MANAGEMENT SCIENCE
Vol. 50, No. 12, December 2004, pp. 1660–1673
ISSN 0025-1909 | EISSN 1526-5501 | 04 | 5012 | 1660

informs

DOI 10.1287/mnsc.1040.0291
© 2004 INFORMS

# Nonlinear Pricing of Information Goods

Arun Sundararajan
Leonard N. Stern School of Business, New York University, 44 West 4th Street, New York, New York 10012,
asundara@stern.nyu.edu, http://oz.stern.nyu.edu/

This paper analyzes optimal pricing for information goods under incomplete information, when both unlimited-usage (fixed-fee) pricing and usage-based pricing are feasible and administering usage-based pricing may involve transaction costs. It is shown that offering fixed-fee pricing in addition to a nonlinear usage-based pricing scheme is always profit improving in the presence of nonzero transaction costs, and there may be markets in which a pure fixed-fee is optimal. This implies that the optimal pricing strategy for information goods is almost never fully revealing. Moreover, it is proved that the optimal usage-based pricing schedule is independent of the value of the fixed fee, a result that simplifies the simultaneous design of pricing schedules considerably and provides a simple procedure for determining the optimal combination of fixed-fee and nonlinear usage-based pricing. The introduction of fixed-fee pricing is shown to increase both consumer surplus and total surplus. The differential effects of setup costs, fixed transaction costs, and variable transaction costs on pricing policy are described. These results suggest a number of managerial guidelines for designing pricing schedules. For instance, in nascent information markets, firms may profit from low fixed-fee penetration pricing, but as these markets mature, the optimal pricing mix should expand to include a wider range of usage-based pricing options. Minimum fees, quantity discounts, and adoption levels across the different pricing schemes are characterized, strategic pricing responses to changes in market characteristics are described, and the implications of the paper's results for bundling and vertical differentiation of information goods are discussed.

*Key words*: pricing; price discrimination; information goods; digital goods; adverse selection; screening; principal agent; incentives; mechanical design; incentive compatible; monopoly; incomplete information; second-degree price discrimination; two-part tariff; vertical differentiation; software; Internet
*History*: Accepted by Rajiv D. Banker, information systems; received May 13, 2002. This paper was with the author $2\frac{1}{2}$ months for 1 revision.

## 1. Introduction

Nonlinear usage-based pricing is a popular price-discrimination technique that has been analyzed extensively in the context of the electricity and long-distance telephone markets (Wilson 1993). This form of price discrimination is used by many sellers of information goods. For instance, corporate software manufacturers associate the price they charge each customer with their expected usage of the software by basing prices on the total processing speed of the servers on which the software is licensed to run. ASP and application syndication models enable a variety of more direct usage-based software pricing.[1] Apple's iTunes music service is priced on a per-song basis.

In contrast, there are numerous examples of *fixed-fee* pricing for information goods under which customers pay a fixed periodic price that is independent of usage. Most ISPs charge residential customers a flat monthly subscription fee. The *Wall Street Journal Online* offers unrestricted access for a fixed annual fee. Sprint PCS recently switched from a per-Mb pricing model to giving consumers unlimited wireless Web access for a fixed monthly fee.

Additionally, many sellers of information goods use a combination of fixed-fee and usage-based pricing. A customer of IBM's zSeries software can opt to pay a flat fee for unlimited usage, or to use a reporting tool that tracks and charges for software usage on a monthly basis. In addition to their regular per-minute pricing scheme, Sprint and AT&T both offer fixed-fee long-distance telephony, and MCI sells an unlimited-usage local, long-distance, and Internet service package for a fixed monthly fee. Other information goods featuring both fixed-fee and usage-based pricing include network bandwidth, the OCLC library information services, and industry research reports.

Pricing policies for information goods which include an unlimited-usage fixed-fee conflict with well-known results from nonlinear pricing theory (Maskin and Riley 1984, Wilson 1993), which have shown that under some fairly general assumptions, the optimal pricing policy for a monopolist should always be strictly based on usage. One goal of this paper is to demonstrate that this disconnect between theory and practice can be rigorously explained by recognizing two unique aspects of pricing *information goods*.

[1] While there are a number of diverse models that these ASPs use to price their services, most involve some form of rental pricing (Susarla et al. 2003).

(a) An increase in the usage of an information good by a customer imposes near-zero or zero direct variable costs of production on the seller. This makes unlimited-usage fixed-fee pricing increasingly viable for these goods.

(b) There are typically fixed and variable transaction costs associated with the *administering* of any usage-based pricing schedule. These costs are especially relevant when designing pricing schedules for information goods because they are significant relative to the near-zero variable costs of production that characterize information goods.[2]

The costs of administering usage-based pricing schedules (henceforth called *transaction costs*) that are alluded to above stem from the many different activities that are necessary to viably administer a usage-based pricing schedule. A seller must monitor and record the details of usage for each individual customer. Even if the direct costs of electronic monitoring are low, the related administration of billing, payment and settlement, and dispute resolution is expensive. When charged a constant periodic fee, customers often are comfortable with automated and direct periodic charges (to a credit card, for instance) with no itemized statement. In contrast, when pricing is based on usage, the seller periodically may need to present each customer with an itemized statement of their usage. This requires administering a reliable process for the delivery of these statements. The seller may need to mail and process nonelectronic statements and payments for those customers unable to use, or uncomfortable with, electronic billing.

Even computer-based usage monitoring systems are prone to error (an example familiar to most readers may be their monthly statements of long-distance telephone usage), and this imposes additional dispute resolution costs on the seller, which may be incurred even when a customer merely conjectures that they have been incorrectly billed. The costs of handling related customer service calls can be substantial, even with relatively low error rates. Moreover, errors in billing, or perceptions thereof, can lead to future lost sales. The seller may need to maintain auditable records of usage for each individual customer, in the event of future customer disputes. Clearly, many of these per-customer transaction costs (itemized invoicing, for instance), are triggered by any positive usage, while others (such as usage recording and error resolution) are likely to increase with customer usage. There may also be one-time setup costs or periodic infrastructure costs to install and run reliable processes that support the activities associated exclusively with usage-based pricing.

[2] In contrast, when variable costs are high relative to these costs of administering usage-based pricing, as is the case for many physical goods, these transaction costs affect pricing only minimally.

These drivers of transaction costs are unrelated to the production or delivery of the actual digital good being sold and are simply a consequence of offering a usage-based pricing scheme. Consequently, sellers may want to mitigate these costs by offering some or all of their customers a fixed-fee pricing scheme. This may seem especially attractive for sellers of mass-market digital goods, where transaction costs are significant relative to the potential revenue firms may obtain from each of their customers. Moreover, each customer is likely to be willing to pay more for the option of unlimited usage. There is a trade-off, however—fixed-fee pricing precludes second-degree price discrimination based on usage, and this is likely to adversely affect seller revenues.

The model in this paper analyzes this trade-off by deriving the optimal combination of the unlimited usage fixed-fee and the usage-based nonlinear pricing function, with very general assumptions about customer preferences and transaction costs and under incomplete information. It shows that *any* positive fixed or variable transaction costs make it optimal for the monopolist to offer their customers the option of a fixed-fee pricing scheme. It also establishes that the optimal choice of the usage-based pricing schedule is independent of the value of the fixed-fee—a result that simplifies the simultaneous design of pricing schedules significantly. Managerial implications for pricing design, volume discounting, adoption patterns, and market evolution are also discussed.

The optimal pricing of information systems has been studied quite extensively, most often with a focus on congestion pricing. This body of work includes a queuing model of ASP pricing by Cheng and Koehler (1999), an analysis of pricing service facilities with nonlinear delay costs by Dewan and Mendelson (1990), a model of usage-based pricing in a network based on the theoretical framework of Gupta et al. (1997), the seminal paper by Mendelson (1985) on pricing computer services by internalizing delay externalities, which was followed by a model of variable priority pricing under asymmetric information for queues by Mendelson and Whang (1990), and a model of optimal IS pricing with network externalities by Westland (1992). These papers focus specifically on either contrasting usage-based pricing with alternate schemes, or pricing information systems under asymmetric information, which makes their models most relevant to this paper. A more detailed survey is available from the author on request.

This paper adds to the literature by presenting a new model directly contrasting fixed-fee and nonlinear pricing of information goods under a more general analytical framework and underlining the importance of transaction costs (highlighted briefly by Varian 2000) in the design of optimal pricing schedules for information goods. Two related and active areas

of information-goods research into pricing and market segmentation—bundling and vertical differentiation (versioning)—have indirectly shed some light on the trade-offs between fixed-fee and usage-based pricing, and this relationship is discussed further in §6.

## 2. Model

### 2.1. Firm and Customers

A monopoly firm sells an information good[3] that may be used by customers in varying quantities. Variable costs of production to the firm—creating copies or providing access to the product—are zero. Customers are heterogeneous, indexed by their type $\theta \in [\underline{\theta}, \bar{\theta}]$. The preferences of a customer of type $\theta$ are represented by the function

$$W(q, \theta, p) = U(q, \theta) - p, \tag{1}$$

where $q$ is the quantity of the product used and $p$ is the total price paid by the customer. The function $U(q, \theta)$ is referred to as the customer's utility function. Numbered subscripts to functions denote partial derivatives with respect to the corresponding argument. For instance, $U_1(q, \theta)$ is the partial derivative of $U$ with respect to its first argument, and $U_{12}(q, \theta)$ is the cross-partial of $U$ with respect to its first and second arguments. This notation is preserved throughout the paper.

