Non-linear pricing of information goods

Arun Sundararajan[1]

Department of Information, Operations and Management Sciences

Leonard N. Stern School of Business, New York University

44 West 4th Street, New York, NY 10012

asundara@stern.nyu.edu

January 2002

[1] I thank Roy Radner for a number of helpful discussions and comments, and Bing Jing and Ravi Mantena for helpful comments on an earlier draft of this paper. All errors remain mine.

Center for Digital Economy Research
Stern School of Business
Working Paper IS-02-01

Non-linear pricing of information goods

January 2002

This paper analyzes optimal pricing for information goods under incomplete information, when both unlimited-usage (fixed-fee) pricing and usage-based pricing are feasible. For a general set of customer characteristics, it is shown that in the presence of contract administration costs, offering fixed-fee pricing in addition to a non-linear usage-based pricing scheme is always profit-improving, and there may be markets in which a pure fixed-fee is optimal. Moreover, it is proved that the optimal usage-based pricing schedule is independent of the value of the fixed-fee. These results imply that the optimal pricing strategy is never fully revealing. A procedure for determining the optimal combination of fixed-fee and non-linear usage-based contracts is presented.

Applying these general results to specific business contexts suggests a number of operational guidelines for designing pricing schedules, and managerial insights for setting pricing policy. For instance, in nascent information markets, firms are most likely to 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. The effects of changes in product value and administration costs on the adoption levels of different pricing schemes, optimal quantity discounts, firm profitability and total welfare are analyzed. 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.

Center for Digital Economy Research
Stern School of Business
Working Paper IS-02-01

1 Introduction

Non-linear usage-based pricing is a popular price-discrimination technique, and has been analyzed extensively in the context of the electricity and long-distance telephone markets (Wilson, 1993). Many producers of information goods have also adopted this form of pricing. For instance, most corporate software manufacturers tie the prices of their products to the total processing speed of the servers on which the software is licensed to run, which indirectly bases the price paid by the customer on the total expected usage of the software. The emergence of application service providers has enabled a variety of more direct usage-based pricing schemes for software products[1]. All the major cellular telephone service providers have extensive menus of usage-based pricing plans. Many research firms price their services on a per-report basis.

In contrast, there are numerous examples of fixed-fee pricing for information goods, under which customers pay a fixed price that is independent of usage. Most Internet service providers charge their residential customers a flat monthly subscription fee. Jupiter Media-Metrix charges its clients $22,000 per year, independent of research usage. The Wall Street Journal offers unlimited access to its online version for a fixed annual fee. NetFlix gives its consumers unlimited DVD rentals for a flat fee of $20 per month. These fixed-fees are feasible because the firm selling the product incurs low or zero direct variable costs from increases in product usage.

Additionally, a number of companies have started using a combination of fixed-fee and usage-based pricing. For instance, IBM recently introduced options for both usage-based and flat-fee pricing on its zSeries software. Under this new pricing model, a customer can opt to pay a fixed fee, or to use a reporting tool[2] from IBM which tracks software usage on a monthly basis, and under which the customer may adjust their desired maximum usage periodically. Hewlett-Packard has announced that it will add pay-per-use pricing to its fixed-fee schedules for its Superdome and NetServer family of servers, charging customers based on the monthly average of the daily maximum active processors. In addition to its regular per-minute pricing scheme, Sprint offers a fixed-fee residential long-distance plan. Other information goods that feature both fixed-fee and transaction-based pricing include the OCLC information services used by most major libraries and Internet bandwidth, where a mixture of usage-independent and per-Mbps pricing structures is common.

These pricing policies which include fixed-fee options conflict with 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, and will be fully revealing. This disconnect between pricing theory and practice, though puzzling, could be due to one or more of the following reasons. These theoretical results may not apply to markets for information goods, where variable costs from usage are zero. These models may ignore transaction or administration costs associated with usage-based pricing, which make fixed-fees preferable in practice. Perhaps there are specific customer characteristics which favor fixed-fee pricing, and others which favor usage-based pricing. It is also possible that many industry pricing schemes are simply flawed. These are currently unresolved research questions, which this paper addresses.

Specifically, this paper models pricing decisions by a monopolist in a market for information goods, with customer heterogeneity and asymmetric information. Under fairly general assumptions about customer utility functions and preference distributions, the optimal combination of the unlimited usage fixed-fee and the usage-based nonlinear pricing function is derived. We establish that any positive cost of monitoring usage causes the monopolist to offer customers a fixed-fee option. It is shown that the choice of the usage-based pricing function is independent of the fixed-fee. In the process, we generalize and add to some recent results about renting versus buying information goods (Varian, 2000). The general formulation is applied to two examples, and the sensitivity of a number of variables (the level of the fixed fee, the shape of the usage-based pricing function, the optimal quantity discount rate, the proportion of customers adopting each pricing scheme, the monopolist's profits, consumer surplus and total welfare) to the model's parameters (usage-monitoring costs, marginal value from usage and shape of the customer distribution) is analyzed.

Researchers have studied the optimal pricing in information systems for many years, most often focusing on congestion pricing. This body of work includes a queuing model of ASP pricing by Cheng and Koehler (1999), a model of software renting under monopolistic competition by Choudhary, Tomek and Chaturvedi (1998), a general analysis of pricing service facilities with nonlinear delay costs, by Dewan and Mendelson (1990), a simulation model of usage-based pricing in a network by Gupta et. al. (2001), which is based on the theoretical framework of Gupta et al. (1995), the seminal paper by Mendelson (1985) on pricing computer services by internalizing delay externalities, which was followed by a model of variable queue pricing under asymmetric information by Mendelson and Whang (1990), and the general model of optimal IS pricing with network externalities by Westland (1992). Space constraints preclude a more detailed survey or analysis – these papers focus specifically either on contrasting usage-based pricing with alternate schemes, or on IS pricing under asymmetric information – which makes their models most relevant to this paper. Recent work-in-progress that is related includes Afeche and Mendelson's (2001) study of uniform versus auctioned pricing, and Lin and Whinston's (2001) work on incorporating risk preferences and variable usage into software contracts.

However, this paper represents the first formal work directly analyzing fixed-fee and non-linear pricing of information goods within a general framework. Two related and active areas of research into pricing and market segmentation – bundling and vertical differentiation (versioning) – have indirectly shed some light on this issue. For instance, Bakos and Brynjolfsson (1999) show that for a general set of valuation distributions, increasing the size of a bundle statistically increases the average valuation placed on a randomly chosen item, and reduces the variance relative to its mean, which allows a seller to charge a higher average fixed fee if consumer valuations are unbounded. Other results suggest that mixed bundling may be optimal when customer utility is bounded and linear in consumption (Chuang and Sirbu, 1998) and that customer heterogeneity favors bundling in specific competitive markets (Kephart and Fay, 2000). Typically, pure bundling is considered synonymous with fixed-fee pricing, and mixed bundling is equated to a mixture of a fixed-fee and usage-based pricing. However, this interpretation may not always be complete. Even if larger bundles are more valuable to customers on average, profit improvements may be possible from a mixture of fixed-fee and usage-based pricing for the larger bundle. This is likely to depend on the extent to which bundling reduces the information asymmetry about product valuation between the producer and consumers, and the fraction of the bundle consumed by the average customer. This paper provides results that can help in addressing these issues. We discuss the implications of our results for bundling models in Section 5.

The stream of literature on vertical differentiation for information goods (for instance, Bhargava and Choudhary, 2001, Dewan, Jing and Seidmann, 2000, Jones and Mendelson, 1998, Varian and Shapiro, 1998, and Weber, 2001, among others) studies price discrimination when customers have heterogeneous preferences for quality. In general, the results suggest that multiple versions of information goods can increase firm profits, though a result of Bhargava and Choudhary shows that under certain cost structures, a monopolist maximizes profits by offering a single version. Intrinsically, quality and quantity discrimination have similar objectives and use similar models, and the implications of the results of this paper for versioning information goods are also discussed in Section 5.

The rest of the paper is organized as follows. Section 2 describes the model. Section 3 provides the general results of the paper. These results are applied to two examples in Section 4. The final section provides pricing guidelines, discusses pricing policy implications and outlines current research.

2 Model

The model is described in three parts – the characteristics of the firm and customers, the feasible pricing schemes, and the sequence of interactions between the firm and their customers.

2.1 Firm and customers

A monopoly firm sells a homogeneous digital product, which may be used by customers in varying quantities. The product already exists (possibly as a result of expending a fixed cost). The variable cost to the firm of creating a copy of, or providing access to the product is zero. In addition, the variable cost to the firm from usage of the product by a customer is zero (apart from usage-based contract administration costs, discussed in section 2.3).

Customers are heterogeneous, indexed by their type $\theta \in [\underline{\theta}, \overline{\theta}]$. The preferences of a customer of type θ are represented by the linearly separable function

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

where q is the quantity of the product used by the customer, and p is the total price paid by the customer. Given a pricing schedule, each customer chooses a usage level q that maximizes w. The function U is referred to as the customer's utility function. Numbered subscripts to $U(q,\theta)$ denote partial derivatives with respect to the respective 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 shall be used for partial derivatives of functions throughout the paper, except when functions have only one argument, in which case the usual $g'(.)$ notation is used.

The utility function $U(q,\theta)$ has the following five properties:

1. Finite maximum usage: For each θ, $U_1(q,\theta) = 0$ at exactly one finite value of q, denoted $\alpha(\theta)$; $U_1(q,\theta) > 0$ for $q < \alpha(\theta)$, and $U_1(q,\theta) < 0$ for $q > \alpha(\theta)$.

2. Higher types get higher utility: $U_2(q,\theta) > 0$ for all q, θ.

3. Spence-Mirrlees single-crossing: For each θ, $U_{12}(q,\theta) > 0$ for all $q \leq \alpha(\theta)$.

4. Strict concavity in usage: $U_{11}(q,\theta) < 0$

5. Non-increasing absolute risk-aversion: $\dfrac{\partial}{\partial\theta}\left(\dfrac{-U_{11}(q,\theta)}{U_1(q,\theta)}\right) \leq 0$

Property 1 captures the fact that customers use a finite quantity of any information good, even if the marginal price of additional usage is zero. This is because value from usage is typically bounded by a constraint on some related resource – attention or computing power being two common examples – and the implicit presence of a substitute use for this resource. Analogously, sometimes the increased usage of an information good may necessitate the purchase of additional costly complementary assets like hardware. The examples in section 4 discuss these interpretations further.

Property 2 specifies that higher-types always get more value than lower types from the usage of a specific quantity. Property 3 is the standard Spence-Mirrlees single-crossing condition, required only in the region where U is increasing in quantity. The condition implies that in this region, higher types get a higher increase in value than lower types, from the same increase in quantity. Property 4 – the concavity of U in usage – is standard. Property 5, which is also relatively common, implies that the absolute risk aversion – a measure of the curvature of U – is either constant or is decreasing in type.

The utility derived from maximal usage by type θ is represented using the function

$$v(\theta) = \max_q U(q,\theta) = U(\alpha(\theta),\theta), \tag{2}$$

where $\alpha(\theta) = \arg\max_q U(q,\theta)$. Properties 1 and 2 together imply that if $\theta > \hat{\theta}$, then $\alpha(\theta) > \alpha(\hat{\theta})$, and that consequently, $v(\theta) > v(\hat{\theta})$. An example of the utility function, and the corresponding values of $v(\theta)$ and $\alpha(\theta)$, is depicted in Figure 1.

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 $\theta \in [\underline{\theta}, \overline{\theta}]$. The reciprocal of the hazard rate of the distribution of types, $\frac{1-F(\theta)}{f(\theta)}$, is assumed to be non-increasing in θ for all θ in $[\underline{\theta}, \overline{\theta}]$. This is a standard assumption in adverse selection problems, and is satisfied by most commonly used probability distributions. It also implies that for any fixed value of $\hat{\theta}$, $\frac{F(\hat{\theta})-F(\theta)}{f(\theta)}$ is non-increasing[3] in θ for all θ in $[\underline{\theta}, \hat{\theta}]$.

2.2 Pricing

The firm can offer one or both of two kinds of pricing (subsequently referred to as 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 administration costs associated with a fixed-fee contract – the customer simply pays the firm the price T, and usage is not monitored.

Usage-based: A usage-based contract assigns a specific price to each level of usage q. Since the firm cannot explicitly distinguish between customer types prior to contracting, 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 $t \in [\underline{\theta}, \overline{\theta}]$. This menu must satisfy a slightly modified version of the standard incentive-compatibility (IC) and individual rationality (IR) constraints:

(IC) (1) For each θ, $U(q(\theta), \theta) - \tau(\theta) \geq U(q(\hat{\theta}), \theta) - \tau(\hat{\theta})$, for all $\hat{\theta} \neq \theta$, $q(\hat{\theta}) \leq \alpha(\theta)$

(2) For each θ, $U(q(\theta), \theta) - \tau(\theta) \geq U(\alpha(\theta), \theta) - \tau(\hat{\theta})$, for all $\hat{\theta} \neq \theta$, $q(\hat{\theta}) > \alpha(\theta)$

(IR) For all θ, $U(q(\theta), \theta) - \tau(\theta) \geq 0$.

The (IC) constraints are separated into two sets because $U_1(q, \theta)$ is negative for $q > \alpha(\theta)$. Consequently, a customer of type θ who evaluates a quantity-price pair $(q(\hat{\theta}), \tau(\hat{\theta}))$for which $q(\hat{\theta}) > \alpha(\theta)$ will view its value as $U(\alpha(\theta), \theta) - \tau(\hat{\theta})$, rather than as (the strictly lower value) $U(q(\hat{\theta}), \theta) - \tau(\hat{\theta})$. This is analogous to an assumption of free disposal. If $q(\theta) > q(\hat{\theta})$ implies that $\tau(\theta) > \tau(\hat{\theta})$, then the first set of constraints imply the second set.

When the menu of quantity-price pairs satisfies (IC) and (IR), every customer of type θ will choose the pair $q(\theta), \tau(\theta)$.

The firm bears contract administration costs at the rate of c per unit usage, for each customer who chooses to adopt a usage-based contract. Pricing and product managers have indicated to the author that a reluctance to offer usage-based contracts may often stem from the associated administration costs – the costs that arise from having to monitor, record and analyze customer usage (sometimes using a remote software agent), implement non-uniform billing and collection on a periodic basis, the associated customer support costs from errors, billing disputes, clarifications and corrections, the additional accounting costs, and so on. Clearly, in practice, these costs will be non-decreasing in the level of usage by the customer; for simplicity, it is assumed in the model that they are linear in q.

