V'(\underline{q}) - c] = (1 - \alpha)k^*(\underline{q})[\overline{\theta}V'(k^*(\underline{q})\underline{q}) - \underline{\theta}V'(\underline{q})] + (1 - \alpha)\frac{\partial k^*(\underline{q})}{\partial \underline{q}}[\overline{\theta}V'(k^*(\underline{q})\underline{q})\underline{q} - \underline{\theta}V(\underline{q})].$$

The last term of the second equation is equal to zero from the definition of $k^*(\underline{q})$ (see the proof of Proposition 1), so the expression in the proposition obtains.

[17] It is easy to verify that there exist functions V and parameter values such that this is true, ensuring that I am not working on the empty set.

Clearly, $\bar{q} = \bar{q}^*$. Then I show that $\underline{q} < \underline{q}^*$. From the definition of $k^*(\underline{q})$, $\bar{\theta}V'(k^*(\underline{q})\underline{q}) = [\underline{\theta}V(\underline{q})]/\underline{q}$. By concavity of V, $[\underline{\theta}V(\underline{q})]/\underline{q} > \underline{\theta}V'(\underline{q})$. Therefore, the right-hand side of the second equation above is positive, implying that $\underline{q} < \underline{q}^*$.

Finally, verify that $V(\bar{q}) - V(k^*(\underline{q})\underline{q}) > 0$ (see the proof of Proposition 1). I know that $\underline{\theta}V'(\underline{q}) > c$. Further, I know that $\bar{\theta}V'(k^*(\underline{q})\underline{q}) = [\underline{\theta}V(\underline{q})]/\underline{q}$, which is strictly greater than $\underline{\theta}V'(\underline{q})$ by strict concavity of V. Hence, $\bar{\theta}V'(k^*(\underline{q})\underline{q}) > c$, which implies that $\bar{q} > k^*(\underline{q})\underline{q}$, since $\bar{\theta}V'(\bar{q}) = c$. *Q.E.D.*

Proof of Proposition 6. Step 1: Note first that as in the multiple-purchase case, a simple argument is sufficient to show that constraints (14) and (17) are slack. Then, for a consumer of type $\underline{\theta}$ to prefer the bundle meant for him to some fraction $1/n$ of the bundle meant for a consumer of type $\bar{\theta}$, the point $(1/n)\bar{q}$ on the price line $t = \bar{p}q$ must lie above the indifference curve of the consumer of type $\underline{\theta}$. Figure A1 shows a situation in which (13) for $n = ½$ is not satisfied, whereas (13) for $n = ¾$ is satisfied. Extending this argument, constraints (13) and (16) for $n > 1$ imply that the segment on the price line $t = \bar{p}q$ to the left of the bundle $\{\bar{q}, \bar{t}\}$ must lie above or on the indifference curves of both types of consumers. Similarly, constraints (12) and (15) for $n > 1$ imply that the segment on the price line $t = \underline{p}q$ to the left of the bundle $\{\underline{q}, \underline{t}\}$ must lie above or on the indifference curves of both types of consumers.

Step 2: Prove by contradiction that $\bar{p} \geq \underline{p}$. Suppose that $\bar{p} < \underline{p}$ at an optimum. Since $\underline{q} < \bar{q}$ by necessity, it implies that there exists a nonempty set of points on the line $t = \bar{p}q$ that lies strictly below the indifference curve of the consumer of type $\underline{\theta}$ passing through $\{\underline{q}, \underline{t}\}$. That is, there exists $n > 1$ for which constraint (13) is not satisfied, which contradicts the optimality of $\bar{p} < \underline{p}$.

Step 3: Prove that $\underline{\theta}V'(\underline{q}^j) = \underline{p}^j$ and $\bar{\theta}V'(\bar{q}^j) = \bar{p}^j$. Given that $\bar{p} \geq \underline{p}$, it is easy to draw figures supporting the following. First, suppose that all the constraints are satisfied and that $\bar{\theta}V'(\bar{q}) > \bar{p}$ (it cannot be strictly smaller than $\bar{p}$, because then some incentive constraints would be violated). Then the monopolist can increase its profit by increasing $\bar{q}$ and $\bar{p}$ without jeopardizing any constraints. Hence, $\bar{\theta}V'(\bar{q}) = \bar{p}$. Similarly, if $\underline{\theta}V'(\underline{q}) > \underline{p}$, the monopolist can increase its profit by increasing $\underline{q}$ and $\underline{p}$ (along the indifference curve of consumer $\bar{\theta}$ if (15) for $n = 1$ is binding) without jeopardizing any constraint. Hence, $\underline{\theta}V'(\underline{q}) = \underline{p}$.