The utility function $U(q, \theta)$ has the following properties, for each $\theta \in [\underline{\theta}, \bar{\theta}]$:

1. Increasing and concave value:
$$U(0, \theta) = 0;\ U_1(q, \theta) \geq 0,\ U_{11}(q, \theta) < 0 \quad \text{for all } q.$$
2. Higher customer types get higher utility:
$$U_2(q, \theta) > 0 \quad \text{for all } q > 0.$$
3. Spence-Mirrlees single-crossing condition:
$$U_{12}(q, \theta) > 0 \quad \text{for all } q.$$
4. Nonincreasing absolute risk aversion:
$$\frac{\partial}{\partial \theta}(-U_{11}(q, \theta)/U_1(q, \theta)) \leq 0.$$
5. Finite maximum value:
$$\lim_{q \to \infty} U(q, \theta) = v(\theta) < \infty.$$

Property 2 simply states that type $\theta$ orders customers based on the value they get from the product. In addition, Property 3 implies that higher types get a higher increase in value than lower types from the same increase in usage. Property 4 states that higher types are also increasingly less risk averse. Property 5 bounds the maximum value a customer can derive from the product, ensuring that the monopolist cannot make infinite profits by offering unlimited-usage pricing. As indicated above, the utility derived from maximal usage by type $\theta$ is represented using the function

$$v(\theta) = \lim_{q \to \infty} U(q, \theta). \tag{2}$$

The firm does not observe the type of any customer but knows $F(\theta)$, the probability distribution of types in the customer population, and the corresponding density function $f(\theta)$, which is strictly positive for all $\theta \in [\underline{\theta}, \bar{\theta}]$.

### 2.2. Pricing Schedules

The information good is priced using one or both of two kinds of pricing schedules (also called *contracts*).

*Fixed-Fee.* A fixed-fee contract specifies a price $T$ to be paid by the customer in exchange for unlimited usage of the information good. There are no transaction costs associated with a fixed-fee contract—the customer simply pays the firm the deterministic, prespecified price $T$.

*Usage-Based.* A usage-based contract assigns a specific price to each level of usage $q$. The firm cannot explicitly distinguish between customer types prior to contracting. Thus, the entire menu of quantity-price pairs must be available to all customers. The revelation principle ensures that the firm can restrict its attention to direct mechanisms—that is, usage-based contracts in which one specific quantity-price pair is designed for each customer, and in which it is rational and optimal for the customer to choose the quantity-price pair that was designed for him or her.[4] The usage-based contract is represented by a menu of quantity-price pairs $(q(t), \tau(t))$, where $t \in [\underline{\theta}, \bar{\theta}]$. This menu must satisfy two standard constraints:

[IC]: For each $\theta$,

$$U(q(\theta), \theta) - \tau(\theta) \geq U(q(t), \theta) - \tau(t), \quad \text{for all } t \in [\underline{\theta}, \bar{\theta}].$$

[IR]: For each $\theta$,

$$U(q(\theta), \theta) - \tau(\theta) \geq 0.$$

When the menu of quantity-price pairs satisfies (IC) and (IR), every customer of type $\theta$ will choose the pair $q(\theta), \tau(\theta)$. For brevity, a usage-based contract satisfying these constraints is simply referred to as *incentive compatible*. An incentive-compatible usage-based contract is said to be *optimal* for a subinterval $[\theta_L, \theta_H]$ if it yields profits that are at least as high as any other incentive-compatible usage-based contract designed exclusively for customers in the subinterval $[\theta_L, \theta_H]$. When no subinterval is mentioned, optimality applies to the entire interval $[\underline{\theta}, \bar{\theta}]$.

The firm bears transaction costs of $C(q)$ for each customer who adopts the *usage-based* contract and uses a

[3] This may either be a homogeneous good (such as bandwidth) or a bundle of related heterogeneous quantity units (for instance, a library of MP3 songs, where each successive unit of consumption is a download of a different song).

[4] This kind of formulation is standard in models of price screening—see, for instance, Anderson (1996, §2). A good exposition of mechanism design, the revelation principle, and its applications to pricing can be found in Fudenberg and Tirole (1991, Chapter 7).

**Table 1 Summary of Key Notation**

| Symbol | Explanation |
|---|---|
| $U(q, \theta)$ | Utility that customer type $\theta$ gets from usage level $q$. |
| $v(\theta)$ | Maximum utility that customer type $\theta$ can get from usage. $v(\theta) = \lim_{q\to\infty} U(q, \theta)$. |
| $[\underline{\theta}, \bar{\theta}]$ | Range of possible customer types $\theta$. |
| $f(\theta), F(\theta)$ | Density and distribution functions of the customer type distribution. |
| $C(q)$ | Transaction costs of administering a usage-based pricing schedule for a customer who uses quantity $q$. In §4, $C(q)$ takes the form $K + c(q)$ for $q > 0$. |
| $q(\theta), \tau(\theta)$ | Usage-based contract (menu of quantity-price pairs) that is incentive compatible. For a specific $\theta$, $q(\theta)$ is the quantity and $\tau(\theta)$ is the price for that quantity. |
| $q^*(\theta), \tau^*(\theta)$ | Optimal incentive-compatible usage-based contract. |
| $T$ | Unlimited-usage fixed-fee price. |
| $\theta_K$ | Lowest customer type that uses a positive quantity under the usage-based contract. |
| $\theta_F$ | Lowest customer type that is indifferent between fixed-fee and usage-based pricing. |

quantity $q$. The drivers of these costs are discussed at length in §1. Key notation is summarized in Table 1.

### 2.3. Interaction Between the Firm and Its Customers

The sequence of interaction between the firm and its potential customers is as follows:

1. The firm designs and posts either an incentive-compatible usage-based contract $(q(\cdot), \tau(\cdot))$, a fixed-fee contract $T$, or both.

2. Each customer either chooses to purchase using one of the two pricing schedules or chooses not to purchase. If a customer chooses the fixed-fee contract, a fixed payment of $T$ is made to the firm. Because the usage-based pricing schedule $(q(\cdot), \tau(\cdot))$ is incentive compatible, a customer of type $\theta$ who chooses this pricing schedule uses a quantity $q(\theta)$ and makes a payment of $\tau(\theta)$.

The problem for a customer of type $\theta$ is to choose between paying a fixed fee $T$ for maximal usage (and a corresponding value $v(\theta)$), paying $\tau(\theta)$ for a usage level $q(\theta)$, or not participating. The problem of the firm is to choose the kinds of contracts (fixed fee, usage based, both) to offer and to design these contract(s) to maximize ex ante expected profits.

## 3. The Optimality of Offering Fixed-Fee Pricing

This section describes how fixed-fee pricing affects customer choice and establishes that in the presence of nonzero transaction costs, a fixed-fee pricing scheme always improves profits for the seller of an information good. First, a preliminary result is established.

Lemma 1. *If $q(\theta)$, $\tau(\theta)$ is an incentive-compatible contract, then*

(a) $q_1(\theta) \geq 0$, $\tau_1(\theta) \geq 0$.

(b) $U(q(\theta), \theta) - \tau(\theta)$ *is nondecreasing in* $\theta$.

Unless specified otherwise, all proofs are in the appendix. The expression $U(q(\theta), \theta) - \tau(\theta)$ is the surplus obtained by a customer of type $\theta$ from the contract $q(\cdot), \tau(\cdot)$ and is commonly referred to as the *informational rent* for type $\theta$.

### 3.1. The Impact of a Fixed Fee on Customer Choice

The main result of this subsection establishes that when a fixed fee is offered along with *any* incentive-compatible usage-based contract, then customers typically bifurcate into two intervals, with lower types adopting the usage-based contract and higher types adopting the fixed fee.

Suppose the firm offers a fixed-fee contract $T$ along with a usage-based contract $(q(\cdot), \tau(\cdot))$, which is incentive compatible in the *absence* of $T$. The surplus that a customer of type $\theta$ gets from choosing the fixed-fee contract is $v(\theta) - T$. Therefore, a customer of type $\theta$ will choose the fixed-fee contract if and only if

$$v(\theta) - T \geq U(q(\theta), \theta) - \tau(\theta), \tag{3}$$

where it is assumed that an indifferent customer chooses the fixed-fee contract. Note that Equation (3) is equivalent to

$$v(\theta) - U(q(\theta), \theta) + \tau(\theta) \geq T. \tag{4}$$

The expression on the left-hand side of (4) has a simple economic interpretation. It is the difference between the maximum value $v(\theta)$ obtainable by type $\theta$ from the information good and the informational rent $[U(q(\theta), \theta) - \tau(\theta)]$ that type $\theta$ gets from using under the usage-based contract. Consequently, it is the maximum fixed fee that the firm can charge if they want type $\theta$ to adopt the fixed fee. Lemma 2 shows that this maximum amount is increasing in $\theta$.

Lemma 2. *For any incentive-compatible usage-based contract $(q(\cdot), \tau(\cdot))$, the function*

$$\psi(\theta) = v(\theta) - U(q(\theta), \theta) + \tau(\theta) \tag{5}$$

*is strictly increasing for all $\theta$ in $[\underline{\theta}, \bar{\theta})$.*

This lemma leads to the following proposition.