2.3 Interaction between the firm and customer

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

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

2. Each customer either chooses a contract, or chooses not to purchase. If a customer chooses the fixed-fee contract, a fixed payment of T is made to the firm. Since the usage-based contract $(q(.), \tau(.))$ has been designed to satisfy (IC) and (IR), if a customer of type θ chooses the usage-based contract, then this customer uses a quantity $q(\theta)$, and makes a payment of $\tau(\theta)$ to the firm..

It is possible that when the customer chooses a quantity-price pair $(q(\theta), \tau(\theta))$, a copy of the digital product designed to be used upto a maximum quantity of $q(\theta)$ is delivered to the customer. This is commonly observed in the software industry. In equilibrium, since the usage-based contract satisfies (IC) and (IR), the model is consistent with this scenario as well.

The problem for a customer of type θ is to choose between paying a fixed fee T for a usage level $\alpha(\theta)$, 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.

The notation introduced in this section, along with some additional notation used in Section 3, is summarized in Table 1.

3 Optimal Contracts

Some (standard) terminology applied to the usage-based contracts subsequently:

Incentive-compatible: A usage-based contract $(q(.), \tau(.))$ is said to be incentive-compatible if it satisfies (IC) and (IR)

Fully-revealing: An incentive-compatible usage-based contract $(q(.), \tau(.))$ is said to be fully-revealing if $q(\theta)$ is strictly monotonic. Under a fully-revealing contract, the type of a customer is revealed by his choice of q.

Optimal: An incentive-compatible usage-based contract is said to be optimal for a sub-interval $[\theta_1, \theta_2]$ if it yields firm profits that are at least as high as any other incentive-compatible usage-based contract designed exclusively for customers in the sub-interval $[\theta_1, \theta_2]$. When no sub-interval is mentioned, optimality applies to the entire interval $[\underline{\theta}, \overline{\theta}]$

3.1 Optimal usage-based contracts for a subset of customers

The results in this sub-section demonstrate that individual customer surplus and usage increase as the interval of customers for which the firm designs a usage-based contract is narrowed. We also describe a lemma used in all subsequent proofs of the paper's propositions.

Suppose the firm wishes to maximize its ex-ante expected profits from customers in the sub-interval $[\underline{\theta}, \theta_U]$, where $\theta_U \leq \overline{\theta}$, by offering them a usage-based contract. Let the function pairs $(q(\theta, \theta_U), \tau(\theta, \theta_U))$ and $(q^*(\theta, \theta_U), \tau^*(\theta, \theta_U))$ denote any incentive-compatible usage-based contract, and the optimal incentive-compatible usage-based contract respectively, for the sub-interval $[\underline{\theta}, \theta_U]$. When the contract is for the entire interval $[\underline{\theta}, \overline{\theta}]$, the second argument in q and τ is dropped.

The firm's maximization problem is therefore:

$$\max_{q(.,\theta_U), \tau(.,\theta_U)} \int_{\underline{\theta}}^{\theta_U} [\tau(\theta, \theta_U) - cq(\theta, \theta_U)] f(\theta) d\theta,$$

subject to the set of constraints (IC) and (IR). Proposition 1 characterizes how usage and individual surplus vary with the interval $[\underline{\theta}, \theta_U]$.

Proposition 1 *As θ_U decreases:*

(a) the usage level $q^*(\theta, \theta_U)$ of any customer of type $\theta < \theta_U$ strictly increases, and

(b) the surplus $U(q^*(\theta, \theta_U), \theta) - \tau^*(\theta, \theta_U)$ of any customer of type $\theta < \theta_U$ strictly increases.

Proofs not in the main text are provided in Appendix A. The proof of Proposition 1 uses the following lemma:

Lemma 1 *The firm's problem when it wishes to maximize profits from customers in the sub-interval $[\underline{\theta}, \theta_U]$ by offering them a usage-based contract can be reduced to:*

$$\max_{q(.,\theta_U)} \int_{\underline{\theta}}^{\theta_U} [U(q(\theta,\theta_U),\theta) - cq(\theta,\theta_U) - U_2(q(\theta,\theta_U),\theta)\frac{F(\theta_U) - F(\theta)}{f(\theta)}]f(\theta)d\theta \tag{3}$$

subject to

$$q_1(\theta, \theta_U) \geq 0 \; \forall \theta, \tag{4}$$

and the optimal contract $(q^*(.,\theta_U), \tau^*(.,\theta_U))$ for the sub-interval $[\underline{\theta}, \theta_U]$ satisfies:

$$\begin{aligned} U_1(q^*(\theta,\theta_U),\theta) &= c + U_{12}(q^*(\theta,\theta_U),\theta)\frac{F(\theta_U) - F(\theta)}{f(\theta)} \; \forall \theta \leq \theta_U; \\ \tau^*(\theta,\theta_U) &= U(q^*(\theta,\theta_U),\theta) - \int_{\underline{\theta}}^{\theta} U_2(q^*(x,\theta_U),x)dx \; \forall \theta \leq \theta_U. \end{aligned} \tag{5}$$

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

Mathematically, with a little work, Lemma 1 can be deduced from Proposition 4 of Maskin and Riley (1984), though the specification of customer preferences and marginal costs is different in their model. For completeness, a (slightly different) proof of this lemma is provided in Appendix A.

The quantity $\int_{\underline{\theta}}^{\theta} U_2(q^*(x,\theta_U),x)dx$ is often referred to as the *informational rent* for type θ. By Lemma 1, it is exactly equal to the surplus that type θ gets, and can be interpreted as the rent that the monopolist needs to 'pay' customers of type θ in exchange for them revealing their true type. One might expect that lowering of the number of (IC) constraints will enable the firm to optimally ensure incentive compatibility at a lower level of informational rent. However, as the value of $q^*(\theta, \theta_U)$ increases, the surplus that a type higher than θ can get by choosing the quantity-price pair $(q^*(\theta, \theta_U), \tau^*(\theta, \theta_U))$ also increases. As a result, as the firm increases the usage level $q^*(\theta, \theta_U)$ designed for type θ in order to increase revenues, it also needs to increase the surplus offered to each of the types higher than θ.

Since $(q^*(\theta, \theta_U), \tau^*(\theta, \theta_U))$ is the optimal usage-based contract designed specifically for the sub-interval $[\underline{\theta}, \theta_U]$, the firm's profits *from this sub-interval* $[\underline{\theta}, \theta_U]$ are clearly at least as high under $(q^*(\theta, \theta_U), \tau^*(\theta, \theta_U))$ than they would be under $(q^*(\theta), \tau^*(\theta))$, the contract which is designed to optimize profits from the entire interval.

3.2 The impact of a fixed-fee on customer choice

The main result of this sub-section is to establish that when a fixed-fee contract is offered along with *any* incentive-compatible usage-based contract, then the customers typically bifurcate into two sub-intervals, with lower types adopting the usage-based contract, and higher types adopting the fixed-fee contract. The result provides a foundation for establishing the existence of a profit-improving fixed fee (which is proved in Section 3.3), and for then deriving the optimal mix of fixed-fee and usage-based contracts, which is addressed in Section 3.4.

Suppose the firm has in place an incentive-compatible usage-based contract $(q(.), \tau(.))$ for the entire customer interval $[\underline{\theta}, \overline{\theta}]$, which may or may not be optimal. For any incentive-compatible contract, the proof of Lemma 1 establishes that both $q(\theta)$ and $\tau(\theta)$ are non-decreasing, and that informational rent is non-decreasing. These observations provide some intuition for the results that follow.

We now analyze how customer choice is affected if the firm also introduces a fixed-fee contract T and leaves the usage-based contract $(q(.), \tau(.))$ *unchanged.* The surplus that a customer of type θ gets from choosing the fixed-fee contract is $v(\theta) - T$. Therefore, a customer of type θ will switch to the fixed fee contract if and only if

$$v(\theta) - T \geq U(q(\theta), \theta) - \tau(\theta), \tag{6}$$

where it is assumed that an indifferent customer chooses the fixed-fee contract. Note that equation (6) is equivalent to

$$v(\theta) - U(q(\theta), \theta) + \tau(\theta) \geq T. \tag{7}$$

The expression on the LHS of (7) has a simple economic interpretation. It is the difference between the maximum value $v(\theta)$, obtained from consumption of the optimal quantity $\alpha(\theta)$, and the informational rent $[U(q(\theta), \theta) - \tau(\theta)]$ that type θ gets from the incentive-compatible usage-based contract. Consequently, it is the maximum amount that the firm can charge for the fixed contract if the firm wishes type θ to adopt this fixed-fee contract. Lemma 2 shows that this amount is non-decreasing as θ increases.

Lemma 2 *For any incentive-compatible usage-based contract $(q(.), \tau(.))$, the function $v(\theta) - U(q(\theta), \theta) + \tau(\theta)$ is:*

(a) non-decreasing in θ for all θ in $[\underline{\theta}, \overline{\theta}]$.

(b) strictly increasing at $\hat{\theta}$ if $q(\hat{\theta}) < \alpha(\hat{\theta})$ for some $\hat{\theta}$ in $[\underline{\theta}, \overline{\theta}]$.

(c) strictly increasing for all θ in $[\underline{\theta}, \overline{\theta})$ if $q(\theta) < \alpha(\theta)$ for all θ in $[\underline{\theta}, \overline{\theta})$.

This lemma leads to the following proposition:

Proposition 2 *If the firm introduces a fixed-fee T in addition to an existing usage-based contract $(q(.), \tau(.))$ 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 $v(\overline{\theta}) - T < U(q(\overline{\theta}), \overline{\theta}) - \tau(\overline{\theta})$, then all customers continue to adopt the usage-based contract, and

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

switch to the fixed-fee contract, where

$$\theta_F = \min\{\theta : v(\theta) - U(q(\theta), \theta) + \tau(\theta) = T\}. \tag{8}$$

The proof of this proposition is simple, and proceeds as follows. Combining (7) and the fact that $v(\theta) - U(q(\theta), \theta) + \tau(\theta)$ is non-decreasing (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 since $v(\theta) - U(q(\theta), \theta) + \tau(\theta)$ is non-decreasing in $[\underline{\theta}, \overline{\theta}]$, this ensures that there will be at least[5] one type θ for which $v(\theta) - U(q(\theta), \theta) + \tau(\theta) = T$. Since θ_F is the lowest such value of θ, and indifferent types adopt the fixed-fee contract, this proves part (c), which completes the proof.

The customers in $[\theta_F, \overline{\theta}]$ who switch to the fixed-fee contract will have usage-levels of $\alpha(\theta)$, which will be at least as high as $q(\theta)$. In addition, so long as $q(\theta_F) < \alpha(\theta_F)$, equation (8) implies that:

$$T - \tau(\theta_F) = v(\theta_F) - U(q(\theta_F), \theta_F) > 0, \tag{9}$$

and since $v(\theta) - U(q(\theta), \theta) + \tau(\theta)$ is non-decreasing, there will be a positive fraction of these customers for whom $T > \tau(\theta)$, and who therefore pay a strictly *higher* total price. Let θ_H be the highest value of θ which satisfies $\tau(\theta) = T$. Assuming that such a value of θ_H exists in $[\underline{\theta}, \overline{\theta}]$, Lemma 2 implies that $\theta_H > \theta_F$. Therefore, while the fraction $[\theta_F, \theta_H]$ pay a higher price, the fraction $(\theta_H, \overline{\theta}]$ will pay a lower price than they were paying under the usage-based contract. The firm therefore gains revenue from the fraction $[\theta_F, \theta_H]$, but may lose revenue from a fraction $(\theta_H, \overline{\theta}]$. The firm lowers its costs in the interval $[\theta_F, \overline{\theta}]$, as a consequence of having no contract administration costs from these customers. This is illustrated in Figure 2.

3.3 Profit-improving fixed-fees

This sub-section establishes that the profits from the optimal usage-based contract in the absence of a fixed fee can always be strictly improved by the introduction of a fixed-fee. The proposition is stated below:

Proposition 3 *If $c > 0$, then it is always strictly profit-improving for the firm to offer a fixed-fee contract.*

Again, the proof of this proposition is relatively straightforward, and proceeds as follows. Let $(q^*(.), \tau^*(.))$ be the optimal usage-based contract in the *absence* of a fixed-fee. This is simply the contract which satisfies (5) over the entire interval $[\underline{\theta}, \overline{\theta}]$. According to Lemma 1, under this usage-based contract, the usage level of type $\overline{\theta}$ is $q^*(\overline{\theta})$, which satisfies

$$U_1(q^*(\overline{\theta}), \overline{\theta}) = c. \tag{10}$$

Since $U_1(q, \theta)$ is decreasing in q, and $U_1(\alpha(\overline{\theta}), \overline{\theta}) = 0$ by definition, (10) implies that $q^*(\overline{\theta}) < \alpha(\overline{\theta})$, so long as $c > 0$, which implies that $v(\overline{\theta}) - U(q(\overline{\theta}), \overline{\theta}) > 0$. Therefore, there exists a value of $T > \tau(\overline{\theta})$ such that

$$v(\overline{\theta}) - [U(q^*(\overline{\theta}), \overline{\theta}) - \tau^*(\overline{\theta})] > T. \tag{11}$$

Since $\tau^*(.)$ is non-decreasing, the value of T defined in (11) is strictly greater than $\tau^*(\theta)$, for all θ in $[\underline{\theta}, \overline{\theta}]$. Suppose the firm offers a fixed-fee contract priced at this value T. We know from Lemma 2(c) that for an optimal usage-based contract, $v(\theta) - U(q^*(\theta), \theta) + \tau^*(\theta)$ is strictly increasing in $[\underline{\theta}, \overline{\theta})$. Therefore, there is exactly one type θ_F in for which

$$v(\theta_F) - U(q^*(\theta_F), \theta_F) + \tau^*(\theta_F) = T. \tag{12}$$

Proposition 2 establishes that all customers of type $\theta \geq \theta_F$ will switch to the fixed-fee contract. This raises the firm's revenues, since $T > \tau^*(\theta)$. In addition, the firm no longer bears the contract administration cost $cq^*(\theta)$ for customers of type $\theta \geq \theta_F$. Consequently, the firm's profits from the customers in the interval $[\theta_F, \overline{\theta}]$ are strictly increased by introducing the fixed-fee contract T.