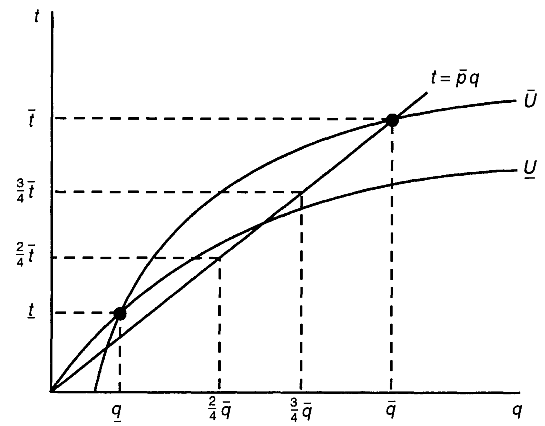
Step 4: Prove that $\underline{q}^j < \underline{q}^*$ and $\bar{q}^j < \bar{q}^*$. It is sufficient to check that $\underline{\theta}V'(\underline{q}) > c$, since $\bar{\theta}V'(\bar{q}) = \bar{p} \geq \underline{p}$ implies that if $\underline{\theta}V'(\underline{q}) > c$, then $\bar{\theta}V'(\bar{q}) > c$. If $\underline{\theta}V'(\underline{q}) \leq c$, the monopolist would be better off by selling only to consumers of type $\bar{\theta}$. Hence, $\underline{\theta}V'(\underline{q}) > c$ at any optimum for which $\underline{q} > 0$.

Step 5: The equalities $\underline{\theta}V'(\underline{q}^j) = \underline{p}^j$ and $\bar{\theta}V'(\bar{q}^j) = \bar{p}^j$ imply that all constraints but one are satisfied. The remaining constraint is (15) for $n = 1$. If it is not binding, the fact that the prices $\underline{p}^j$ and $\bar{p}^j$ are defined by the respective Lerner indices follows from the fact that $\underline{\theta}V'(\underline{q}^j) = \underline{p}^j$ and $\bar{\theta}V'(\bar{q}^j) = \bar{p}^j$. Indeed, this implies that the objective function can be written

$$\alpha[\underline{p}(\underline{q}) - c]\underline{q} + (1 - \alpha)[\bar{p}(\bar{q}) - c]\bar{q}.$$

Since this is separable in $\underline{q}$ and $\bar{q}$, it is as if the monopolist made two separate maximizations, which are standard monopoly-pricing maximizations. Finally, numerical examples show that there exist parameter values such that the constraint (15) for $n = 1$ is slack at the solution given by the two Lerner indices; for

FIGURE A1



other parameter values, however, it is violated, implying that it may be binding at the optimum. In that case, the monopolist maximizes the above expected profit with respect to $\underline{q}$ and $\overline{q}$, under the additional constraint

$$\overline{\theta}[V(\overline{q}) - \overline{q}V'(\overline{q})] = \overline{\theta}V(\underline{q}) - \underline{\theta}\underline{q}V'(\underline{q}).$$

Q.E.D.

References

HAMMOND, P.J. "Markets as Constraints: Multilateral Incentive Compatibility in Continuum Economies." *Review of Economic Studies,* Vol. 54 (1987), pp. 399–412.

INNES, R. AND SEXTON, R. "Customer Coalitions, Monopoly Price Discrimination and Generic Entry Deterrence." *European Economic Review,* Vol. 37 (1993), pp. 1569–1597.

——— AND ———. "Strategic Buyers and Exclusionary Contracts." *American Economic Review,* Vol. 84 (1994), pp. 566–584.

——— AND ———. "Divide and Conquer Price Discrimination in Entry Games with Strategic Buyers." In D. Martimort, ed., *Agricultural Markets: Mechanisms, Failures, and Regulations.* New York: North-Holland, 1996.

KWOKA, J.E., JR. "Market Segmentation by Price-Quality Schedules: Some Evidence from Automobiles." *Journal of Business,* Vol. 65 (1992), pp. 615–628.

MASKIN, E. AND RILEY, J. "Monopoly with Incomplete Information." *RAND Journal of Economics,* Vol. 15 (1984), pp. 171–196.

MCMANUS, B. "Two-Part Pricing with Costly Arbitrage." Mimeo, Department of Economics, University of Virginia, 1998.

MUSSA, M. AND ROSEN, S. "Monopoly and Product Quality." *Journal of Economic Theory,* Vol. 18 (1978), pp. 301–317.

SÄLLSTRÖM, S. "Kvalitet—en nationalekonomisk studie." Mimeo, Department of Economics, Stockholm School of Economics, 1991. (In Swedish.)

TIROLE, J. *The Theory of Industrial Organization.* Cambridge, Mass.: MIT Press, 1988.

WILSON, R. *Nonlinear Pricing.* Oxford: Oxford University Press, 1992.