Proposition 1. *If the firm introduces a fixed fee $T$ in addition to an existing usage-based contract $(q(\cdot), \tau(\cdot))$ which is incentive-compatible in the absence of $T$, this affects customer choice in exactly one of the following three ways.*

(a) *If $v(\underline{\theta}) - T \geq U(q(\underline{\theta}), \underline{\theta}) - \tau(\underline{\theta})$, then all customers adopt the fixed-fee contract.*

(b) *If $v(\bar{\theta}) - T < U(q(\bar{\theta}), \bar{\theta}) - \tau(\bar{\theta})$, then all customers continue to adopt the usage-based contract.*

(c) *If $v(\underline{\theta}) - T < U(q(\underline{\theta}), \underline{\theta}) - \tau(\underline{\theta})$ and $v(\bar{\theta}) - T \geq U(q(\bar{\theta}), \bar{\theta}) - \tau(\bar{\theta})$, then customers of type $\theta \in [\underline{\theta}, \theta_F)$ continue to adopt the usage-based contract, and customers of*

*type $\theta \in [\theta_F, \bar{\theta}]$ switch to the fixed-fee contract, where*

$$\theta_F = \min\{\theta:\ v(\theta) - U(q(\theta), \theta) + \tau(\theta) = T\}. \qquad (6)$$

Proof. Combining (4) and the fact that $\psi(\theta)$ is increasing (as shown in Lemma 2) establishes that if type $\hat{\theta}$ adopts the fixed-fee contract, then so do all types $\theta > \hat{\theta}$. In addition, if type $\hat{\theta}$ does not adopt the fixed-fee contract, then neither does any type $\theta < \hat{\theta}$. This proves parts (a) and (b). If the conditions for (c) hold, then because $\psi(\theta)$ is increasing in $[\underline{\theta}, \bar{\theta}]$, this ensures that there will be at least one type $\theta$ for which $\psi(\theta) = T$. Because $\theta_F$ is the lowest such value of $\theta$ and indifferent types adopt the fixed-fee contract, this proves part (c), which completes the proof. □

As illustrated by Figure 1, $T$ is always higher than $\tau(\theta_F)$ but may be lower than $\tau(\bar{\theta})$. Therefore, while a fraction $[\theta_F, \theta_T]$ of customer types always pay a higher price, there may be a fraction of customer types $[\theta_T, \bar{\theta}]$ that pay a lower price under the fixed fee. The firm gains revenue from the former set but may lose revenue from the latter set. It also lowers costs in the interval $[\theta_F, \bar{\theta}]$ as a consequence of having no transaction costs from all customers adopting $T$.

### 3.2. Profit-Improving Fixed Fees

This subsection establishes that the profits of a seller of information goods can always be strictly improved by the introduction of an unlimited-usage fixed fee. First, I establish that under the optimal usage-based contract in the absence of a fixed fee, the firm's profit from each customer is positive and nondecreasing in their type.

Lemma 3. *If $q^*(\theta), \tau^*(\theta)$ is the optimal usage-based contract in the absence of a fixed fee, then*

(a) *$\tau^*(\theta) - C(q^*(\theta))$ is nondecreasing in $\theta$,*

(b) *$\tau^*(\theta) - C(q^*(\theta)) \geq 0$ for all $\theta$.*

The main result of the section now follows.

Proposition 2. *If transaction costs are nonzero—that is, if $C(q) > 0$ for $q > 0$—then it is always profit improving for the seller of an information good to offer a fixed-fee contract.*

Proof. Let $q^*(\theta), \tau^*(\theta)$ be the optimal usage-based contract in the *absence* of a fixed fee. If $q^*(\bar{\theta}) = 0$, then from Lemma 1, $q^*(\theta) = 0$ for all $\theta$, profits are zero, and the fixed fee $T = v(\underline{\theta})$ strictly improves profits. If $q^*(\bar{\theta}) > 0$, then $C(q^*(\bar{\theta})) > 0$. Now, choose any fixed fee $T$ such that

$$[\tau^*(\bar{\theta}) - C(q^*(\bar{\theta}))] < T < \tau^*(\bar{\theta}). \qquad (7)$$

Because $v(\bar{\theta}) \geq U(q, \bar{\theta})$ for all $q$, it follows that

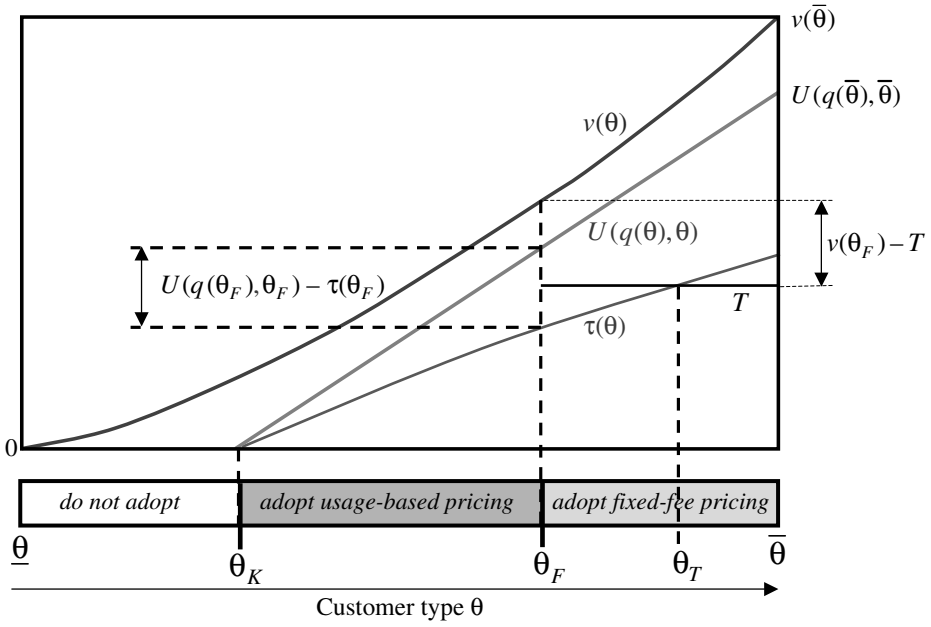
$$v(\bar{\theta}) - T > U(q^*(\bar{\theta}), \bar{\theta}) - \tau^*(\bar{\theta}), \qquad (8)$$

and Proposition 1 ensures that a fraction $[\theta_F, \bar{\theta}]$ of customer types (perhaps all) will adopt the fixed fee. From Equation (7), we know that $T > \tau^*(\bar{\theta}) - C(q^*(\bar{\theta}))$. Using Lemma 3, this implies that

$$\tau^*(\theta) - C(q^*(\theta)) < T \quad \text{for all } \theta \in [\theta_F, \bar{\theta}]. \qquad (9)$$

As a consequence, the profits from each customer type in $[\theta_F, \bar{\theta}]$ are strictly increased by the introduction of the fixed fee $T$. Proposition 1 ensures that for $\theta_F > \underline{\theta}$, customers in $[\underline{\theta}, \theta_F)$ continue to adopt the usage-based contract $q^*(\theta), \tau^*(\theta)$, and profits from this segment remain unchanged. Therefore, the firm's overall

**Figure 1 The Impact of Fixed-Fee Pricing on Customer Choice**



profits are increased by the introduction of $T$, which completes the proof. □

In the absence of changes to $(q^*(\cdot), \tau^*(\cdot))$, the increase in the firm's profits is feasible as a result of two separate effects. The first is the elimination of transaction costs for the adopters of the fixed fee. The second is an increase in total surplus from the higher usage levels of these adopters, which may induce a net increase in revenues. Recall that the firm bears no additional variable production costs from this increase in usage of the information good.

# 4. The Optimal Combination of Fixed-Fee and Usage-Based Pricing

Proposition 2 establishes the desirability of fixed-fee pricing for information goods under very general conditions—for any positive transaction cost function $C(q)$ and for any absolutely continuous customer type distribution $F(\theta)$. In this section, the structure of the optimal fixed-fee and usage-based pricing schedules is established in more detail. The transaction cost function $C(q)$ is assumed to take the following form:

$$\begin{aligned} &C(q) = 0 \quad \text{for } q = 0, \\ &C(q) = K + c(q) \quad \text{for } q > 0, \end{aligned} \tag{10}$$

where $K \geq 0$, $c_1(q) \geq 0$ (nondecreasing variable costs), and variable costs are "not too concave":

$$\frac{c_{11}(q)}{c_1(q)} > \frac{U_{11}(q, \theta)}{U_1(q, \theta)} \quad \text{for all } \theta. \tag{11}$$

The condition (11) above is met by any linear or convex cost function, a cost function that is positive and constant for all $q > 0$, as well as a range of concave cost functions.[5] As described in §1, many per-customer transaction costs may be triggered by any positive usage level, simply because a usage-based pricing schedule has to be administered. Other transaction costs are proportionate to the level of usage of the information good. The flexible specification in (10) allows both these kinds of costs, and also admits varying levels of economies of scale.

The type distribution is restricted to having a nonincreasing inverse hazard rate:

$$\frac{\partial}{\partial \theta}\left(\frac{1 - F(\theta)}{f(\theta)}\right) \leq 0 \quad \text{for all } \theta.$$

This property is satisfied by most commonly used unimodal distributions, and the restriction is standard in models of price discrimination.

[5] Equation (11) ensures that the firm's profit function is strictly quasi-concave. It may not be a necessary condition for the results that follow, but in its absence, the optimal contracts cannot be mathematically characterized, because the point-wise optimization problem is not guaranteed to have a unique local maximum.

## 4.1. Optimal Usage-Based Pricing in the Absence of a Fixed Fee

The result in this subsection characterizes the structure of the optimal usage-based contract $q^*(\theta), \tau^*(\theta)$ when the firm *does not* offer a fixed-fee contract.

Proposition 3. *The optimal usage-based contract* $(q^*(\theta), \tau^*(\theta))$ *in the absence of a fixed fee takes the following form*:

$$q^*(\theta) = 0 \quad \textit{for } \theta < \theta_K; \tag{12}$$

$$\tau^*(\theta) = 0 \quad \textit{for } \theta < \theta_K; \tag{13}$$

$$q^*(\theta) = q^0(\theta) \quad \textit{for } \theta \geq \theta_K; \tag{14}$$

$$\tau^*(\theta) = U(q^*(\theta), \theta) - \int_{\theta_K}^{\theta} U_2(q^*(x), x)\, dx \quad \textit{for } \theta \geq \theta_K, \tag{15}$$

*where* $q^0(\theta)$ *and* $\theta_K$ *are defined by*

$$U_1(q^0(\theta), \theta) = c_1(q^0(\theta)) + U_{12}(q^0(\theta), \theta)\frac{1 - F(\theta)}{f(\theta)} \quad \forall \theta \tag{16}$$

$$\theta_K = \min\left\{\theta\colon U(q^0(\theta), \theta) - U_2(q^0(\theta), \theta)\frac{1 - F(\theta)}{f(\theta)} \geq [K + c(q^0(\theta))]\right\} \tag{17}$$

*Also, if* $U_{122}(q, \theta) \leq 0$, *then* $q_1^*(\theta) > 0$ *for all* $\theta > \theta_K$, *and the optimal contract is therefore fully revealing for all customers who use nonzero quantities*.