Proposition 2 also shows that all customers of type $\theta < \theta_F$, will continue to adopt the usage-based contract, with no impact on the firm's profits. Therefore, total firm profits strictly increase. This completes the proof.

In the absence of changes to $(q^*(.), \tau^*(.))$, the increase in the firm's profits comes from two separate sources – an increase in revenues from a higher price, and a decrease in costs from lower administration costs. The former is feasible due to an increase in total surplus induced by any fixed-fee contract – each customer who adopts the fixed-fee contract chooses their globally optimal usage level $\alpha(\theta)$, and T can be chosen so that part of this increase in utility is shared by the firm.

3.4 Optimal fixed-fee and usage-based contracts

The main result of this sub-section is to show that the optimal usage-based contract in the presence of a fixed-fee contract is *independent* of the value of the optimal fixed-fee contract. As a consequence, the simultaneous derivation of the optimal combination of usage-based and fixed-fee contracts is simplified considerably.

Proposition 3 has established the desirability of a fixed-fee contract, but does not indicate what the optimal value of T should be. Besides, the contract $(q^*(.), \tau^*(.))$ considered in Proposition 3 was optimal in the *absence of* T. When a fraction $[\theta_F, \overline{\theta}]$ of the customers no longer adopt this usage-based contract, the firm may be able to redesign the contract for the remaining customer types $[\underline{\theta}, \theta_F]$ in a profit-improving way. This will change the value of the lowest type θ_F who is indifferent. Consequently, in order to evaluate the net profit impact of each feasible fixed contract, one needs to consider optimally redesigned usage-based contracts for a range of sub-intervals. In addition, it is not yet guaranteed that a combination of this form – a value of T, and the optimal unconstrained usage-based contract for the corresponding interval $[\underline{\theta}, \theta_F]$ – is in fact the global optimum. For instance, a higher value of T, and a correspondingly constrained incentive-compatible contract may actually yield higher profits.

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 $(q^*(.), \tau^*(.))$ by solving:

$$\begin{aligned} U_1(q^*(\theta),\theta) &= c + U_{12}(q^*(\theta),\theta)\frac{1-F(\theta)}{f(\theta)}\ \forall\theta; \\ \tau^*(\theta) &= U(q^*(\theta),\theta) - \int_{\underline{\theta}}^{\theta} U_2(q^*(x,\theta_U),x)dx\ \forall\theta. \end{aligned} \tag{13}$$

(b) Find the optimal interval $[\theta_F^*, \overline{\theta}]$ which should adopt the fixed-fee contract by solving:

$$\theta_F^* = \arg\max_{\theta_F} \int_{\underline{\theta}}^{\theta_F} [\tau^*(\theta) - cq^*(\theta)]f(\theta)d\theta + [1 - F(\theta_F]\,[v(\theta_F) - U(q^*(\theta_F),\theta_F) + \tau^*(\theta_F)]\,. \tag{14}$$

(c) Determine the optimal fixed-fee contract:

$$T^* = v(\theta_F^*) - U(q^*(\theta_F^*),\theta_F^*) + \tau^*(\theta_F^*). \tag{15}$$

The proof of Proposition 4, which is presented in Appendix A, proceeds in three stages. First, we characterize the sub-problem of designing an optimal usage-based contract for the customer interval $[\underline{\theta}, \theta_F]$, subject to the requirement that every customer in $[\theta_F, \overline{\theta}]$ prefers a fixed-fee contract with a *fixed value* of T, and that all customers in $[\underline{\theta}, \theta_F)$ either choose the usage-based contract, or choose a usage level of zero. Next, the necessary conditions for any solution to this sub-problem (which are derived by point-matched optimization of the Lagrangian, as in Banker and Datar, 1989) are used to characterize the solution to the problem of simultaneously choosing the best fixed-fee contract and usage-based contract, subject to the requirement that every customer in $[\theta_F, \overline{\theta}]$ prefers the fixed-fee contract. Finally, the first-order necessary conditions for this problem immediately lead to the main result.

Proposition 4 is a surprising result. It shows that when the firm uses 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, and under an additional set of constraints. The result also reduces a fairly complex problem into a relatively simple sequence. Proposition 1 shows that if the firm had to design the optimal usage-based contract exclusively for a smaller interval, that the redesigned usage-based contract would always be different from $(q^*(.), \tau^*(.))$, and would result in higher usage and surplus for all the customers in the sub-interval. Furthermore, when the firm takes into account the introduction of the fixed contract, this introduces a new (and infinite) set of individual rationality inequality constraints. These constraints make establishing sufficiency difficult – and the entire set of constraints changes as one varies both the level of the fixed-fee contract T, and the sub-interval $[\theta_F, \overline{\theta}]$ that the firm wants to induce to adopt T – both of which necessitate changes in the optimal usage-based contract.

Proposition 4 establishes that this complicated sequence can be reduced to a simple problem of determining a globally optimal usage-based contract (which has a unique solution when $U_{122} \leq 0$), 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 readjustment of the optimal usage-based contract) does not reduce the surplus of any customer, relative to the scenario in which only a usage-based contract is offered. Since the surplus of those customers adopting the fixed-fee contract increases[6], this means that consumer surplus strictly increases as well. Proposition 3 ensures that firm profits also strictly increase, implying that total surplus increases as well.

4 Examples

The results derived in section 3 are applied in two examples, which explore the profitability and welfare impact of fixed-fee and usage-based contracts for different customer preferences, contract administration costs and customer type distributions. The first example is described in more detail than the second, to illustrate clearly how to use the results of Proposition 4, and so that the different effects on usage, profits and welfare can be precisely isolated and explained.

The additional notation used in both examples is summarized in Table 2.

4.1 Uniform customer distribution

The first example uses a quadratic utility function of the form:

$$U(q,\theta) = (\beta+\theta)q - \frac{1}{2}q^2, \tag{16}$$

and assumes that customer types θ are uniformly distributed in $[0,1]$, that is, $f(\theta)=1$ and $F(\theta)=\theta$.

(16) represents a customer utility function which is additively separable into two parts – a customer benefit from usage which increases linearly, and a (non-price) customer cost from usage, which is increasing and convex in usage. The customer cost could arise out of the need to acquire complementary assets as one increases usage. For instance, suppose the information good in question is Internet bandwidth. As a company increases it's use of bandwidth, it has to increase the capacity of switches and routers in its offices, it needs to acquire more disk space for caching, and it has to get more personnel for network administration. All of these costs are generally independent of the type of customer, and simply depend on usage levels. Similarly, if the information good in question is corporate software, more usage may require more processing time on a shared set of servers. As usage increases, this utilization of shared processing power imposes an increasing cost on the performance of the customer's other software.

The parameter β influences the marginal benefit of usage linearly and uniformly across customer types. It is used to illustrate how pricing varies as marginal utility from usage is shifted. Also, it is easily verified that $U_{122}(q,\theta) = 0$, and therefore, by Lemma 1, the first-order conditions in (5) describe the unique optimal usage-based contract.

The maximal usage $\alpha(\theta)$ and the corresponding value $v(\theta)$ are obtained by unconstrained maximization of $U(q,\theta)$:

$$\alpha(\theta) \;=\; \beta+\theta; \tag{17}$$

$$v(\theta) = \frac{(\beta+\theta)^2}{2}.$$

Applying (13) from Proposition 4, one gets:

$$q^*(\theta) = 2\theta + \beta - (c+1). \tag{18}$$

Therefore, if $\beta > c+1$, all types are offered positive usage, and if $\beta \le c+1$, only a fraction of the types have $q^*(\theta) \ge 0$. We focus on the latter case[7]. Define $\theta_L(\beta, c)$ as the lowest type which chooses a non-negative usage level:

$$\theta_L(\beta, c) = \frac{c+1-\beta}{2}. \tag{19}$$

Under this assumption, the optimal usage levels are:

$$\begin{aligned} q^*(\theta) &= 2\theta + \beta - (c+1) \text{ for } \theta \ge \theta_L(\beta, c); \\ q^*(\theta) &= 0 \text{ for } \theta < \theta_L(\beta, c), \end{aligned} \tag{20}$$

and the optimal usage-based price for each type is:

$$\tau^*(\theta) = U(q^*(\theta), \theta) - \int_{\theta_L(\beta,c)}^{\theta} U_2(q^*(x), x)dx, \tag{21}$$

which solves to:

$$\begin{aligned} \tau^*(\theta) &= \frac{\beta^2 + 2\beta(c+1) - 3(c+1)^2}{4} + 2\theta(c+1) - \theta^2 \text{ for } \theta \ge \theta_L(\beta, c); \\ \tau^*(\theta) &= 0 \text{ for } \theta < \theta_L(\beta, c). \end{aligned} \tag{22}$$

Since $q^*(\theta)$ is linear in θ, one can invert (20) and substitute it into (22) to yield price as a function of quantity:

$$p(q) = \frac{1}{2}[(1+\beta+c)q - \frac{q^2}{2}]. \tag{23}$$

Therefore, the optimal usage-based pricing scheme is strictly concave in quantity. If one interprets $\frac{(1+\beta+c)}{2}$ as the standard unit price, and $\frac{q^2}{4}$ as the level of quantity discount, this implies that the unit price increases when either administration costs c or marginal utility β increase. Also, since the absolute value of the discount is not sensitive to changes in c or β, this implies that as a percentage, quantity discounts decrease as either c or β increase.

The optimal contract and pricing function are depicted in Figure 3.

The first-order condition for (14) from Proposition 4 is:

$$\frac{\partial}{\partial \theta_F}\left(\int\limits_{\theta_L(\beta,c)}^{\theta_F} [\tau^*(\theta) - cq^*(\theta)]d\theta + [1-\theta_F]\left[v(\theta_F) - U(q^*(\theta_F),\theta_F) + \tau^*(\theta_F)\right]\right) = 0, \qquad (24)$$

which, when solved, yields $\theta_F^*(\beta, c)$, the type which is indifferent between the fixed-fee and usage-based contract:

$$\theta_F^*(\beta,c) = 1 + 2c - \sqrt{c(2+2\beta+3c)}, \qquad (25)$$

It is easily verified analytically that the objective function in (24) is strictly concave in θ_F, which establishes that the necessary condition (24) is sufficient. The optimal fixed-fee contract is obtained by solving (15) from Proposition 4:

$$T^*(\beta,c) = \frac{(1+\beta+c)^2}{4} - c[1-\theta_F^*(\beta,c)] \qquad (26)$$

Figure 3 also depicts the value of the fixed-fee contract $T^*(\beta, c)$, superimposed on the usage-based contract, and the corresponding value of $\theta_F^*(\beta, c)$.

Before discussing profits and welfare, we analyze the effects of varying c and β on customer adoption. Clearly, $\theta_L(\beta, c)$ is strictly increasing in c. In addition, $\theta_F^*(\beta, c)$ is strictly decreasing in c, for c between 0 and $\frac{1+\beta}{3}$. Consequently, as c increases in this range, the set of customers adopting the usage-based contract shrinks, while the set of customers adopting the fixed-fee contract increases. However, the total fraction of customers who buy the product decreases. This is illustrated in Figure 4(a). The value of $c = \frac{1+\beta}{3}$ is also exactly the value at which $\theta_L(\beta, c) = \theta_F^*(\beta, c)$, at which point no customers adopt the usage-based contract, and all customers for whom $\theta \geq \theta_F^*(\beta, c)$ adopt the fixed-fee contract.

Not surprisingly, increasing β has a strictly positive effect on adoption in general. Both $\theta_L(\beta, c)$ and $\theta_F^*(\beta, c)$ are strictly decreasing in β – the former implies that the total number of adopters strictly increases in β, and the latter implies that the number of adopters of the fixed-fee contract also increases in β. The function $\theta_L(\beta, c)$ decreases more rapidly[8] than $\theta_F^*(\beta, c)$, and therefore, the number of customers choosing the usage-based contract also increases in β. This is illustrated in Figure 4(b).

Clearly, the profits from the optimal combination of fixed-fee and usage-based contracts should be higher than either the profits from offering only the optimal usage-based contract or the profits from offering only a fixed-fee contract. The profits from offering only the optimal usage-based contract are:

$$\Pi_U(\beta, c) = \int_{\theta_L(\beta,c)}^{1} [\tau^*(\theta) - cq^*(\theta)]d\theta = \frac{(1+\beta-c)^3}{12}. \tag{27}$$

Also, if the firm offered only a fixed-fee contract, it would set the price at $\frac{2(1+\beta)^2}{9}$, and its profits[9] would be:

$$\Pi_F(\beta) = \frac{2(1+\beta)^3}{27}. \tag{28}$$

The firm's profits from the optimal combination of contracts are:

$$\Pi(\beta, c) = \int_{\theta_L(\beta,c)}^{\theta_F^*(\beta,c)} [\tau^*(\theta) - cq^*(\theta)]d\theta + T^*(\beta, c)[1 - \theta_F^*(\beta, c)]. \tag{29}$$

The expression for $\Pi(\beta, c)$ is omitted since it is rather cumbersome. The three functions are illustrated in Figure 5(a). As expected, $\Pi_F(\beta) = \Pi(\beta, \frac{1+\beta}{3})$, and $T^*(\beta, \frac{1+\beta}{3}) = \frac{2(1+\beta)^2}{9}$, which confirms that the firm should offer only the fixed-fee contract for $c \geq \frac{1+\beta}{3}$. Also, $\Pi_U(\beta, 0) = \Pi(\beta, 0)$, which confirms that the firm should offer only the usage-based contract for $c = 0$. The corresponding levels of consumer surplus are depicted in Figure 5(b).

Figure 5(a) suggests that the firm's profits are strictly decreasing in c. This is confirmed by differentiating (29) with respect to c:

$$\Pi_2(\beta, c) = \frac{2c[(1+\beta)^2 + 9(1+\beta)c + 9c^2]}{\sqrt{c(2+2\beta+3c)}} - \frac{(1+\beta)^2 + 14(1+\beta)c + 21c^2}{4}, \tag{30}$$

which is negative for $0 \leq c \leq \frac{1+\beta}{3}$. However, the corresponding result does not hold for either consumer surplus or total surplus. Total surplus is:

$$S(\beta, c) = \int_{\theta_L(\beta,c)}^{\theta_F^*(\beta,c)} [U(q^*(\theta), \theta) - cq^*(\theta)]d\theta + \int_{\theta_F^*(\beta,c)}^{1} v(\theta)d\theta, \tag{31}$$

and consumer surplus is:

$$C(\beta, c) = \int_{\theta_L(\beta,c)}^{\theta_F^*(\beta,c)} [U(q^*(\theta), \theta) - c\tau^*(\theta)]d\theta + \int_{\theta_F^*(\beta,c)}^{1} [v(\theta) - T^*(\beta, c)]d\theta. \tag{32}$$

Again, the expressions for these functions are omitted, since they are algebraically complex and provide no additional insight. Both $S(\beta,c)$ and $C(\beta,c)$ are first decreasing, and then increasing in c, and both have a slope of zero at $c = \frac{1+\beta}{3}$. Figure 5(c) plots consumer surplus and total surplus for fixed β and varying c as a fraction of $\frac{1+\beta}{3}$. For all values of β in the interval of interest, one can show that total surplus $S(\beta,c)$ decreases for a little over half the range of c, and then increases subsequently until $c = \frac{1+\beta}{3}$, after which it is constant. Consumer surplus $C(\beta,c)$ decreases for a small fraction of the interval, and then also increases steadily until $c = \frac{1+\beta}{3}$, after which it is constant[10]. Moreover, while $S(\beta,c)$ and $\Pi(\beta,c)$ are maximized for $c = 0$, consumer surplus is maximized at $c = \frac{1+\beta}{3}$.

These observations are explained in some detail in Figure 6. When c increases, there are two sets of effects on total surplus. Firstly, 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 as c increases, thereby reducing total surplus. In addition, there is a *direct cost effect* – the costs borne by the firm per unit of usage increase, which changes both firm profits and total surplus – however, the decrease in quantity demanded by the adopters of the usage-based contract may offset this unit-cost increase. However, there are also two *positive indirect effects*. An increase in c increases the number of adopters of the fixed-fee contract. This implies that there is a larger fraction of the customer population whose usage levels are at their globally optimal value $\alpha(\theta)$. In addition, these customers no longer impose the contract administration costs $cq^*(\theta)$ on the firm. Both of these increase total welfare.

Figure 5(c) indicates that the negative effects dominates for about the first half of the range $0 \leq c \leq \frac{1+\beta}{3}$, after which, the positive indirect effect dominates. As c approaches $\frac{1+\beta}{3}$, the fraction of customers adopting the usage-based contract approaches zero, which is why the changes in total surplus also tend to zero. Finally, at $c = \frac{1+\beta}{3}$ (and beyond), there are no more customers adopting the usage-based contract, so further increases in c have no effect on surplus.

Similarly, as c increases from zero, there are two effects on consumer surplus – a negative indirect effect as fewer customers adopt any contract, and a positive indirect effect as more customers shift to the fixed-fee contract. In this case, the positive effect dominates almost immediately. Once all the adopting customers have shifted to the fixed-fee contract (which occurs at $c = \frac{1+\beta}{3}$), there is no further impact from increasing c. Further implications of these results are discussed in Section 5.

4.2 Positively-skewed customer distribution

The next example uses a simpler quadratic utility function of the form:

$$U(q, \theta) = \theta q - \frac{1}{2}q^2, \tag{33}$$

and assumes that customer types θ are exponentially distributed with mean β, that is, $f(\theta) = \frac{e^{-\theta/\beta}}{\beta}$, and $F(\theta) = 1 - e^{-\theta/\beta}$. Relative to a flat distribution, the exponential distribution is positively-skewed – the mean and median are higher than the mode – in fact, $f(\theta)$ is strictly decreasing in θ. This represents a scenario where there are a relatively higher number of customers who have a low utility from usage, and relatively fewer higher types[11]. Many markets for information goods are well characterized by a positively-skewed type distribution, with a number of customers who wish to use the good only occasionally, and relatively fewer 'power-users' whose usage levels are very high. Examples include residential Internet access, music and online financial information. In addition, certain types of corporate software display similar demand characteristics.

Varying β varies both the mean and the shape of the distribution. An increase in β shifts customer types towards the right, resulting in fewer lower-type customers and more higher-type customers. Consequently, the average customer type increases, as does the average demand.

Equation (17) implies that $v(\theta) = \frac{\theta^2}{2}$ and $\alpha(\theta) = \theta$. Proceeding as in section 4.1, applying Proposition 4 yields the optimal usage-based contract:

$$\begin{aligned} q^*(\theta) &= \theta - (\beta + c) \text{ for } \theta \geq \theta_L(\beta, c); \qquad (34) \\ q^*(\theta) &= 0 \text{ for } \theta < \theta_L(\beta, c), \end{aligned}$$

where $\theta_L(\beta, c) = (\beta + c)$, and:

$$
\begin{aligned}
\tau^*(\theta) &= (\beta + c)(\theta - (\beta + c)) \text{ for } \theta \geq \theta_L(\beta, c); \\
&= 0 \text{ for } \theta < \theta_L(\beta, c).
\end{aligned} \tag{35}
$$

The optimal usage-based contract is therefore linear in usage. The first-order condition for (14) in Proposition 4 solves to:

$$\theta_F^*(\beta, c) = \frac{(\beta + c)^2}{2c}, \tag{36}$$

and the optimal fixed-fee contract simplifies to:

$$T^*(\beta, c) = \frac{\beta(\beta + c)^2}{2c}. \tag{37}$$

It can be shown that $\frac{\partial \theta_L(\beta,c)}{\partial c} > 0$ and that $\frac{\partial \theta_F^*(\beta,c)}{\partial c} < 0$ for all $c > 0$. Consequently, increases in c shrink the fraction of customers who adopt the usage-based contract. This segment shrinks to zero when $c = \beta$, which is the point at which $\theta_F^*(\beta, c) = \theta_L(\beta, c)$; for $c > \beta$, only a fixed-fee contract is offered.

Since $\frac{\partial \theta_F^*(\beta,c)}{\partial \beta} > \frac{\partial \theta_L(\beta,c)}{\partial \beta} > 0$ for all $\beta > c$, an increase in β increases the number of types adopting the usage-based contract, and increases the number of non-adopting types[12]. However, the shape of the distribution also changes with β, and therefore changes in interval widths do not directly correspond to changes in customer density. If $\beta > c$, the fraction of customers adopting the fixed-fee contract is $1 - F(\theta_F^*(\beta, c))$, and the fraction of customers adopting the usage-based contract is $F(\theta_F^*(\beta, c) - F(\theta_L(\beta, c))$. Since

$$\frac{\partial}{\partial \beta}[1 - F(\theta_F^*(\beta, c)] = \frac{(c^2 - \beta^2)e^{-\frac{(\beta+c)^2}{2\beta c}}}{2\beta^2 c} \tag{38}$$

is strictly negative for $\beta > c$, an increase in β results in a lower fraction of customers adopting the fixed fee contract. Note that these are comparative statics results – the shift away from the fixed-fee contract is after taking into account the adjustments that the firm will make to its pricing schedule as a consequence of this increase in β – clearly, both the fixed fee and the usage-based fees will also increase as β increases. Also,

$$\frac{\partial}{\partial \beta}[F(\theta_F^*(\beta, c)) - F(\theta_L(\beta, c))] = \frac{(\beta^2 - c^2)e^{-(1-\frac{b}{2c}+\frac{c}{b})} + 2c^2 e^{-\frac{\beta+c}{\beta}}}{2\beta^2 c} \tag{39}$$

is strictly positive for $\beta \geq c$, an increase in β causes an increase in the fraction of customers adopting the usage-based contract. When $\beta \leq c$, (38) indicates that as β increases towards c, the fraction of customers adopting the fixed-fee contract (which is the only contract offered in this case) increases. Figure 7 illustrates the customer intervals which adopt the fixed-fee contract, the usage-based contract and neither, as β varies.

The firm's profits from offering just the optimal usage-based contract, from offering just a fixed-fee contract, and from offering the optimal combination of contracts, are compared below:

$$\Pi_U(\beta, c) = \int_{\theta_L(\beta,c)}^{\infty} [\tau^*(\theta) - cq^*(\theta)]f(\theta)d\theta = \beta^2 e^{-\frac{\beta+c}{\beta}}; \tag{40}$$

$$\Pi_F(\beta) = \max_p [p(1 - F(\sqrt{2p})] = \frac{2\beta^2}{e^2}; \tag{41}$$

$$\Pi(\beta, c) = \int_{\theta_L(\beta,c)}^{\theta_F^*(\beta,c)} [\tau^*(\theta) - cq^*(\theta)]f(\theta)d\theta + T^*(\beta, c)[1 - \theta_F^*(\beta, c)] = \beta^2 e^{-\frac{\beta+c}{\beta}} + \beta c e^{-\frac{(\beta+c)^2}{2\beta c}}.$$

Figure 8(a) depicts how these functions vary with c. As expected, $\Pi_U(\beta, 0) = \Pi(\beta, 0)$, and $\Pi_F(\beta) = \Pi(\beta, \beta)$, which confirm that only the optimal usage-based contract should be offered for $c = 0$, and that only a fixed-fee contract should be offered for $c \geq \beta$. It can also be confirmed that $T^*(\beta, \beta) = 2\beta$, the optimal fixed-fee contract in the absence of any usage-based contract. Figure 8(b) charts how consumer surplus varies with c.

Under the optimal combination of fixed-fee and usage-based contracts, consumer surplus $C(\beta, c)$ and total surplus $S(\beta, c)$ solve to:

$$C(\beta, c) = \beta^2 e^{-\frac{\beta+c}{\beta}} + \beta(\beta + c)e^{-\frac{(\beta+c)^2}{2\beta c}}; \tag{42}$$

$$S(\beta, c) = 2\beta^2 e^{-\frac{\beta+c}{\beta}} + \beta(\beta + 2c)e^{-\frac{(\beta+c)^2}{2\beta c}}. \tag{43}$$

As was the case in section 4.1, after decreasing for a while, both consumer surplus and total surplus increase with c, consumer surplus is maximized at $c = \beta$, and total surplus decreases over a wider range of c values than consumer surplus. Again, these results are discussed in more detail in Section 5

5 Discussion and Conclusions

We establish that a firm should offer its customers a combination of usage-based and unlimited usage pricing schedules, so long as the cost of administering the usage-based schedule is not unduly high. As the costs of administering and supporting these contracts increase, the desirability of non-linear pricing diminishes. However, the magnitude of these costs still play a crucial role in determining the optimal fixed-fee.

These conclusions contrast with earlier results on pricing information goods, and from nonlinear pricing theory in general. For instance, in Varian (2000), it is proved that when customers are of two types and utility is linear in usage, a 'buy only' pricing regime (which corresponds to offering only an unlimited usage fixed-fee in our model) is strictly preferable to any pricing regime that includes renting (usage-based pricing in our model) so long as the transaction costs of renting are positive (non-zero administration costs c in our model). The result is intuitively appealing, and highlights the importance of considering usage transaction costs when pricing information goods. However, our results show that for a continuum of customer types, it is in fact optimal to offer both usage-based and fixed fee pricing, for a range of positive administration (or transaction) costs – we establish that the two kinds of pricing can optimally co-exist, by generalizing the specification of customers preferences and the distribution of customer heterogeneity.

In contrast, results from nonlinear pricing theory under assumptions similar to those made in this paper (see, for instance, Maskin and Riley, 1984, or Wilson, 1993) suggest that the optimal monopoly pricing structure is purely usage-based. These models do not generally explicitly consider administration costs. We show that the optimality of a pure-usage based contract is highly sensitive to the absence of these administrative costs – in fact, Proposition 3 has established that when there are no marginal production costs from additional usage, a purely usage-based pricing scheme is no longer optimal for any $c > 0$. Given the fact that monitoring usage, billing and collection are always expensive, this is an important new conclusion for any business which is pricing information goods.

We have also proved 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 weak – they cover a vast spectrum of utility functions and distributions. Applying Proposition 4 is relatively straightforward, as illustrated by the examples in Section 4. This will enable relatively easier development of focused and rigorous models for specific information pricing problems.

5.1 Customer adoption and pricing structure

As the unit cost of monitoring usage increases relative to the value customers place on usage, results from both the examples in Section 4 show that pricing should be altered to reduce the number of usage-based adopters, and that this should be accomplished not only by reducing the fraction of lower-valuation customers who adopt (by increasing unit prices), but also by shifting the high-end usage-based customers to the fixed-fee contract. When the opposite occurs – when the value of usage increases relative to the cost of monitoring usage – a firm should alter its pricing in a way that increases both the number of usage-based adopters, as well as its total market coverage. However, the optimal adjustment of price should be such that there is a larger shift towards usage-based pricing. In fact, the optimal price adjustments in response to an increase in product value can lower number of fixed-fee adopters, which indicates that the increase in fixed fee should be more than proportionate to the increase in value.

The optimal shape of the usage-based pricing schedule is also sensitive to these parameters. As the unit valuation of consumers increases, a firm will clearly increase its unit prices. However, as shown in Section 4.1, the percentage of quantity discount offered should progressively decrease. The intuition here is that in equilibrium, the increase in marginal value will cause an increase in the level of usage chosen by each customer, since the optimal unit price increase is a fraction of the increase in value. In addition, it is optimal for the firm to induce a higher fraction of the market to adopt its fixed-fee contract. The relative benefits of the quantity discount for the firm are consequently lower overall, and this naturally leads to a decrease in the discount. Additionally, when the unit cost of monitoring usage increases, a firm should increase unit prices and decrease in the percentage of quantity discount. While the direction of the result is similar, the intuition is different – in this case, the firm actually wants to induce lower levels of usage for all adopters of the usage-based contract, and accomplishes this by increasing unit price – there is no corresponding increase in unit value driving increased usage.

5.2 Changes in customer characteristics and market evolution

Many markets for information goods are characterized by a relatively high concentration of occasional users (or customers who place a low value on usage), and a smaller concentration of high-usage customers. This was the kind of customer distribution analyzed in section 4.2. Facing a market of this kind, a firm may be tempted to maximize its market coverage by offering pay-per-use pricing for the large number of low-end customers (perhaps through an ASP, in the case of an enterprise software manufacturer), and high-priced unlimited usage contracts for the high-end customers. However, our results indicate that while both types of contracts may co-exist, it is often optimal to price high enough exclude a substantial fraction (substantially higher than 50% in our examples) of the possible customers, all of whom are on the low-end of the market. While one may gain higher market coverage from an attractive usage-based pricing scheme, these benefits are outweighed by the cannibalization of revenues that could have been garnered from higher fees the high-end customers. This is a particularly important insight for firms that are experimenting with pricing targeted at low-end usage, and monitored on per-cycle basis, for instance. The current pricing model many software companies use – which segments customers through a usage-based pricing scheme by tying licensing fees to processor speed, and in which prices are high enough to exclude a number of small businesses – may be more profitable in the long run. Administration costs are reduced to near zero by implementing usage limits into the software, and consequently, the fact that high-end software is priced almost purely based on some measure of usage is consistent with our model's results.