Proposition 3 indicates that increases in the transaction costs tend to increase prices correspondingly and shrink the fraction of adopters of the information good. Furthermore, for any $K > 0$, the optimal usage-based contract is a nonlinear two-part tariff. Because $U_1(0, \theta) > 0$, it is clear from Equation (17) that $q^0(\theta_K) > 0$. As a consequence, there is a minimum price $\tau^*(\theta_K)$ for usage above zero but lower than $q^*(\theta_K)$, and variable pricing beyond that. This is a commonly observed pricing structure for digital goods, and Proposition 3 establishes that it it always induced when per-customer transaction costs have a usage-independent component $K > 0$.

## 4.2. Independence of Fixed-Fee and Usage-Based Pricing

The result of this subsection shows that the optimal usage-based pricing schedule in the presence of an unlimited-usage fixed fee is *independent* of the value of the fixed fee. As a consequence, the simultaneous derivation of the optimal combination of usage-based and fixed-fee pricing is simplified considerably.

Using Proposition 3, we know how to design the pricing schedule $(q^*(\theta), \tau^*(\theta))$ that is optimal in the *absence of* a fixed fee. While Proposition 2 has established the desirability of a fixed-fee contract $T$ in addition, it does not indicate what the optimal value of $T$ should be. When a fraction $[\theta_F, \bar{\theta}]$ of

the customers no longer adopt the optimal usage-based contract $(q^*(\theta), \tau^*(\theta))$, the firm may want to redesign pricing for the remaining customer types $[\underline{\theta}, \theta_F]$ in a profit-improving way. This may change the value of the lowest type $\theta_F$ who is indifferent. Consequently, to evaluate the net profit impact of each feasible fixed fee, one needs to consider optimally redesigned usage-based contracts for a continuum of subintervals. Moreover, it is not guaranteed that a combination of this form is in fact optimal—for instance, a higher value of $T$ and a correspondingly constrained incentive-compatible contract may be more profitable. Therefore, to find the optimal combination, the firm needs to vary $T$ while simultaneously considering all feasible incentive-compatible contracts (and their profits from corresponding adoption) under the constraints imposed by the existence of each $T$. Proposition 4 describes the solution to this problem.

Proposition 4. *The optimal usage-based contract in the presence of the optimal fixed fee is independent of the value of the fixed fee, and is identical to the optimal usage-based contract in the absence of any fixed fee. Consequently, the optimal combination of fixed-fee and usage-based contracts can be constructed as follows.*

(a) *Determine the optimal usage-based contract* $(q^*(\cdot), \tau^*(\cdot))$ *using Proposition* 3.

(b) *Find the optimal interval of types* $[\theta_F^*, \bar{\theta}]$ *that should adopt the fixed-fee contract by solving*

$$\theta_F^* = \arg\max_{\theta_F} \int_{\theta_K}^{\theta_F} [\tau^*(\theta) - C(q^*(\theta))] f(\theta)\, d\theta + [1 - F(\theta_F)][v(\theta_F) - U(q^*(\theta_F), \theta_F) + \tau^*(\theta_F)]. \quad (18)$$

(c) *Determine the optimal fixed-fee contract*

$$T^* = v(\theta_F^*) - U(q^*(\theta_F^*), \theta_F^*) + \tau^*(\theta_F^*). \quad (19)$$

Proposition 4 is a surprising result. It shows that when the firm offers the optimal fixed-fee contract, and this contract is adopted by a positive fraction of customers, the optimal usage-based contract offered to the remaining customers remains unchanged, even though the usage-based contract is being designed for a different (and smaller) interval of customers. Were the seller to design a usage-based contract exclusively for this smaller interval, ignoring the fixed fee, it would always be different from $(q^*(\theta), \tau^*(\theta))$. When the firm does take the introduction of the fixed fee into account, this introduces a new (and infinite) set of individual rationality inequality constraints, which change as one varies either the level of the fixed fee $T$ or the subinterval $[\theta_F, \bar{\theta}]$ that the firm wants to induce to adopt $T$. Either of these changes necessitates a redesign of optimal usage-based pricing. Proposition 4 reduces this complicated sequence to a simple problem of determining a globally optimal usage-based pricing schedule and then solving an unconstrained maximization problem in a single variable.

An immediate corollary of Proposition 4 is that the introduction of the optimal fixed-fee contract (and the consequent adjustment of customer usage) does not reduce the surplus of any customer, relative to the scenario in which only usage-based pricing is offered. Because the surplus of those customers adopting the fixed-fee increases, this means that consumer surplus strictly increases as well. Proposition 2 ensures that firm profits also strictly increase, implying that total surplus also increases on account of the fixed fee.

Apart from per-customer transaction costs $C(q)$, the seller may also bear a *setup cost* of administering a usage-based contract, which is incurred if the seller wishes to offer usage-based pricing to any fraction of customers. This kind of cost increases the likelihood that the seller will offer just fixed-fee pricing and forego usage-based pricing entirely. However, if the seller still chooses to offer usage-based pricing, this cost does not alter the optimal pricing schedules and does not affect any of the results derived above. This is discussed further in the paper's extended appendix (available at http://mansci.pubs.informs.org/ecompanion.html).

## 5. Example and Discussion

The general results derived above are applied to a simple example, and some managerial guidelines are drawn from this exercise. In the example, the customers' utility function is assumed to be

$$U(q, \theta) = (w + \theta)q - \tfrac{1}{2}q^2 \quad \text{for } q \le w + \theta, \quad (20)$$

$$U(q, \theta) = \frac{(w + \theta)^2}{2} \quad \text{for } q > w + \theta. \quad (21)$$

It is easily verified that $U_{122}(q, \theta) = 0$, and therefore, the conditions in Proposition 3 describe the unique optimal usage-based contract. The transaction cost function takes the form $C(q) = K + cq$. In addition, customer types are assumed to have the beta distribution[6] with parameters $a = 0$, $b \ge 1$, and support $\theta \in [0, 1]$.

Applying Proposition 3 yields the optimal usage-based contract $(q^*(\theta), \tau^*(\theta))$, summarized in Table 2. The corresponding expressions for $\theta_K$ (the lowest type adopting the usage-based contract, as specified in Proposition 3) and $\theta_F^*$ (the lowest type adopting the fixed-fee contract, as specified in Proposition 4) are

[6] The general form of the beta density function is

$$B(\theta; a, b) = \frac{\theta^{a-1}(1 - \theta)^{b-1}}{\beta(a, b)},$$

where $\beta(a, b) = \int_0^1 x^{a-1}(1 - x)^{b-1}\, dx$ is the beta function with parameters $a$ and $b$.

**Table 2 Optimal Contracts and Indifferent Customer Types in the Example**

| Example: $U(q,\theta)=(w+\theta)q-\frac{1}{2}q^2$; $C(q)=K+cq$; $F(\theta)=1-(1-\theta)^b$ | |
|---|---|
| Optimal usage-based contract: | $q^*(\theta)=(\theta+w)-\left(c+\dfrac{1-\theta}{b}\right)$; |
| | $\tau^*(\theta)=\dfrac{1+bc+w}{1+b}q^*(\theta)-\dfrac{[q^*(\theta)]^2}{2(1+b)}$ |
| (A) Uniform type distribution: $b=1$, $w\geq 0$, $c\geq 0$, $K\geq 0$ | |
| Lowest adopter of usage-based contract | $\theta_K=\dfrac{2K+1-(w-c)^2}{2(1+w-c)}$ |
| Lowest adopter of fixed-fee contract | $\theta_F^*=1-\left(\sqrt{4c^2+2c\left(1+w-\dfrac{c}{2}\right)+2K}-2c\right)$ |
| (B) Positively-skewed type distribution: $b>1$, $w=0$, $c\geq 0$, $K=0$ | |
| Lowest adopter of usage-based contract | $\theta_K=\dfrac{1+bc}{1+b}$ |
| Lowest adopter of fixed-fee contract | $\theta_F^*=1-\dfrac{b(\sqrt{4b^2c^2+c(2-c)(2b-1)}-2bc)}{2b-1}$ |

summarized in Table 2 for two separate cases. In the first case, customer types are uniformly distributed ($b=1$). In the second case, the customer type distribution is positively skewed ($b>1$), and the values of $K$ and $w$ are normalized to zero. A detailed example in which customer types are exponentially distributed is presented in the paper's extended appendix.

### 5.1. Transaction Costs and Adoption Levels

The changes induced in adoption as transaction costs vary are illustrated in Figure 2, for the case of uniformly distributed $\theta$. An increase in the fixed transaction cost $K$ results in a strict increase in $\theta_K$ and a strict decrease in $\theta_F^*$, thereby increasing the fraction of customers who adopt the fixed fee, reducing the set of customers who adopt the usage-based contract, and also reducing total adoption. A directionally identical change occurs with an increase in $c$. Therefore, a reduction in either the fixed transaction costs of administering usage-based pricing (from an increase in the adoption of online billing, for instance) or of the variable transaction costs of administering usage-based pricing (from increased ease of using online customer support, for instance) should induce sellers to alter their pricing structures in a manner that shifts users away from fixed-fee pricing.

### 5.2. Profits, Surplus, and Welfare

Based on the expressions derived in Table 2, deriving expressions for profits, customer surplus and total surplus is straightforward. These expressions are algebraically cumbersome and are omitted for brevity, but are illustrated in Figure 3 for $K=0$. Profits are strictly decreasing in both $c$ and $K$ upto a point, after which they are constant, because no more customers adopt usage-based pricing. Consumer surplus decreases with $c$ and $K$ initially, then increases, and is often higher at those values of $(c,K)$ for which only usage-based contracts are offered than it is for $c=K=0$. Interestingly, total surplus (the sum of profits and consumer surplus) decreases initially in $c$, but then increases as $c$ increases, though it attains its maximum at $c=K=0$.