An interesting example of a shift towards more low-end customers occurred in segments of the enterprise software market during the dotcom boom in the late 1990's. There was a sudden jump in the number of early-stage companies that needed database and application server software to run their ecommerce sites, but whose site traffic and sales volumes required relatively low usage levels. Many of these companies ended up spending substantial fractions of their budgets on licenses – far more than they would have spent if there were lower-priced usage-based contracts available. In fact, companies like Oracle probably suffered some revenue loss when hosting companies like Exodus started renting out existing software licenses on a monthly basis. Had Oracle responded by offering its own ASP-based monthly scheme, or some other kind of per-use pricing targeted at lower-end customers, it is likely that while they may have gained market share from some mySQL users, they would have lowered revenue overall.

Often, however, the change in customer distribution over time is in the other direction – towards the high end. In the early stages of technology markets, is common for a relatively small fraction of early innovators to constitute a bulk of total usage, and for there to be a high concentration of occasional experimenters on the low end of the market. As the market matures, the usage distribution evens out, and for successful products, tends to increase on average. For instance, monthly usage levels per customer in the online services market have been steadily increasing over the last few years – as of late 2001, average AOL usage had more than doubled to about 40 hours per month – and generalizing Jupiter Media-Metrix survey data from May 2001 on overall residential online usage across the USA indicates that the distribution of usage levels per customer is flattening out, especially below the mean. This is the kind of distributional change corresponding to an increase in β in our example from section 4.2, as illustrated in Figure 7. When usage patterns in a market matures in this manner, it is optimal for the provider to penetrate the market with a pricing scheme that is biased highly towards inducing adoption of a relatively low fixed-fee, and to gradually increase the number of usage-based customers over time, as the average value from the service becomes increasingly higher than the cost of monitoring usage.

AOL's switch to fixed-fee pricing at a relatively early stage in the market (late 1998) is consistent with this prescription. However, examining AOL's pricing structure as of the end of 2001 suggests that substantially less than 20% of their customers are likely to be adopters of their usage-based contract[13]. This could be due to the somewhat low total dollar value of their service relative to usage monitoring costs. It could also be due to competitive pressure (there are competitors like NetZero which offer basic Internet access for less than half of AOL's fixed fee), though AOL does have substantial market power in residential Internet access. Our results suggest that as AOL strengthens its customer lock-in, it would be optimal for them to continue to gradually increase their unlimited usage price, and to also introduce usage-based pricing schedules that span a wider range of their customer base.

5.3 Implications for bundling and versioning

As mentioned in Section 1, Bakos and Brynjolfsson (1999) have shown that it is always optimal for a monopoly seller of information goods to create increasingly large bundles. Two differences between our model and theirs are that we place a finite maximum on consumption (which arises from substitution effects, and related costs of consumption), and that we assume that there is always heterogeneity in per-unit customer valuations (which precludes the complete elimination of incomplete information). Consequently, our model admits a combination of fixed-fee and usage-based pricing as the optimal solution[14]. In this regard, our results are consistent with the predictions of Chuang and Sirbu (1999), albeit in a more generalized model.

Overall, however, our results do complement the basic rationale for increasing bundling. One key insight of Bakos and Brynjolfsson was that an increase in the gross size of the bundle of information goods increased the average per-unit customer valuation, and was therefore profit-improving for a firm under a fixed-fee. The bundle here can be interpreted as the *potential set* of goods the consumer might use, and not the set of goods *actually consumed.* The unit value from usage is likely to increase with the size of this bundle, and this will clearly increase either a fixed-fee or a usage-based price, consequently increasing firm profits. In addition, if the firm chooses the right combination of fixed-fee and usage based contracts, it will increase the fraction of customers who choose a usage-based contract as this per-unit value increases, and so long as the size of the bundle is finite, is likely be able to extract more surplus from these customers than it would have either under a pure fixed-fee, or with a smaller bundle. This suggests, for instance, that as the market for digital music matures and addresses its rights management issues, larger bundles of songs offered on a per-use basis will result in higher profits per song than smaller bundles. PressPlay and MusicNet are therefore likely to be more profitable if they bundle their offering into a unified service, even if they stick with their usage-based pricing model[15].

If one were to interpret q in our model as quality instead of quantity, and θ as the value different customer types place on quality, the model is identical to one of vertical differentiation, with a continuum of possible product versions, a continuum of customer types, and linear costs of versioning. In the context of information goods with multiple features, where quality is proportionate to the number of features, Proposition 3 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 Bhargava and Choudhury (2001). However, our results also indicate that if versioning is costless, it is optimal to offer as many distinct versions as is possible, and to price each version individually. In addition, for low costs of versioning, it is optimal to offer both the 'unlimited features' version, as well as a set of limited feature versions, and to price each of these differently.

5.4 Ongoing research: piracy and competition

Apart from those considered in this study, there are two other important factors that influence the pricing of information goods. The first is the threat of digital piracy. If one assumes a constant expected value from pirated goods that is uniform across customers, this can be incorporated into our model by simply increasing the RHS of the IR constraints.

However, the expected value of pirated information goods is typically a function of both user preferences as well as some measure of the level of legal usage. Consequently, the IR constraints would need to include an additional 'piracy value' function, which depends on total legal usage, and on customer type. The former dependence could be internalized into the price using a technique analogous to a tax mechanism (Groves, 1973, Groves and Radner, 1972), and the latter dependence may be handled using recent results due to Jullien (2000). A model that addresses this extension is part of ongoing research.

The second factor of importance is the presence of competition. The cost structure of information goods often leads to natural monopoly, but there are instances in which price competition is significant. Current results on Bertrand competition with non-linear pricing suggests that undifferentiated price competition may not be sustainable for information goods. For instance, results from Mandy (1992) show that the equilibrium outcome is either minimum average cost pricing (which is not well-defined for information goods, since the average cost is always strictly decreasing in quantity), or pricing under which customers pay an average price equal to marginal cost (which implies a price of zero for information goods). Fishburn, Odlyzko and Siders (1997) provide one way around this, by modeling a repeated game in which one player chooses only a fixed-fee, and the other chooses only a linear usage-based price. Another approach, which may yield more general results, is if one models competition between differentiated information goods. This is another focus of ongoing research. We hope to address these open issues in the near future.

Notes

[1]While there are a number of diverse models that these ASP's use to price their services, most involve some form of rental pricing (Susarla, Barua and Whinston, 2001)

[2]IBM intends to release a comprehensive license manager to enable a variety of flexible pricing models. Currently, customers can download the SCRT (sub-capacity reporting tool), which will track software usage and CPU activity, and generate a pricing report based on this data each month.

[3]Assume the converse – that $\frac{1-F(\theta)}{f(\theta)}$ is non-increasing, but that for some $\hat{\theta}$, $\frac{F(\hat{\theta})-F(\theta)}{f(\theta)}$ is increasing in some interval $[\theta_1, \theta_2]$. This implies that $\frac{F(\hat{\theta})-F(\theta_1)}{f(\theta_1)} < \frac{F(\hat{\theta})-F(\theta_2)}{f(\theta_2)}$. Since $F(\theta_1) < F(\theta_2)$, this implies that $f(\theta_1) > f(\theta_2)$, which in turn implies that $\frac{1-F(\hat{\theta})}{f(\theta_1)} < \frac{1-F(\hat{\theta})}{f(\theta_2)}$, which when added to $\frac{F(\hat{\theta})-F(\theta_1)}{f(\theta_1)} < \frac{F(\hat{\theta})-F(\theta_2)}{f(\theta_2)}$ implies that $\frac{1-F(\theta)}{f(\theta)}$ is increasing somewhere in $[\theta_1, \theta_2]$, a contradiction.

[4]This kind of formulation is standard in models of price screening – see, for instance, section 2 of Anderson (1996). A good exposition of mechanism design, the revelation principle and its applications to pricing can be found in chapter 7 of Fudenberg and Tirole (1991) – in particular section 7.2 describes the revelation principle, and section 7.1 discusses a non-lienar pricing example.

[5]In addition, if the contract $(q(.), \tau(.))$ is optimal, we know from Lemma 1 that all types (except possibly $\overline{\theta}$ – and this occurs only when $c = 0$) will get an allocation $q(\theta) < \alpha(\theta)$, which means that there will be exactly one value of θ for which $v(\theta) - U(q(\theta), \theta) + \tau(\theta) = T$.

[6]Except for those customers of type θ_F^* who are indifferent.

[7]The case analyzed is a little more complex, algebraically – however, it illustrates more interesting trade-offs between fixed-fee and usage-based contracts. Analysis of the other case is available on request.

[8]There are no customers who adopt the usage-based pricing scheme for $c \geq \frac{1+\beta}{3}$, and hence our range of interest for β is $3c - 1 \leq \beta \leq c + 1$. In this range, $\frac{\partial \theta_L(\beta,c)}{\partial \beta} = -\frac{1}{2}$, and $\frac{\partial \theta_F^*(\beta,c)}{\partial \beta} = -\frac{\sqrt{c}}{\sqrt{2+2\beta+3c}}$. Therefore, $\frac{\partial \theta_L(\beta,c)}{\partial \beta} < \frac{\partial \theta_F^*(\beta,c)}{\partial \beta}$ if $2 + 2\beta > c$, or if $\beta > \frac{c-2}{2}$. Since $\frac{c-2}{2} < 3c - 1$ for $c \geq 0$, $\frac{\partial \theta_L(\beta,c)}{\partial \beta} < \frac{\partial \theta_F^*(\beta,c)}{\partial \beta}$ over the entire range – or $\theta_L(\beta, c)$ decreases more rapidly than $\theta_F^*(\beta, c)$.

[9]In the presence of just a fixed-fee p, type θ adopts if $v(\theta) \geq p$, or if $\frac{(\beta+\theta)^2}{2} \geq p$, which implies that the lowest type adopting is $\sqrt{2p} - \beta$. Consequently, the firm's profits from a price p are $p(1 + \beta - \sqrt{2p})$. First order conditons yield the optimal price $p = \frac{2(1+\beta)^2}{9}$.

[10]More precisely, for this analysis, c is set to be $k\left(\frac{1+\beta}{3}\right)$, and k is varied between 0 and 1. Solving $\frac{\partial}{\partial k} C(\beta, k\frac{1+\beta}{3}) = 0$ indicates that consumer surplus is minimized at about $k = 0.0931$, and is maximized in

this range for $k = 1$, where the slope is also zero. The k-values at which the maximum and minimum are obtained are invariant to the magnitude of c, so long as $c \leq \frac{1+\beta}{3}$. Similar analysis for $S((\beta, k\frac{1+\beta}{3})$ indicates that total surplus is minimized at $k = 0.5177$, and is maximized in the range for $k = 0$. Again, the slope is 0 at $k = 1$, but the value of total surplus is higher at $k = 0$.

[11]In fact, the exponential distribution is the most positively skewed distribution one can use while preserving the requirement that the reciprocal of the hazard rate be non-increasing. It has a constant hazard rate $\frac{1}{\beta}$ – which is also an analytically attractive property.

[12]β is a measure of the average unit value θ that customers place on the information good, as was the case in section 4.1 – which places its comparison with c in context. In addition, it also represents the average optimal usage of a customer (since the utility function is quadratic, $\alpha(\theta) = \theta$). These interpretations are discussed further in section 5.

[13]As of the end of 2001, AOL offered a fixed fee of \$23.95. Their best usage-based contract offered 3 hours for \$4.95, and charged \$2.50 perhour thereafter. This suggests that any customer with a usage level higher than 10 hours a month would choose the fixed-fee contract. Less than 20% of all online users in the U.S. have this level of usage, and AOL users display higher usage levels than the average online user.

[14]If customers were homogeneous in our model, the optimal pricing schedule would simply be a fixed fee, and customer surplus would be zero, as obtained in the limit by Bakos and Brynjolfsson (1999).

[15]As of January 2002, PressPlay was offering a four-tier usage-based pricing scheme, with different usage levels for \$9.95, \$14.95, \$19.95 and \$24.95.

References

1. Afeche, P. and Mendelson, H., 2001. Auctioning priorities vs. uniform pricing in delay-sensitive service markets. *Thirteenth Workshop on Information Systems and Economics.*

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

3. Bakos, Y. and Brynjolfsson, E., 1999. Bundling information goods: pricing, profits and efficiency. *Management Science* 45 (12), 1613-1630.

4. Banker, R. and Datar, S., 1989. Sensitivity, precision and the linear aggregation of accounting signals. *Journal of Accounting Research* 27, 21-39.

5. Banker, R., Khosla, I. and Sinha, K. 1998. Quality and competition. *Management Science* 44, 1179-1192.

6. Bhargava, H. and Choudhary, V., 2001. Information goods and vertical differentiation. Working Paper, GSIA, Carnegie-Mellon University.

7. Cheng, H. and Koehler, G., 1999. Optimal pricing policies of web-enabled application services. Available at *http://www.cba.ufl.edu/dis/papers/9910100.pdf*

8. Cheng, H. , Freimer, M., Richmond, W., and Sumita, U., 1998. Optimal repair of degrading computers under asymmetric information and incentive incompatibility. *Computers and Operations Research* 25, 719-728.

9. Choudhary, V., Tomak, K. and Chaturvedi, A., 1998. Economic benefits of software renting. *Journal of Organizational Computing and Electronic Commerce* 8, 277-305.

10. Chuang, J. and Sirbu, M., 1998. Network delivery of information goods: optimal pricing of articles and subscriptions. In Hurley, Kahin, and Varian, editors, *The Economics of Digital Information Goods.* MIT Press, Cambridge, MA.

11. Chuang, J. and Sirbu, M., 1999. Optimal bundling strategy for digital information goods: network delivery of articles and subscriptions. *Information Economics and Policy* 11, 147-176.

12. Dewan, R., Jing, B. and Seidmann, A., 2000. Adoption of Internet-based product customization and pricing strategies. *Proceedings of the 31st Hawaii International Conference on System Sciences.*

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

14. Fishburn, P., Odlyzko, A., and Siders, R., 1997. Fixed fee versus unit pricing for information goods: competition, equilibria, and price wars. *First Monday* 2 (7). Available at *http://www.firstmonday.dk/issues/issue2_7/.*

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

16. Groves, T., 1973. Incentives in teams. *Econometrica* 41, 617-631.

17. Groves, T. and Radner, R., 1972. Allocation of resources in a team. *Journal of Economic Theory* 4, 415-441.

18. Gupta, A., Stahl, D. and Whinston, A., 1997. A stochastic equilibrium model of Internet pricing. *Journal of Economic Dynamics and Control* 21, 697-722.

19. Gupta, A., Lindel, L., Stahl, D. and Whinston, A., 2001. Benefits and costs of adopting usage-based pricing in a subnetwork *Information Technology and Management* 2, 175-191

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

21. Jullien, B., 2000. Participation constraints in adverse selection problems. *Journal of Economic Theory* 93, 1-47.

22. Kephart, J., and Fay, S., 2000. Competitive bundling of categorized information goods. *Proceedings of the First ACM Conference on Electronic Commerce.*

23. Lin, L. and Whinston, A., 2001. Information technology contracts and hedging of risks for service providers. *Thirteenth Workshop on Information Systems and Economics.*

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

25. Maskin, E. and Riley, J., 1984. Monopoly with incomplete information. *Rand Journal of Economics* 15, 171-196.

26. Mendelson, H., 1985. Pricing computer services: queuing effects. *Communications of the ACM* 28, 312-321.

27. Mendelson, H. and Whang, S., 1990. Optimal incentive-compatible priority pricing for the M/M/1 queue. *Operations Research* 38, 870-883.

28. Susarla, A., Barua, A., and Whinston, A., 2001. A primer on ASP contracts. CREC Working Paper, University of Texas at Austin.

29. Varian, H. and Shapiro, C., 1998. *Information Rules.* Harvard Business School Press, Cambridge, MA.

30. Varian, H., 2000. Buying, selling and renting information goods. *Journal of Industrial Economics* XLVIII, 473-488.

31. Weber, T., 2001. Mixed versioning of information goods under incomplete information. *Proceedings of the 22nd International Conference on Information Systems.*

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

33. Wilson, R., 1993. *Nonlinear Pricing.* Oxford University Press.

A Appendix: Proofs

Proof of Lemma 1

The proof has four parts:

(a) Reduction of IC1 and IC2 to a simpler form: IC1 and IC2 are satisfied if θ is the solution to:

$$\max_{x \in [\underline{\theta}, \theta_U]} U(q(x, \theta_U), \theta) - \tau(x), \tag{44}$$

for all θ in $[\underline{\theta}, \theta_U]$. The necessary and sufficient conditions for (44) are:

$$U_1(q(\theta, \theta_U), \theta) q_1(\theta, \theta_U) - \tau'(\theta) = 0; \tag{45}$$

$$U_{11}(q(\theta, \theta_U), \theta)(q_1(\theta, \theta_U))^2 + U_1(q(\theta, \theta_U), \theta) q_{11}(\theta, \theta_U) - \tau''(\theta) \leq 0. \tag{46}$$

(45) is the first-order necessary condition, and (46) is the second-order sufficient condition. Differentiating (45) with respect to θ yields:

$$\begin{aligned} \tau''(\theta) = \; & U_{11}(q(\theta, \theta_U), \theta)(q_1(\theta, \theta_U))^2 + \\ & U_{12}(q(\theta, \theta_U), \theta) q_1(\theta, \theta_U) + U_1(q(\theta, \theta_U), \theta) q_{11}(\theta, \theta_U), \end{aligned}$$

which when substituted into (46) yields:

$$U_{12}(q(\theta, \theta_U), \theta) q_1(\theta, \theta_U) \geq 0. \tag{47}$$

The Spence-Mirrlees conditions ensure that $U_{12}(q(\theta, \theta_U), \theta) > 0$, which implies that IC1 and IC2 have reduced to:

$$\tau'(\theta) = U_1(q(\theta, \theta_U), \theta) q_1(\theta, \theta_U); \tag{48}$$

$$q_1(\theta, \theta_U) \geq 0 \tag{49}$$

Since $U_1(q(\theta, \theta_U), \theta) > 0$, the fact that $q_1(\theta, \theta_U) \geq 0$ implies that $\tau'(\theta) \geq 0$.

(b) Redefining the firm's objective function: Define the surplus of customer type θ as

$$s(\theta) = U(q(\theta, \theta_U), \theta) - \tau(\theta). \tag{50}$$

Differentiating with respect to θ, and substituting (48) yields:

$$s'(\theta) = U_2(q(\theta, \theta_U), \theta). \tag{51}$$

(51) and Property 2 of $U(q,\theta)$ imply that surplus is strictly increasing in type. Consequently, if IR is satisfied for the lowest type $\underline{\theta}$, it is satisfied for all others. Since the firm is profit maximizing, it will choose $s(\underline{\theta}) = 0$, which implies that:

$$s(\theta) = \int_{x=\underline{\theta}}^{\theta} U_2(q(x,\theta_U),x)dx. \tag{52}$$

$s(\theta)$ is referred to as the informational rent of type θ. Since $U_2(q(x,\theta_U),x) > 0$, $s(\theta)$ is strictly increasing if $q(\theta,\theta_U) > 0$. Now, (50) and (52) imply that

$$\tau(\theta) = U(q(\theta,\theta_U),\theta) - \int_{x=\underline{\theta}}^{\theta} U_2(q(x,\theta_U),x)dx. \tag{53}$$

Therefore, the firm's objective function, which is:

$$\max_{q(.),\tau(.)} \int_{\theta=\underline{\theta}}^{\theta_U} [\tau(\theta) - cq(\theta,\theta_U)]f(\theta)d\theta \tag{54}$$

can be rewritten as:

$$\max_{q(.)} \int_{\theta=\underline{\theta}}^{\theta_U} [U(q(\theta,\theta_U),\theta) - cq(\theta,\theta_U)]f(\theta)d\theta - \int_{\theta=\underline{\theta}}^{\theta_U} \left(\int_{x=\underline{\theta}}^{\theta} U_2(q(x,\theta_U),x)dx \right) f(\theta)d\theta. \tag{55}$$

Integrating the second part of (55) by parts (define $G(\theta) = \int_{x=\underline{\theta}}^{\theta} U_2(q(x,\theta_U),x)dx$, note that $F(\underline{\theta}) = 0$, and use $\int_{\theta=\underline{\theta}}^{\theta_U} G(\theta)f(\theta)d\theta = F(\theta_U)G(\theta_U) - \int_{\theta=\underline{\theta}}^{\theta_U} F(\theta)dG(\theta)$) yields:

$$\max_{q(.)} \int_{\theta=\underline{\theta}}^{\theta_U} [U(q(\theta,\theta_U),\theta) - cq(\theta,\theta_U)]f(\theta)d\theta - \int_{\theta=\underline{\theta}}^{\theta_U} U_2(q(\theta,\theta_U),\theta)[F(\theta_U) - F(\theta)]d\theta, \tag{56}$$

which can be rewritten as:

$$\max_{q(.)} \int_{\theta=\underline{\theta}}^{\theta_U} [U(q(\theta,\theta_U),\theta) - cq(\theta,\theta_U) - U_2(q(\theta,\theta_U),\theta)H(\theta,\theta_U)]f(\theta)d\theta, \tag{57}$$

where $H(\theta,\theta_U) = \frac{F(\theta_U)-F(\theta)}{f(\theta)}$.

(c) Unique solution to unconstrained problem: By (57) and (49), the firm's problem is now:

$$\max_{q(.)} \int_{\theta=\underline{\theta}}^{\theta_U} [U(q(\theta,\theta_U),\theta) - cq(\theta,\theta_U) - U_2(q(\theta,\theta_U),\theta)H(\theta,\theta_U)]f(\theta)d\theta \tag{58}$$

$$\text{subject to:} \quad q_1(\theta,\theta_U) \geq 0. \tag{59}$$

If the unconstrained version of this problem has a unique solution for which $q_1(\theta, \theta_U) \geq 0$, then this is the solution to the constrained problem as well.

The unconstrained problem can be solved by optimizing (58) pointwise to construct $q(., \theta_U)$. The first-order conditions for this problem are:

$$U_1(q(\theta, \theta_U), \theta) - c - U_{12}(q(\theta, \theta_U), \theta)H(\theta, \theta_U) = 0 \ \forall \theta, \tag{60}$$

which reduce to:

$$U_1(q(\theta, \theta_U), \theta) = c + U_{12}(q(\theta, \theta_U), \theta)\frac{F(\theta_U) - F(\theta)}{f(\theta)} \ \forall \theta. \tag{61}$$

The conditions (61) are sufficient if the function:

$$\pi(q, \theta, \theta_U) = U(q, \theta) - cq - U_2(q, \theta)H(\theta, \theta_U) \tag{62}$$

is strictly quasiconcave in q. Differentiating (62) with respect to q yields:

$$\pi_1(q, \theta, \theta_U) = U_1(q, \theta) - c - U_{12}(q, \theta)H(\theta, \theta_U). \tag{63}$$

Consequently, if $\pi_1(q, \theta, \theta_U) = 0$, (63) implies that:

$$H(\theta, \theta_U) = \frac{U_1(q, \theta) - c}{U_{12}(q, \theta)}. \tag{64}$$

Also, Property 5 of $U(q, \theta)$ assumes that:

$$\frac{\partial}{\partial \theta}\left(\frac{-U_{11}(q, \theta)}{U_1(q, \theta)}\right) \leq 0, \tag{65}$$

which can be expanded to:

$$\frac{-U_{112}(q, \theta)}{U_1(q, \theta)} - \frac{-U_{11}(q, \theta)U_{12}(q, \theta)}{(U_1(q, \theta))^2} \leq 0, \tag{66}$$

or

$$U_{112}(q, \theta) \geq \frac{U_{11}(q, \theta)U_{12}(q, \theta)}{U_1(q, \theta)}. \tag{67}$$

Now, differentiating (63) with respect to q yields:

$$\pi_{11}(q, \theta, \theta_U) = U_{11}(q, \theta) - U_{112}(q, \theta)H(\theta, \theta_U). \tag{68}$$

Therefore, if $\pi_1(q,\theta,\theta_U)=0$, (64) and (68) imply that

$$\pi_{11}(q,\theta,\theta_U)=U_{11}(q,\theta)-U_{112}(q,\theta)\frac{U_1(q,\theta)-c}{U_{12}(q,\theta)}, \tag{69}$$

which when combined with (67) yields

$$\pi_{11}(q,\theta,\theta_U)\leq U_{11}(q,\theta)-\frac{U_{11}(q,\theta)U_{12}(q,\theta)}{U_1(q,\theta)}\frac{U_1(q,\theta)-c}{U_{12}(q,\theta)}, \tag{70}$$

which simplifies to:

$$\pi_{11}(q,\theta,\theta_U)\leq U_{11}(q,\theta)\frac{c}{U_1(q,\theta)}. \tag{71}$$

Since $U(q,\theta)$ is strictly concave in q, the RHS of (71) is strictly negative. Consequently, if $\pi_1(q,\theta,\theta_U)=0$, then $\pi_{11}(q,\theta,\theta_U)<0$, which establishes that $\pi(q,\theta,\theta_U)$ is strictly quasi-concave in q, which in turn ensures that for the unconstrained problem of (58), first-order conditions (61) yield the unique solution.

(d) Monotonicity of $\mathbf{q}(\boldsymbol{\theta},\boldsymbol{\theta}_U)$ in $\boldsymbol{\theta}$: Assume that $U_{122}(q(\theta,\theta_U),\theta)\leq 0$. Differentiating both sides of (61) yields:

$$\begin{aligned}
&U_{11}(q(\theta,\theta_U),\theta)q_1(\theta,\theta_U)+U_{12}(q(\theta,\theta_U),\theta)= \\
&U_{112}(q(\theta,\theta_U),\theta)q_1(\theta,\theta_U)H(\theta,\theta_U)+U_{122}(q(\theta,\theta_U),\theta)H(\theta,\theta_U) \\
&+U_{12}(q(\theta,\theta_U),\theta)H_1(\theta,\theta_U),
\end{aligned} \tag{72}$$

which implies that:

$$q_1(\theta,\theta_U)=\frac{U_{12}(q(\theta,\theta_U),\theta)[1-H_1(\theta,\theta_U)]-U_{122}(q(\theta,\theta_U),\theta)H(\theta,\theta_U)}{U_{112}(q(\theta,\theta_U),\theta)H(\theta,\theta_U)-U_{11}(q(\theta,\theta_U),\theta)}. \tag{73}$$

From (68) and the fact that $\pi(q,\theta,\theta_U)$ has been shown to be strictly quasiconcave, we know that the denominator of (73) is strictly positive. Also, we know that $H_1(\theta,\theta_U)\leq 0$, since the reciprocal of the hazard rate has been assumed to be non-increasing. Since $H(\theta,\theta_U)>0$ for all θ in the interior of $[\underline{\theta},\theta_U]$, the numerator of (73) is strictly positive (since we have assumed that $U_{122}(q(\theta,\theta_U),\theta)\leq 0$ for this part) which implies that $q_1(\theta,\theta_U)>0$.

Since we know that for the optimal $q^*(.,\theta_U)$, necessary condition (61) has to hold for all $q^*(\theta,\theta_U)>0$, this establishes that $q^*(\theta,\theta_U)$ is strictly increasing in θ when it is positive. This also means that if $U_{122}(q(\theta,\theta_U),\theta)\leq 0$, then the first order conditions (58) define the unique optimal contract $q^*(\theta,\theta_U)$.