The economic intuition behind these observations is explained in Figure 4, for changes in $c$, with $K$ normalized to zero. There are two sets of effects that an increase in $c$ has on total surplus. First, there is a *negative indirect effect*—owing to the reduction in both the number of adopters of the usage-based contract, and the quantity used by each, total customer utility reduces. Second, there is a *direct cost effect*—the transaction costs borne by the firm per unit of usage increases, which changes both firm profits and total surplus. However, the decrease in usage by the adopters of usage-based pricing may offset this cost increase. There are also two *positive indirect effects*, however. An increase in $c$ increases the number of adopters of the fixed-fee contract, and all of these customers enjoy higher utility levels, at their maximum value $v(\theta)$. In addition, these customers no longer impose the transaction costs $K+cq^*(\theta)$ on the firm. The negative effects dominates for lower values of $c$, after which the positive indirect effects dominate.

### 5.3. Customer Value and Market Evolution

The changes in adoption when either $w$ or $b$ vary are more subtle. An increase in $w$ results in a decrease in both $\theta_K$ and $\theta_F^*$, as shown in Figure 5. This indicates an increase in both the total number of adopters, as well as the fraction of adopters of the fixed fee. For parameter values of interest.[7] $\theta_K$ decreases more rapidly than $\theta_F^*$ with a marginal increase in $w$, and therefore

[7] More precisely, for any $c$ and $K$ such that there are at least some adopters of the usage-based contract, it can be shown that $d\theta_F^*/dw > -\frac{1}{2}$, while $d\theta_K/dw < -\frac{1}{2}$, which implies that $d\theta_F^*/dw > d\theta_K/dw$.

**Figure 2 How Increasing Transaction Costs $K$ and $c$ Affects Customer Adoption**

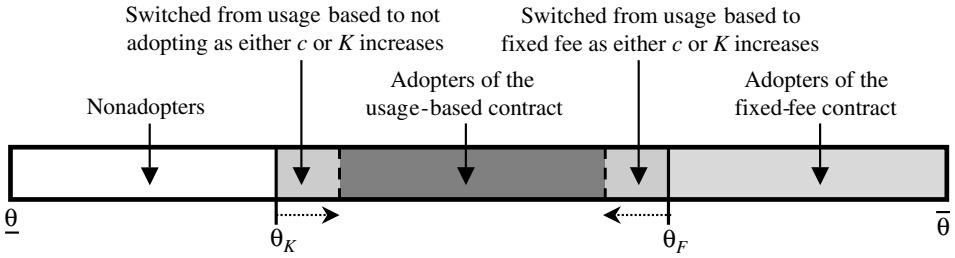



**Figure 3 Changes in Profits and Surplus as Transaction Costs Vary**

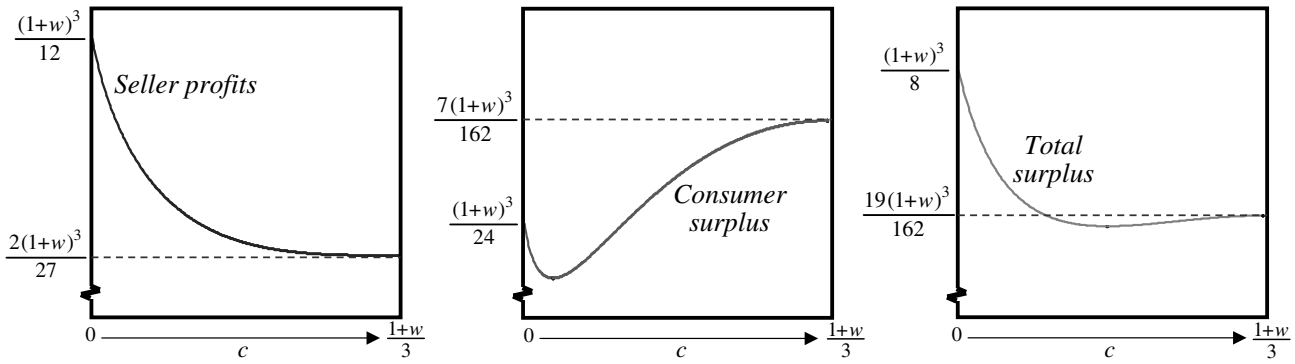



**Figure 4 The Drivers of Changes in Total Surplus: A Closer Look**

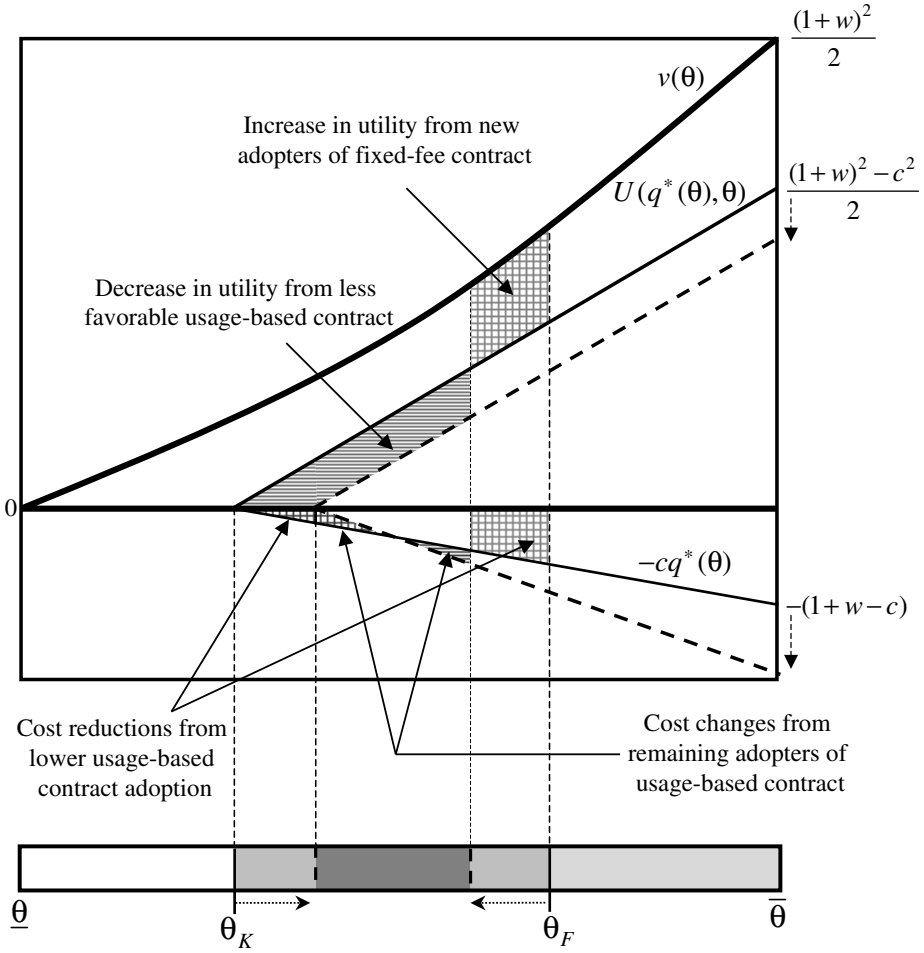
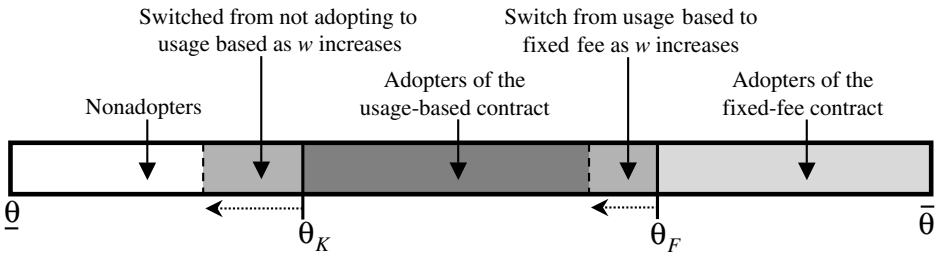


**Figure 5 The Impact of an Increase in Marginal Customer Value $w$ on Customer Adoption**




the number of adopters $(\theta_F^* - \theta_K)$ of the usage-based contract increases as well. Changes in $b$ alter the *shape* of the type distribution as well, and their effect on adoption is ascertained by examining the signs of the total derivatives $(d/db)[1 - F(\theta_F^*)]$ and $(d/db)[F(\theta_F^*) - F(\theta_K)]$, which measure the net changes on the fraction of customers adopting the fixed-fee and usage-based pricing schedules. The former is positive and the latter is negative, suggesting a shift away from fixed-fee pricing as $b$ decreases.[8]

Early-stage information or technology markets commonly feature a high concentration of occasional experimenters on the low end of the market, along with a small fraction of active early innovators who constitute a bulk of total usage. As the market matures, the distribution of customers over usage levels evens out. For instance, average monthly usage levels per customer in the online services market increased steadily over the first few years of the Internet boom—as of late 2001, average AOL usage had more than doubled to about 40 hours per month—and Jupiter Media-Metrix survey data from May 2001 on overall U.S. residential online usage indicates that the distribution of customers over usage levels has flattened out, especially below the mean. This is the kind of distributional change corresponding to a gradual decrease in $b$. The analysis above indicates that it is optimal for the provider to penetrate such a market initially with a pricing scheme that induces the adoption of a relatively low fixed fee. Over time, they should gradually increase this fixed fee, while inducing an increase in the adoption of usage-based pricing. Moreover, if the product becomes more valuable on average (from the addition of new features over time, for instance, which corresponds to an increase in $w$), this is also optimally responded to by increasing the fraction of customers who adopt usage-based pricing.