Proof of Proposition 1

The first-order conditions in Lemma 1 are:

$$U_1(q^*(\theta,\theta_U),\theta) - c - U_{12}(q^*(\theta,\theta_U),\theta)\tfrac{F(\theta_U)-F(\theta)}{f(\theta)} = 0 \ \forall\theta. \tag{74}$$

Differentiating (74) with respect to θ_U yields:

$$q_2^*(\theta,\theta_U)[U_{11}(q(\theta,\theta_U),\theta) - U_{112}(q(\theta,\theta_U),\theta)\tfrac{F(\theta_U)-F(\theta)}{f(\theta)}] = U_{12}(q^*(\theta,\theta_U),\theta)\frac{f(\theta_U)}{f(\theta)}, \tag{75}$$

which implies that

$$q_2^*(\theta,\theta_U) = \frac{f(\theta_U)U_{12}(q^*(\theta,\theta_U),\theta)}{f(\theta)[U_{11}(q^*(\theta,\theta_U),\theta) - U_{112}(q^*(\theta,\theta_U),\theta)H(\theta,\theta_U)]}, \tag{76}$$

where $H(\theta,\theta_U) = \frac{F(\theta_U)-F(\theta)}{f(\theta)}$. Since (74) is true for all θ, it follows that

$$H(\theta,\theta_U) = \frac{U_1(q^*(\theta,\theta_U),\theta) - c}{U_{12}(q^*(\theta,\theta_U),\theta)}. \tag{77}$$

Also, from (67), we know that

$$U_{112}(q^*(\theta,\theta_U),\theta) \geq \frac{U_{11}(q^*(\theta,\theta_U),\theta)U_{12}(q^*(\theta,\theta_U),\theta)}{U_1(q^*(\theta,\theta_U),\theta)}. \tag{78}$$

(77) and (78) can be combined to show that

$$U_{11}(q^*(\theta,\theta_U),\theta) - U_{112}(q^*(\theta,\theta_U),\theta)H(\theta,\theta_U) \leq \frac{cU_{11}(q^*(\theta,\theta_U),\theta)}{U_1(q^*(\theta,\theta_U),\theta)}. \tag{79}$$

Since $U_1(q^*(\theta,\theta_U),\theta) > 0$, $U_{11} < 0$, $U_{12} > 0$ and $f(\theta)$ is strictly positive, (76) and (79) imply that $q_2^*(\theta,\theta_U) < 0$, which proves part (a) of the proposition. Finally,

$$\frac{\partial}{\partial\theta_U}\int_{x=\underline{\theta}}^{\theta} U_2(q^*(x,\theta_U),x)dx = \int_{x=\underline{\theta}}^{\theta} q_2^*(x,\theta_U)U_{12}(q^*(x,\theta_U),x)dx,$$

and since $U_{12} > 0$, this is strictly negative when $q_2^*(\theta,\theta_U) < 0$, which proves part (b).

Proof of Lemma 2

Define $\psi(\theta) = v(\theta) - U(q(\theta),\theta) + \tau(\theta)$. Differentiating with respect to θ :

$$\psi'(\theta) = v'(\theta) - U_1(q(\theta),\theta)q'(\theta) - U_2(q(\theta),\theta) + \tau'(\theta). \tag{80}$$

Since $v(\theta) = \max_q U(q, \theta)$, we can apply the envelope theorem to show that:

$$v'(\theta) = U_2(\alpha(\theta), \theta). \tag{81}$$

Also, incentive compatibility of $(q(.), \tau(.))$ implies that

$$U_1(q(\theta), \theta)q'(\theta) = \tau'(\theta). \tag{82}$$

Substituting (81)and (82) into (80) yields:

$$\psi'(\theta) = U_2(\alpha(\theta), \theta) - U_2(q(\theta), \theta). \tag{83}$$

Since $\alpha(\theta) \geq q(\theta)$ for all θ, the Spence-Mirrlees condition $U_{12}(q, \theta) > 0$ ensures that $\psi'(\theta) \geq 0$, which proves part (a). Also, if $\alpha(\hat{\theta}) > q(\hat{\theta})$ for some $\hat{\theta}$, then $\psi'(\hat{\theta}) > 0$, which proves part (b). Finally, if $\alpha(\theta) > q(\theta)$ for all θ, then $\psi'(\theta) > 0$ for all θ, implying that $\psi(\theta)$ is strictly increasing, which completes the proof.

Proof of Proposition 4

The proof has three parts.

(a) Necessary conditions for optimal usage-based contract in a fixed sub-interval, with a fixed T

Suppose the firm has in place a fixed-fee contract T, which it wants only (and all) customers with $\theta \geq \theta_F$ to adopt, and it wants to find the optimal usage-based contract for the customer interval $[\underline{\theta}, \theta_F]$, given this fixed fee T. The objective function for this problem was derived as (57) in Lemma 1. Also, so long as the usage-based contract is incentive compatible and does not assign a quantity $\alpha(\theta)$ for any θ in $[\underline{\theta}, \theta_F]$, Lemma 2 implies that constraining customers of type θ_F to be indifferent between the fixed-fee contract T and the usage-based contract.$(q(., \theta_U, T), \tau(., \theta_U, T))$ will ensure that all types $\theta < \theta_F$ will choose the usage-based contract, and all types $\theta \geq \theta_F$ will choose the fixed fee. The third argument in q and τ, introduced just for the purpose of this part of the proof, indicates the dependence of the optimal usage-based contract on not just θ_U, but also the level of the fixed-fee contract T.

We know from Lemma 1 that the surplus that a customer of type θ_F will get from this usage-based contract is

$$U(q(\theta_F,\theta_F,T),\theta_F) - \tau(\theta_F,\theta_F,T) = \int_{\underline{\theta}}^{\theta_F} U_2(q(\theta,\theta_F,T),\theta)d\theta \tag{84}$$

Consequently, the firm's problem is:

$$\Upsilon(\theta_F,T) = \max_{q(.,\theta_U,T)} \int_{\underline{\theta}}^{\theta_F} [U(q(\theta,\theta_F,T),\theta) - cq(\theta,\theta_F,T) - U_2(q(\theta,\theta_F,T),\theta)\tfrac{F(\theta_F)-F(\theta)}{f(\theta)}]f(\theta)d\theta, \tag{85}$$

subject to the constraint:

$$\int_{\underline{\theta}}^{\theta_F} U_2(q(\theta,\theta_F,T),\theta)d\theta - [v(\theta_F) - T] = 0. \tag{86}$$

Denote the Lagrangian for this problem as

$$\begin{aligned} L(q,\theta_F,T) &= \int_{\underline{\theta}}^{\theta_F} \left(U(q(\theta,\theta_F,T),\theta) - cq(\theta,\theta_F,T) - U_2(q(\theta,\theta_F,T),\theta)\tfrac{F(\theta_F)-F(\theta)}{f(\theta)}\right) f(\theta)d\theta \\ &\quad +\lambda(T)\left[\int_{\underline{\theta}}^{\theta_F} U_2(q(\theta,\theta_F,T),\theta)d\theta - v(\theta_F) + T\right]. \end{aligned} \tag{87}$$

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

$$\left[\frac{\partial L}{\partial q} = 0\right] : \left(U_1(q^*(\theta,\theta_F,T),\theta) - c - U_{12}(q^*(\theta,\theta_F,T),\theta)[\tfrac{F(\theta_F)-F(\theta)-\lambda(T)}{f(\theta)}]\right) f(\theta) = 0 \tag{88}$$

for all $\theta \leq \theta_F$, and:

$$\left[\frac{\partial L}{\partial \lambda} = 0\right] : \int_{\underline{\theta}}^{\theta_F} U_2(q^*(\theta,\theta_F,T),\theta)d\theta = v(\theta_F) - T. \tag{89}$$

(b) Optimal fixed-fee and usage-based contract in a fixed sub-interval

Now consider the problem of simultaneously choosing the optimal value of T and the optimal usage-based contract for $[\underline{\theta},\theta_F]$, subject to the requirement that every customer with $\theta \geq \theta_F$ prefers T, and that every customer in $[\underline{\theta},\theta_F)$ prefers the usage-based contract. Suppose the firm determines the function $\Upsilon(\theta_F,T)$ as specified in (85) by solving the problem above for each feasible T, where feasibility implies the values of T for which it is feasible to design an incentive-compatible usage-based contract for $[\underline{\theta}, \theta_F]$. Subsequently, if it chooses the value of T – and the corresponding optimal usage-based contract $(q^*(., \theta_U, T), \tau^*(., \theta_U, T))$ – that maximizes the sum of $\Upsilon(\theta_F, T)$ and the profits from the fixed fee T, which are $T[1 - F(\theta_F)]$ – then it has its solution. Consequently, this problem can be formulated as

$$\max_T \Upsilon(\theta_F, T) + T[1 - F(\theta_F)], \tag{90}$$

where $\Upsilon(\theta_F, T)$ is the value function for the problem specified in (86). The first order conditions for this problem are:

$$\Upsilon_2(\theta_F, T^*) = -[1 - F(\theta_F)] \tag{91}$$

To determine $\Upsilon_2(\theta_F, T)$, one can differentiate (85) with respect to T:

$$\Upsilon_2(\theta_F, T) = \int_{\underline{\theta}}^{\theta_F} q_3^*(\theta, \theta_F, T)\left(U_1(q^*(\theta, \theta_F, T), \theta) - c - U_{12}(q^*(\theta, \theta_F, T), \theta)[\tfrac{F(\theta_F)-F(\theta)}{f(\theta)}]f(\theta)\right) d\theta. \tag{92}$$

While (88) is not a sufficient condition, it is necessary, and therefore holds for the optimal usage-based contract. Since it is true for every θ in $[\underline{\theta}, \theta_F]$, we can substitute it into (92) to get:

$$\Upsilon_2(\theta_F, T) = -\lambda(T) \int_{\underline{\theta}}^{\theta_F} q_3^*(\theta, \theta_F, T) U_{12}(q^*(\theta, \theta_F, T), \theta) d\theta. \tag{93}$$

Also, the optimal usage-based contract for any feasible fixed T satisfies (89). Differentiating both sides of (89) with respect to T yields:

$$\int_{\underline{\theta}}^{\theta_F} q_3^*(\theta, \theta_F, T) U_{12}(q^*(\theta, \theta_F, T), \theta) d\theta = -1. \tag{94}$$

Combining (93) and (94) gives us the following value for $\Upsilon_2(\theta_F, T)$:

$$\Upsilon_2(\theta_F, T) = \lambda(T). \tag{95}$$

(95) is a familiar constraint-relaxation result — that the marginal value of relaxing a constraint is equal to the Lagrangian multiplier for that constraint. Now, substituting (95) into

(91) gives us the value of $\lambda(T)$ in the problem specified by (85) and (86), for the optimal fixed-fee contract T^*:

$$\lambda(T^*) = -[1 - F(\theta_F)]. \tag{96}$$

Since we now have the value of $\lambda(T^*)$, we can characterize the optimal usage-based contract which corresponds to this optimal fixed-fee T^*, by substituting (96) into (88):

$$\left(U_1(q^*(\theta,\theta_F,T^*),\theta) - c - U_{12}(q^*(\theta,\theta_F,T),\theta)[\frac{F(\theta_F) - F(\theta) - (-[1 - F(\theta_F)])}{f(\theta)}]\right) f(\theta) = 0, \tag{97}$$

which simplifies to

$$U_1(q^*(\theta,\theta_F,T^*),\theta) = c + U_{12}(q^*(\theta,\theta_F,T^*),\theta)[\frac{1 - F(\theta)}{f(\theta)}] \; \forall \theta \leq \theta_F. \tag{98}$$

(c) Optimal contracts for entire interval

Equation (98) implies that the usage-based quantity assigned to each type θ in $[\underline{\theta}, \theta_F]$ is independent of θ_F. Also, from Lemma 1, (98) is identical to the specification of the optimal usage-based contract for the *entire interval* $[\underline{\theta}, \overline{\theta}]$. Propositions 2 and 3 ensures that so long as $c > 0$, if the optimal combination of a usage-based contract T^* and fixed-fee contract $(q^*(.), \tau^*(.))$ for the entire interval $[\underline{\theta}, \overline{\theta}]$ involves any customer choosing the usage-based contract, it will result in some sub-interval of the customers $[\theta_F^*, \overline{\theta}]$ choosing the fixed-fee contract, and the others choosing the usage-based contract. However, this global optimum also has to be an optimal solution for the problem under which customers in the pre-specified interval $[\underline{\theta}, \theta_F^*)$ choose the usage-based contract – in other words, it has to be a solution to the problem in part (b), where $\theta_F = \theta_F^*$. As a consequence, it has to satisfy the necessary condition (98). This proves that the optimal usage-based contract is independent of the fixed fee. Part (a) immediately follows immediately from Lemma 1.

The value of the fixed-fee at which type θ_F is indifferent is $v(\theta_F) - U(q^*(\theta_F), \theta_F) + \tau^*(\theta_F)$. Part (b) follows from the fact that firm will choose the profit maximizing value of θ_F^*, and that the corresponding optimal usage-based contract $(q^*(\theta), \tau^*(\theta))$ is independent of the choice of θ_F^*. Part (c) simply computes the fixed-fee T^* at this optimal value of θ_F^*. This completes the proof.

Table 1: **Summary of Notation introduced in Sections 2 and 3**

Symbol	Explanation
$U(q,\theta)$	Utility that customer type θ gets from usage level q.
$v(\theta)$	Maximum utility that customer type θ can get from usage. $v(\theta) = \max_q U(q,\theta)$.
$\alpha(\theta)$	Usage level at which utility for type θ is maximized. $\alpha(\theta) = \arg\max_q U(q,\theta)$.
$[\underline{\theta},\overline{\theta}]$	Range of possible customer types θ.
$f(\theta), F(\theta)$	Density and distribution functions of customer type distribution.
$q(\theta), \tau(\theta)$	Usage-based contract (continuous set of quantity-price pairs) that is incentive compatible for $[\underline{\theta},\overline{\theta}]$.For a specific value of θ, $q(\theta)$ is the quantity and $\tau(\theta)$ is the price for that quantity.
$q^*(\theta), \tau^*(\theta)$	Optimal usage-based contract that is incentive compatible for $[\underline{\theta},\overline{\theta}]$.
$q(\theta,\theta_U), \tau(\theta,\theta_U)$	Usage-based contract that is incentive compatible for $[\underline{\theta},\theta_U]$.
$q^*(\theta,\theta_U), \tau^*(\theta,\theta_U)$	Optimal usage-based contract that is incentive compatible for $[\underline{\theta},\theta_U]$.
T	Unlimited-usage fixed fee price.
c	Unit cost of contract administration for a usage-based contract.
θ_F	Lowest customer type which is indifferent between a specified fixed-fee and usage-based contract
θ_H	Lowest customer type for whom $T = \tau(\theta)$.