### 5.4. Minimum Fees and Quantity Discounting

Inspection of the usage-based contract in Table 2 indicates that the explicit usage-based pricing function of the seller takes the form

$$p(q) = \left(\frac{1 + bc + w}{1 + b}\right)q - \frac{q^2}{2(1+b)} \quad \text{for } q \geq q^*(\theta_K). \tag{22}$$

As discussed in §4.1, when $K > 0$, then $q^*(\theta_K) > 0$, and consequently, this is a nonlinear two-part tariff, which specifies a *minimum fee* of $\tau^*(\theta_K)$ for a usage level between 0 and a prespecified upper limit $q^*(\theta_K)$, and additional variable payments for usage above $q^*(\theta_K)$. Moreover, the function $p(q)$ is strictly concave, indicating an increasing level of volume discounts with usage. A useful measure of the percentage of discounting is the expression $(-p_{11}(q)/p_1(q))$, which measures the rate of decrease $p_{11}(q)$ of the variable price (or analogously, the concavity of the pricing function), normalized for variable price $p_1(q)$. From (22),

$$-\frac{p_{11}(q)}{p_1(q)} = \frac{1}{1 + bc + w - q}. \tag{23}$$

Therefore, as $w$ increases, the percentage discount offered should progressively decrease. Intuitively, the increase in marginal value increases the level of usage chosen by each customer; moreover, it is optimal for the firm to induce a higher fraction of the market to adopt its fixed-fee contract, as illustrated in Figure 4. The relative benefits of the quantity discount for the firm are consequently lower, which leads to a decrease in the discount. Similarly, an increase in $c$ results in a decrease in the percentage discount. While the direction of the result is similar, the intuition is slightly different—in this case, the firm does so because it optimally wants to induce lower usage for all adopters of the usage-based contract, as well as shifting a fraction of them to the fixed fee.

## 6. Summary and Conclusions

This paper has established that in the presence of any positive transaction costs, sellers of information goods should offer their customers a combination of usage-based pricing and unlimited-usage fixed-fee pricing. These conclusions contrast with well-known results from nonlinear pricing theory under assumptions similar to those made in this paper (see, for instance, Maskin and Riley 1984 or Wilson 1993), which suggest that optimal monopoly pricing structure is purely

[8] Similar results are obtained for the exponential distribution, which is also positively skewed, and this analysis is available in the extended appendix of the paper.

usage based and fully revealing. These models do not generally explicitly consider transaction costs. I show that the optimality of pure second-degree price discrimination is highly sensitive to the absence of these transaction costs—Proposition 2 has established that when there are no variable production costs, a purely usage-based pricing scheme is never optimal for any $C(q) > 0$. This is an important new conclusion for any seller developing pricing policy for their information goods.

Proposition 4 proves that the optimal usage-based contract is independent of the fixed fee, which reduces a complex constrained problem to a relatively simpler and more tractable one. The assumptions needed on customer preferences and heterogeneity for this result to work are fairly mild. Applying Proposition 4 is relatively straightforward, as illustrated in §5 and the paper's extended appendix. It is hoped that this result will enable further development of focused and rigorous models for specific information pricing problems.

These results generalize existing pricing guidelines for information goods. For instance, Varian (2000) proved that with two customer types and linear utility, a "buy only" pricing regime (which corresponds to offering only an unlimited usage fixed fee in our model) is strictly preferable to one that includes renting (usage-based pricing in our model) as long as the transaction costs of renting are positive (nonzero transaction costs $C(q)$ in our model). This result is intuitively appealing and highlights the importance of considering transaction costs when pricing information goods. By generalizing this intuition, this paper has established that while fixed fees are always profitable, the two kinds of pricing schemes can optimally coexist.

Moreover, all of the paper's results about the optimal design of pricing schedules continue to hold even if these transaction costs are borne by the *customer* rather than by the seller. For instance, when consuming under a usage-based pricing schedule, the customer may bear costs of monitoring and controlling their cumulative monthly usage. Even under this scenario, it is always optimal for the seller to offer both fixed-fee as well as usage-based pricing, and if the costs borne by customers are high enough, only the fixed fee. This provides one explanation for why Sprint recently switched from per-Mb usage-based pricing to fixed-fee pricing for their wireless Web service. Optimal usage levels $q^*(\theta)$ for this problem are analytically identical to those derived in this paper, and the expressions for optimal total prices can be obtained by simply adjusting the corresponding $\tau^*(\theta)$ expressions downward by the transaction costs borne by the customer at each usage level $q^*(\theta)$.

In the absence of transaction costs, Lemma 4 of the paper confirms that pure usage-based second-degree price discrimination is still optimal, even for zero-variable cost information goods. An example of a digital product category that uses this pricing strategy is back-end corporate software (such as database engines and application servers). Pricing is tied to server processor speed, and transaction costs are eliminated by coding the maximum allowable processor speed into the software delivered.

Some other key managerial insights from the paper are summarized below.

- As advances in electronic business transactions drive down the costs associated with administering usage-based contracts, sellers of information goods should adjust their pricing policy to increase the scope and adoption of their usage-based pricing schedules.
- If the administering of usage-based pricing involves a fixed per-customer transaction cost which is triggered by any positive usage level, the optimal usage-based pricing schedule should include a minimum fee for usage upto a prespecified level, and variable pricing beyond this.
- Typically, the variable pricing described above will feature volume discounts. Moreover, the extent of discounting should decrease as marginal value increases but increase as variable transaction costs reduce, or as the customer distribution becomes less skewed.
- In early-stage information markets characterized by a high concentration of low-usage customers and a small fraction of active early adopters, low fixed-fee penetration pricing is a good strategy. This is especially true if there are setup or periodic infrastructure costs associated with administering usage-based pricing. As the market matures and the distribution of customers across different usage levels evens out, sellers should increase their fixed fees and gradually expand their usage-based pricing options.

Propositions 2 and 4 also complement the basic rationale for bundling information goods, as prescribed by Bakos and Brynjolfsson (1999)—that a larger bundle of information goods increases the average customer valuation per unit good. Therefore, if the bundle is treated as the *potential set* of goods the consumer might use, marginal value from usage for each customer will increase with the size of the bundle. When the seller chooses the right combination of fixed-fee and usage-based contracts, this increase in per-unit value will lead to an increase the fraction of customers who choose a usage-based contract. The seller will also be able to extract more surplus from customers than it would have either under pure bundling or with a smaller bundle.

If one were to interpret $q$ in our model as *quality* instead of quantity, it is identical to one of vertical differentiation, with a continuum of possible product versions, a continuum of customer types, and costs of

versioning according to $C(q)$. In the context of information goods with multiple features, where quality is proportionate to the number of features, Proposition 2 indicates that if versioning is costly, it is always optimal to offer a high-priced version with all possible features that allows customers to self-customize (that is, choose the features that they want). This is consistent with Jones and Mendelson (1998) and Bhargava and Choudhury (2001). However, Proposition 4 indicates that it is often optimal to offer limited-feature versions as well, and that these will be adopted by a subset of customers as long as the cost of versioning is not too high.

The cost structure of information goods often leads to natural monopoly. However, competition is also a significant issue in pricing, and a focus of my ongoing work is competitive nonlinear pricing for information goods. Undifferentiated nonlinear price competition is not sustainable for information goods, because the equilibrium outcome is either marginal cost pricing, or minimum average cost pricing (Mandy 1992). The latter outcome suggests infinite-usage fixed-fee pricing, because the average cost per unit for information goods is always strictly decreasing in usage. A promising alternative is presented by Fishburn et al. (1997), who model a repeated game in which one player chooses only a fixed fee, and the other chooses only a linear usage-based price. Approaches that may yield more general results involve modeling horizontally differentiated information goods (as in the monopoly model of Weber 2001 and the duopoly model of Rochet and Stole 2002), using a model of monopolistic competition (as in Banker et al. 1998), or modeling the presence of an outside good (as in Jullien 2000). I hope to add to this literature in the near future.

An electronic companion to this paper is available at http://mansci.pubs.informs.org/ecompanion.html.

**Acknowledgments**
The author thanks Roy Radner for a number of helpful discussions, Rajiv Banker, an anonymous associate editor, three anonymous referees, Bing Jing, Ravi Mantena, Lars Stole, participants at the second International Industrial Organization Conference, and seminar participants at Carnegie-Mellon University, New York University, the University of Maryland, the University of Pennsylvania, and the University of Rochester for their feedback and comments on earlier drafts of this paper. All errors remain the author's.

## Appendix. Proofs

Proof of Lemma 1. Suppose that $q_1(\theta) < 0$ for some $\theta$. This implies that $q(\hat{\theta}) > q(\hat{\theta}+\varepsilon)$ for some $\hat{\theta} \in [\underline{\theta}, \bar{\theta}]$ and some $\varepsilon > 0$. Applying the condition [IC] at $\hat{\theta}$ and $(\hat{\theta}+\varepsilon)$ respectively yields

$$\text{[IC] at } \hat{\theta}\text{:}\quad U(q(\hat{\theta}),\hat{\theta}) - \tau(\hat{\theta}) \geq U(q(\hat{\theta}+\varepsilon),\hat{\theta}) - \tau(\hat{\theta}+\varepsilon). \tag{A1}$$

$$\text{[IC] at } (\hat{\theta}+\varepsilon)\text{:}\quad U(q(\hat{\theta}+\varepsilon),\hat{\theta}+\varepsilon) - \tau(\hat{\theta}+\varepsilon) \geq U(q(\hat{\theta}),\hat{\theta}+\varepsilon) - \tau(\hat{\theta}). \tag{A2}$$

Combining (A1) and (A2) and eliminating $[\tau(\hat{\theta}) - \tau(\hat{\theta}+\varepsilon)]$ yields

$$U(q(\hat{\theta}),\hat{\theta}+\varepsilon) - U(q(\hat{\theta}+\varepsilon),\hat{\theta}+\varepsilon) \geq U(q(\hat{\theta}),\hat{\theta}) - U(q(\hat{\theta}+\varepsilon),\hat{\theta}). \tag{A3}$$

Because $q(\hat{\theta}) > q(\hat{\theta}+\varepsilon)$, and $\hat{\theta}+\varepsilon > \hat{\theta}$, Equation (A3) implies that $U_2(q,\theta) \leq 0$ for some $\theta \in [\hat{\theta}, \hat{\theta}+\varepsilon]$—a contradiction. This proves that $q_1(\theta) \geq 0$ for all $\theta$. Now, applying first-order conditions for [IC] to hold for customer type $\theta$ yields

$$U_1(q(\theta),\theta)q_1(\theta) - \tau_1(\theta) = 0 \quad \text{for all } \theta, \tag{A4}$$

which ensures that $q_1(\theta) \geq 0 \Rightarrow \tau_1(\theta) \geq 0$, and proves part (a). Furthermore,

$$\frac{d}{d\theta}[U(q(\theta),\theta) - \tau(\theta)] = U_1(q(\theta),\theta)q_1(\theta) + U_2(q(\theta),\theta) - \tau_1(\theta). \tag{A5}$$

Combining (A4) and (A5) and using $U_2(q,\theta) > 0$ establishes part (b), and complete the proof. □

Proof of Lemma 2. Recall that $\psi(\theta) = v(\theta) - U(q(\theta),\theta) + \tau(\theta)$. Differentiating both sides with respect to $\theta$ and using (A4) yields

$$\psi_1(\theta) = v_1(\theta) - U_2(q(\theta),\theta). \tag{A6}$$

Because $v(\theta) = \lim_{q\to\infty} U(q,\theta)$, it follows that $v_1(\theta) = \lim_{q\to\infty} U_2(q,\theta)$. Because $U_{12}(q,\theta) > 0$, and $U(q,\theta)$ is monotonic in $q$, this implies that $v_1(\theta) > U_2(q,\theta)$ for all $q < \infty$, which in turn implies that $\psi_1(\theta) > 0$ so long as $q(\theta) < \infty$, and completes the proof. □

Proof of Lemma 3. Suppose that $\tau^*(\theta) - C(q^*(\theta))$ is strictly decreasing at some $\theta_L \in [\underline{\theta}, \bar{\theta}]$. We can therefore define a type $\theta_H > \theta_L$ as

$$\theta_H = \min\{\theta\colon \theta > \theta_L \text{ and } \tau^*(\theta) - C(q^*(\theta)) = \tau^*(\theta_L) - C(q^*(\theta_L))\}, \tag{A7}$$

with $\theta_H = \bar{\theta}$ if a type $\theta_H \in [\underline{\theta}, \bar{\theta}]$ according to (A7) does not exist. The continuity of $\tau^*(\theta)$, $q^*(\theta)$ and $C(q)$ ensures that

$$\tau^*(\theta) - C(q^*(\theta)) < \tau^*(\theta_L) - C(q^*(\theta_L)) \tag{A8}$$

for all $\theta \in (\theta_L, \theta_H)$. Now, define the contract $q(\theta), \tau(\theta)$ as follows:

$$q(\theta) = q^*(\theta),\ \tau(\theta) = \tau^*(\theta), \quad \text{for } \theta \notin [\theta_L, \theta_H]; \tag{A9}$$

$$q(\theta) = q^*(\beta(\theta)),\ \tau(\theta) = \tau^*(\beta(\theta)), \quad \text{for } \theta \in [\theta_L, \theta_H], \tag{A10}$$

where

$$\beta(\theta) = \theta_L \text{ if } U(q^*(\theta_L),\theta) - \tau^*(\theta_L) \geq U(q^*(\theta_H),\theta) - \tau^*(\theta_H); \tag{A11}$$

$$\beta(\theta) = \theta_H \text{ if } U(q^*(\theta_L),\theta) - \tau^*(\theta_L) < U(q^*(\theta_H),\theta) - \tau^*(\theta_H). \tag{A12}$$

Because $q^*(\theta), \tau^*(\theta)$ is incentive compatible, and $U_2(q, \theta) > 0$, it is easily shown that $q(\theta), \tau(\theta)$ is also incentive compatible. Moreover, because $f(\theta) > \theta$ for all $\theta$, (A8) implies that the seller's profits by offering $q(\theta), \tau(\theta)$ are strictly higher than those from $q^*(\theta), \tau^*(\theta)$, which contradicts the fact that $q^*(\theta), \tau^*(\theta)$ is optimal, and proves part (a).

An identical argument for each subinterval $[\theta_L, \theta_H]$ whose interior in which $\tau^*(\theta) - C(q^*(\theta))$ is strictly negative establishes part (b). The result follows. □

Proof of Proposition 3. This proposition uses the following lemma, a slight generalization of a well-known result (Maskin and Riley 1984), which is proved in the extended appendix.

Lemma 4. *If the fixed component of transaction costs is zero, and therefore $C(q) = c(q)$, then the unique optimal usage-based contract $q(\theta), \tau(\theta)$ for any interval $[\theta_L, \theta_H]$ satisfies the following conditions for all $\theta \in [\theta_L, \theta_H]$:*

$$U_1(q(\theta), \theta) = c_1(q(\theta)) + U_{12}(q(\theta), \theta)\frac{F(\theta_H) - F(\theta)}{f(\theta)}, \quad \text{(A13)}$$

$$\tau(\theta) = U(q(\theta), \theta) - \int_{\theta_L}^{\theta} U_2(q(x), x)\, dx. \quad \text{(A14)}$$

*Also, if $U_{122}(q, \theta) \le 0$, then $q_1(\theta) > 0$ for all $\theta$ such that $q(\theta) > 0$.*

The problem above is termed the *zero-fixed-cost problem for the interval* $[\theta_L, \theta_H]$. An important feature of this lemma is that the allocation $q(\theta)$ is independent of the lower support $\theta_L$. Therefore, for each $\theta \in [\theta_L, \theta_H]$, the value of $q(\theta)$ that is optimal for $[\theta_L, \theta_H]$ is the same as the value of $q(\theta)$ that would be optimal for the interval $[\underline{\theta}, \theta_H]$. Total payment $\tau(\theta)$ increases with $\theta_L$, of course.

Now, let $q^*(\theta), \tau^*(\theta)$ be the optimal usage-based contract for the entire interval, and for a given $K > 0$. Term this the solution to the *positive fixed-cost problem*. Define $\Theta \subset [\underline{\theta}, \bar{\theta}]$ as the set of types who use nonzero quantities under this contract:

$$\Theta = \{\theta\colon q^*(\theta) > 0\}. \quad \text{(A15)}$$

Clearly, $\tau^*(\theta) = 0$ for $\theta \notin \Theta$. From Lemma 1(a), we know that $q_1^*(\theta) \ge 0$, which implies that $\Theta$ is a continuous interval $[\theta_K, \bar{\theta}]$, where $\theta_K \ge \underline{\theta}$ (this is ignoring the trivial case where $\Theta$ is empty and the firm makes no profits). Also, from Lemma 3(b), we know that for each $\theta \in [\theta_K, \bar{\theta}]$,

$$\tau^*(\theta) \ge K + c(q^*(\theta)). \quad \text{(A16)}$$

Next, we establish that $q^*(\theta), \tau^*(\theta)$ is also the optimal contract for the zero-fixed-cost problem for the interval $[\theta_K, \bar{\theta}]$. If we assume the converse, the uniqueness result of Lemma 4 therefore implies the existence of an incentive-compatible contract $q(\theta), \tau(\theta)$ such that

$$\int_{\theta_K}^{\bar{\theta}} [\tau(\theta) - c(q(\theta))] f(\theta)\, d\theta > \int_{\theta_K}^{\bar{\theta}} [\tau^*(\theta) - c(q^*(\theta))] f(\theta)\, d\theta, \quad \text{(A17)}$$

which implies that

$$\int_{\theta_K}^{\bar{\theta}} [\tau(\theta) - c(q(\theta)) - K] f(\theta)\, d\theta > \int_{\theta_K}^{\bar{\theta}} [\tau^*(\theta) - c(q^*(\theta)) - K] f(\theta)\, d\theta. \quad \text{(A18)}$$

Because $\tau^*(\theta) = q^*(\theta) = 0$ for $\theta < \theta_K$, (A18) implies that $q(\theta), \tau(\theta)$ is a more profitable solution to the positive fixed-cost problem than $q^*(\theta), \tau^*(\theta)$—a contradiction. Therefore, $q^*(\theta)$ always satisfies

$$U_1(q^*(\theta), \theta) = c_1(q^*(\theta)) + U_{12}(q^*(\theta), \theta)\frac{1 - F(\theta)}{f(\theta)}, \quad \text{(A19)}$$

which specifies a unique $q^*(\theta)$ for each $\theta$ that is also independent of the value of $\theta_K$. The result follows by choosing the $\theta_K$ that maximizes the firm's profits $\int_{\theta_K}^{\bar{\theta}} [\tau^*(\theta) - c(q^*(\theta)) - K] f(\theta)\, d\theta$, while taking into account the dependence of $\tau^*(\theta)$ on $\theta_K$. □

Proof of Proposition 4. Given an interval $[\theta_K, \bar{\theta}]$ and a type $\theta_F \in (\theta_K, \bar{\theta})$, suppose that with fixed costs of zero, the firm wants to design the optimal usage-based contract $\hat{q}(\theta), \hat{\tau}(\theta)$ and the optimal fixed fee $\hat{T}$, subject to the constraint that all customer types $\theta \ge \theta_F$ adopt the fixed fee, and all customer types in the subinterval $[\theta_K, \theta_F]$ adopt the usage-based contract. We term this a *constrained zero-fixed-cost problem for the interval* $[\theta_K, \bar{\theta}]$ *with a constraint on* $\theta_F$. Proposition 2 has shown that constraining customers of type $\theta_F$ to be indifferent between any fixed-fee contract $T$ and the usage-based contract will ensure that all types $\theta \in [\theta_K, \theta_F)$ will choose the usage-based contract, and all types $\theta \ge \theta_F$ will choose the fixed fee. Therefore, from the proof of Lemma 4, we know that this problem of finding $\hat{q}(\theta), \hat{\tau}(\theta)$ and $\hat{T}$ can be formulated as

$$\max_{q(\cdot),\, T} \left( \int_{\theta_K}^{\theta_F} \left[ U(q(\theta), \theta) - c(q(\theta)) - U_2(q(\theta), \theta) \cdot \frac{F(\theta_F) - F(\theta)}{f(\theta)} \right] f(\theta)\, d\theta \right) + T(1 - F(\theta_F)), \quad \text{(A20)}$$

subject to the constraint

$$\left( \int_{\theta_K}^{\theta_F} U_2(q(\theta), \theta)\, d\theta \right) - [v(\theta_F) - T] = 0. \quad \text{(A21)}$$

Denote the Lagrangian for this problem as

$$L(q(\cdot), T, \lambda) = \int_{\theta_K}^{\theta_F} \left( U(q(\theta), \theta) - c(q(\theta)) - U_2(q(\theta), \theta)\frac{F(\theta_F) - F(\theta)}{f(\theta)} \right) f(\theta)\, d\theta + T[1 - F(\theta_F)] + \lambda \left[ \left( \int_{\theta_K}^{\theta_F} U_2(q(\theta), \theta)\, d\theta \right) - v(\theta_F) + T \right]. \quad \text{(A22)}$$

The first-order necessary conditions for any local maximizer to this constrained problem are

$$\left[\frac{\partial L}{\partial q} = 0\right]\colon \left( U_1(q(\theta), \theta) - c_1(q(\theta)) - U_{12}(q(\theta), \theta) \cdot \left[ \frac{F(\theta_F) - F(\theta) - \lambda}{f(\theta)} \right] \right) f(\theta) = 0 \quad \forall \theta \in [\theta_K, \theta_F], \quad \text{(A23)}$$

$$\left[\frac{\partial L}{\partial T} = 0\right]\colon [1 - F(\theta_F)] + \lambda = 0, \quad \text{(A24)}$$

and

$$\left[\frac{\partial L}{\partial \lambda} = 0\right]\colon \int_{\theta_K}^{\theta_F} U_2(q(\theta), \theta)\, d\theta = v(\theta_F) - T. \quad \text{(A25)}$$

Rearranging (A24) yields the value of the Lagrangian multiplier

$$\lambda = -[1 - F(\theta_F)]. \quad \text{(A26)}$$

Substituting (A26) into (A23) yields

$$\left(U_1(q(\theta), \theta) - c - U_{12}(q(\theta), \theta) \cdot \left[\frac{F(\theta_F) - F(\theta) - (-[1 - F(\theta_F)])}{f(\theta)}\right]\right) f(\theta) = 0, \quad \text{(A27)}$$

which simplifies to

$$U_1(q(\theta), \theta) = c_1(q(\theta)) + U_{12}(q(\theta), \theta)\left[\frac{1 - F(\theta)}{f(\theta)}\right] \quad \forall\, \theta \in [\theta_K, \theta_F]. \quad \text{(A28)}$$

Equation (A28) implies that for any constrained zero-fixed cost problem, the usage-based quantity $q(\theta)$ optimally chosen by each type $\theta$ in $[\theta_K, \theta_F]$ is independent of $\theta_K$, independent of $\theta_F$, and is also independent of the value of the fixed fee $T$.

Now, suppose $q^*(\theta), \tau^*(\theta)$, and $T^*$ are a solution to the seller's original problem of choosing the optimal combination of usage-based pricing and fixed-fee, for a specific $K > 0$, and for the entire interval $[\underline{\theta}, \bar{\theta}]$. Let $[\theta_K, \bar{\theta}]$ be the set of types for which $q^*(\theta) > 0$. If all types $[\theta_K, \bar{\theta}]$ adopt $T^*$, the main result is trivially true. If not, Proposition 2 ensures that there is a customer type $\theta_F^* \in (\theta_K, \bar{\theta})$ who is indifferent between adopting $q^*(\theta), \tau^*(\theta)$ and adopting $T^*$, and Proposition 1 ensures that all types in $[\theta_F^*, \bar{\theta}]$ adopt $T^*$.

It now follows that the contracts $q^*(\theta), \tau^*(\theta)$, and $T^*$ must also be a solution to the problem of choosing the optimal combination of usage-based pricing and fixed-fee, for a *constrained zero-cost-problem* for the interval $[\theta_K, \bar{\theta}]$, and with constraint on the type $\theta_F^*$. Suppose it is not such a solution. This implies the existence of $q(\theta), \tau(\theta)$, and $T$ such that types $[\theta_K, \theta_F^*)$ adopt $q(\theta), \tau(\theta)$, types $[\theta_F^*, \bar{\theta}]$ adopt $T$, and

$$\int_{\theta_K}^{\theta_F^*} [\tau(\theta) - c(q(\theta))] f(\theta)\, d\theta + T[1 - F(\theta_F^*)] > \int_{\theta_L}^{\theta_F^*} [\tau^*(\theta) - c(q^*(\theta))] f(\theta)\, d\theta + T^*[1 - F(\theta_F^*)]. \quad \text{(A29)}$$

Subtracting $K[F(\theta_F^*) - F(\theta_K)]$ from both sides of (A29) implies that $q(\theta), \tau(\theta)$ and $T$ satisfy

$$\int_{\theta_K}^{\theta_F^*} [\tau(\theta) - c(q(\theta)) - K] f(\theta)\, d\theta + T[1 - F(\theta_F^*)] > \int_{\theta_L}^{\theta_F^*} [\tau^*(\theta) - c(q^*(\theta)) - K] f(\theta)\, d\theta + T^*[1 - F(\theta_F^*)], \quad \text{(A30)}$$

which contradicts the fact that $q^*(\theta), \tau^*(\theta)$, and $T^*$ are a solution to the seller's original problem. Because (A28) is necessary for*any* solution to the constrained zero-cost-problem, it follows that

$$U_1(q^*(\theta), \theta) = c_1(q^*(\theta)) + U_{12}(q^*(\theta), \theta)\left[\frac{1 - F(\theta)}{f(\theta)}\right] \quad \forall\, \theta \in [\theta_K, \bar{\theta}]. \quad \text{(A31)}$$

Note that the expressions in (A31) are independent of $\theta_K$ and are also independent of $T^*$. This proves part (a). The value of the fixed-fee at which type $\theta_F$ is indifferent is $v(\theta_F) - U(q^*(\theta_F), \theta_F) + \tau^*(\theta_F)$. Part (b) follows from the fact that the firm will choose the profit maximizing value of $\theta_F^*$, and that the corresponding optimal usage-based contract $(q^*(\theta), \tau^*(\theta))$ is independent of the choice of $\theta_F^*$. Part (c) simply computes the fixed-fee $T^*$ at this optimal value of $\theta_F^*$. □

## References

Armstrong, M. 1996. Multidimensional price screening. *Econometrica* **64** 51–75.

Bakos, Y., E. Brynjolfsson. 1999. Bundling information goods: Pricing, profits and efficiency. *Management Sci.* **45**(12) 1613–1630.

Banker, R., I. Khosla, K. Sinha. 1998. Quality and competition. *Management Sci.* **44** 1179–1192.

Bhargava, H., V. Choudhary. 2001. Information goods and vertical differentiation. *J. Management Inform. Systems* **18** 89–106.

Cheng, H., G. Koehler. 1999. Optimal pricing policies of web-enabled application services. University of Florida, Gainesville, FL.

Cheng, H., M. Freimer, W. Richmond, U. Sumita. 1998. Optimal repair of degrading computers under asymmetric information and incentive incompatibility. *Comput. Oper. Res.* **25** 719–728.

Chuang, J., M. Sirbu. 1999. Optimal bundling strategy for digital information goods: Network delivery of articles and subscriptions. *Inform. Econom. Policy* **11** 147–176.

Dewan, S., H. Mendelson. 1990. User delay costs and internal pricing for a service facility. *Management Sci.* **36** 1502–1517.

Fishburn, P., A. Odlyzko, R. Siders. 1997. Fixed fee versus unit pricing for information goods: Competition, equilibria, and price wars. *First Monday* **2**(7), http://www.firstmonday.dk.

Fudenberg, D., J. Tirole. 1991. *Game Theory.* MIT Press, Cambridge, MA.

Gupta, A., D. Stahl, A. Whinston. 1997. A stochastic equilibrium model of Internet pricing. *J. Econom. Dynamics Control* **21** 697–722.

Jones, R., H. Mendelson. 1998. Product and price competition for information goods. Working paper, University of Rochester, Rochester, NY.

Jullien, B. 2000. Participation constraints in adverse selection problems. *J. Econom. Theory* **93** 1–47.

Mandy, D. 1992. Nonuniform Bertrand competition. *Econometrica* **60** 1293–1330.

Maskin, E., J. Riley. 1984. Monopoly with incomplete information. *RAND J. Econom.* **15** 171–196.

Mendelson, H. 1985. Pricing computer services: Queuing effects. *Comm. ACM* **28** 312–321.

Mendelson, H., S. Whang. 1990. Optimal incentive-compatible priority pricing for the $M/M/1$ queue. *Oper. Res.* **38** 870–883.

Rochet, J.-C., L. A. Stole. 2002. Nonlinear pricing and random participation. *Rev. Econom. Stud.* **69** 277–311.

Susarla, A., A. Barua, A. B. Whinston. 2001. Understanding the "service" component of application service provision: An empirical analysis of satisfaction with ASP services. *MIS Quart.* **27** 91–123.

Varian, H. 2000. Buying, selling and renting information goods. *J. Indust. Econom.* **XLVIII** 473–488.

Weber, T. 2001. Mixed differentiation of information goods under incomplete information. L. Applegate, R. D. Galliers, J. I. DeGross, eds. *Proc. 22nd Internat. Conf. Inform. Systems*, 81–94.

Westland, C. 1992. Congestion and network externalities in the short run pricing of information systems services. *Management Sci.* **38** 992–1099.

Wilson, R. 1993. *Nonlinear Pricing*. Oxford University Press, New York.