Table 2: Summary of Notation introduced in Section 4

Symbol	Explanation
β	Parameter proportional to either marginal utility from usage, or mean of the customer type distribution, or both.
$\theta_L(\beta, c)$	Lowest type who adopts optimal usage-based contract, for parameter values β and c.
$\theta_F^*(\beta, c)$	Lowest type who adopts optimal fixed-fee contract, for parameter values β and c.
$T^*(\beta, c)$	Optimal fixed-fee contract, for parameter values β and c
$\Pi_U(\beta, c)$	Maximum firm profits from offering only a usage-based contract, for parameter values β and c.
$\Pi_F(\beta)$	Maximum firm profits from offering only a fixed-fee contract, for parameter values β and c.
$\Pi(\beta, c)$	Firm profits from offering the optimal combination of fixed-fee contract and usage-based contract, for parameter values β and c.
$C(\beta, c)$	Total customer surplus when firm offers the optimal combination of fixed-fee contract and usage-based contract, for parameter values β and c.
$S(\beta, c)$	Total surplus when firm offers the optimal combination of fixed-fee contract and usage-based contract, for parameter values β and c.

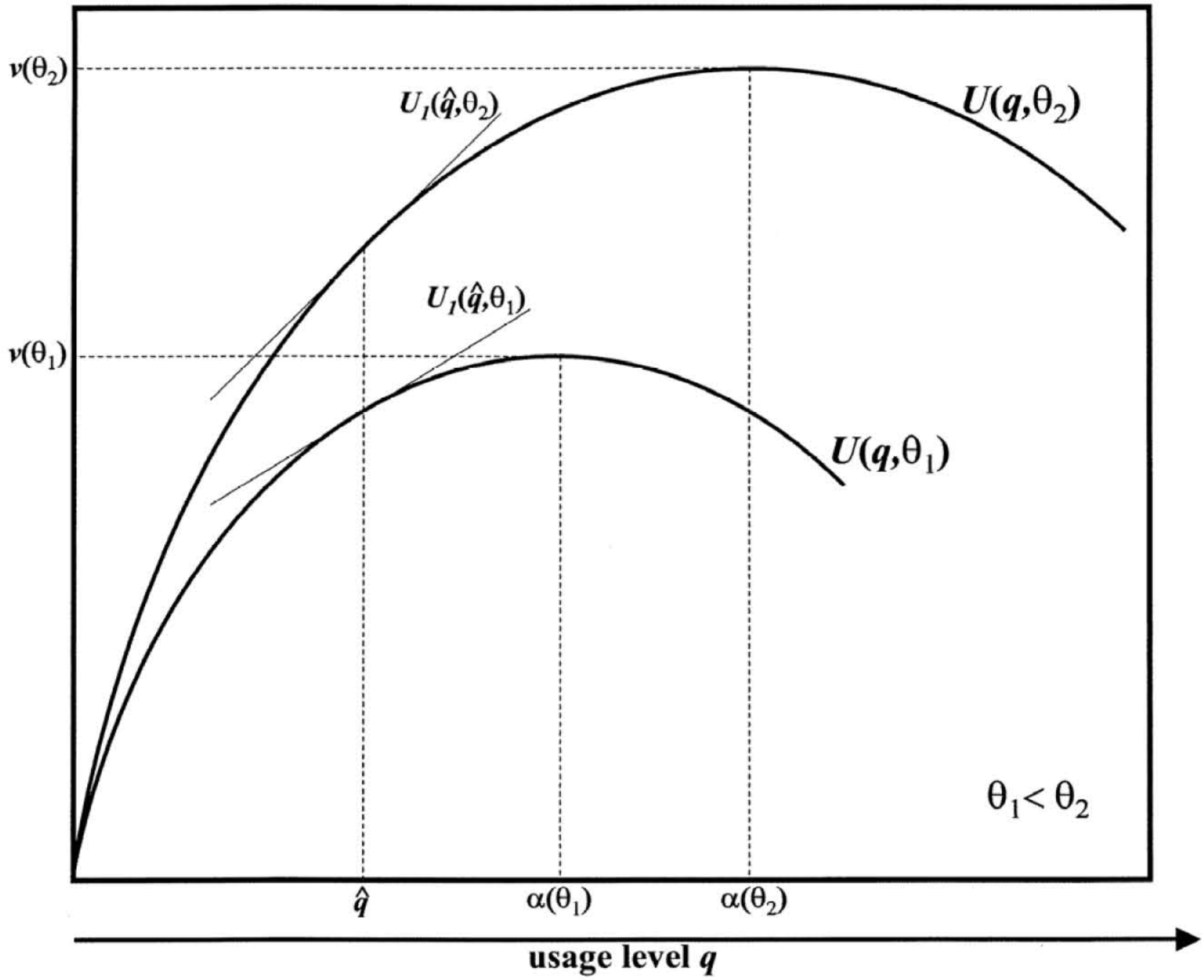


Figure 1: Illustration of $U(q\ ,\theta)$ for two different values of θ

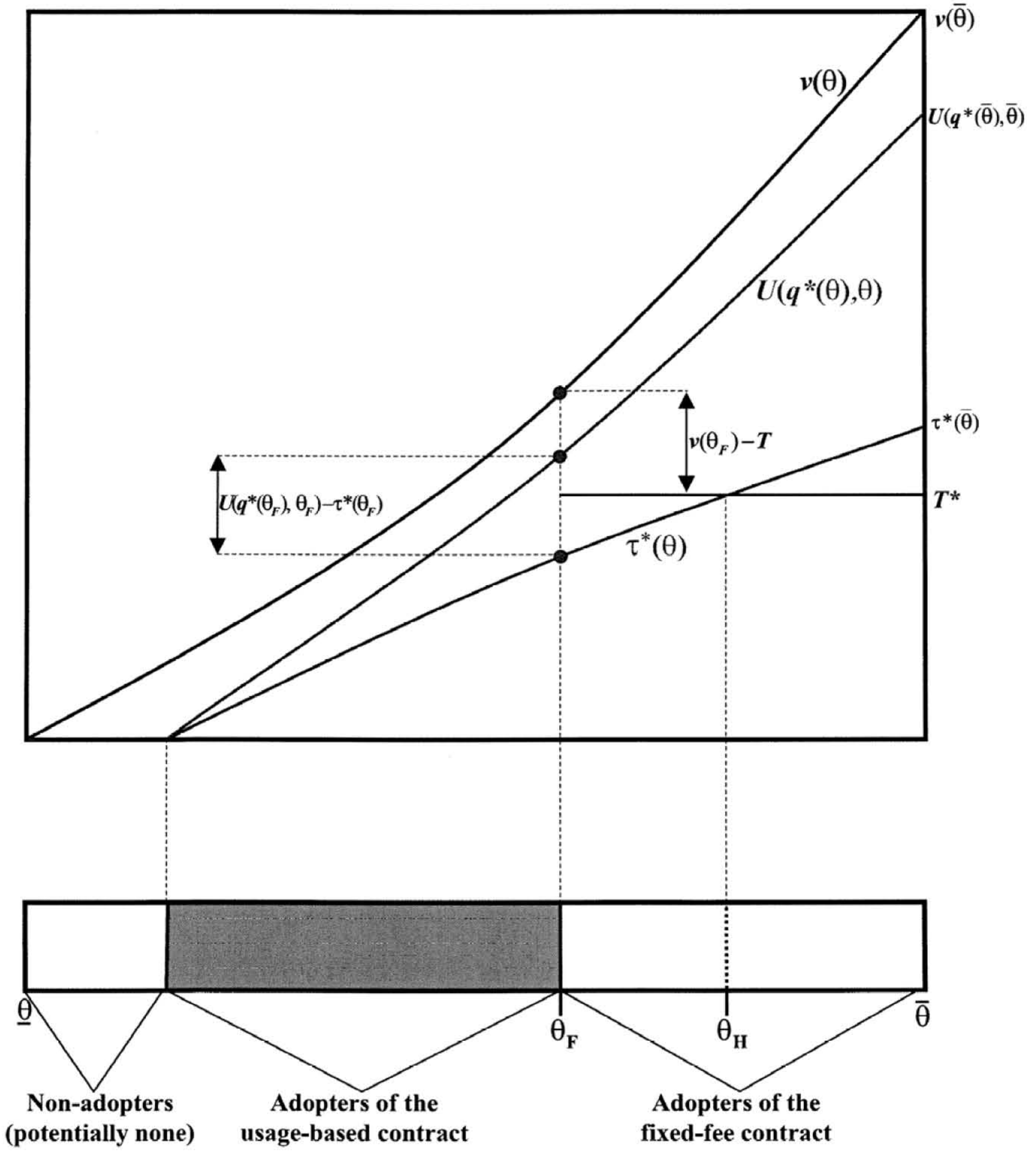


Figure 2: The impact of a fixed-fee contract on user choice

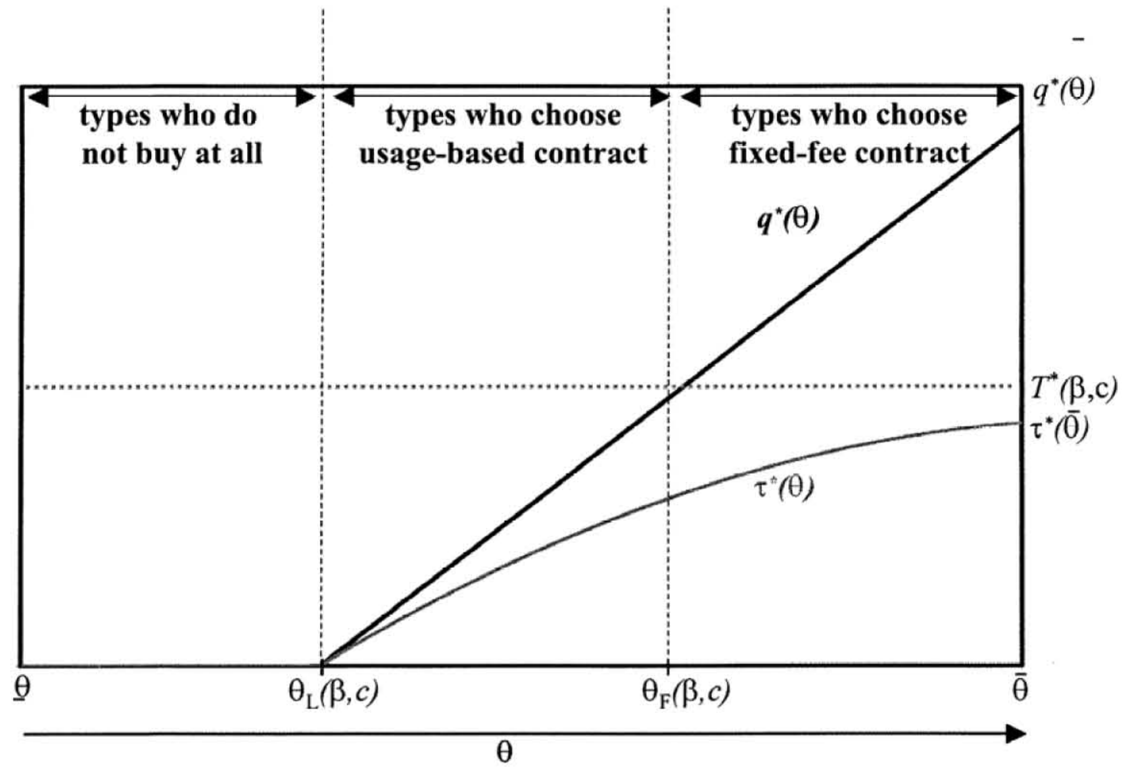


Figure 3a: Optimal Contract $q^*(\cdot),\tau^*(\cdot)$

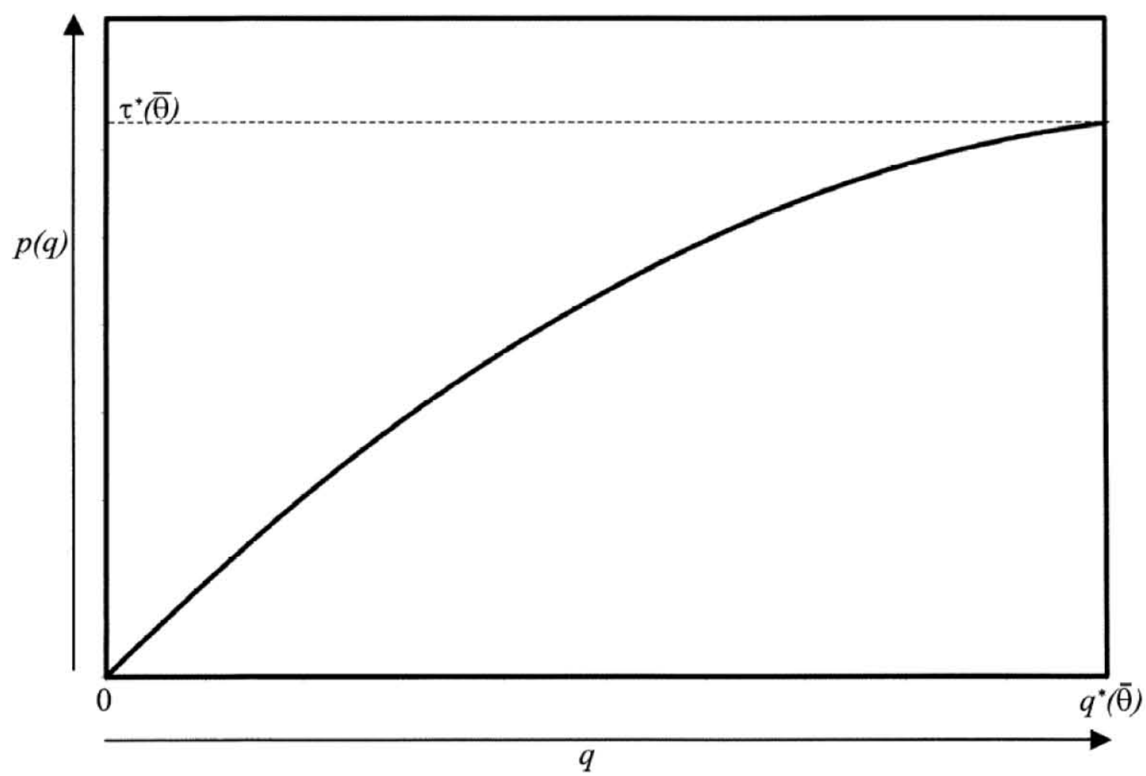


Figure 3b: Price $p(q)$ as a function of usage

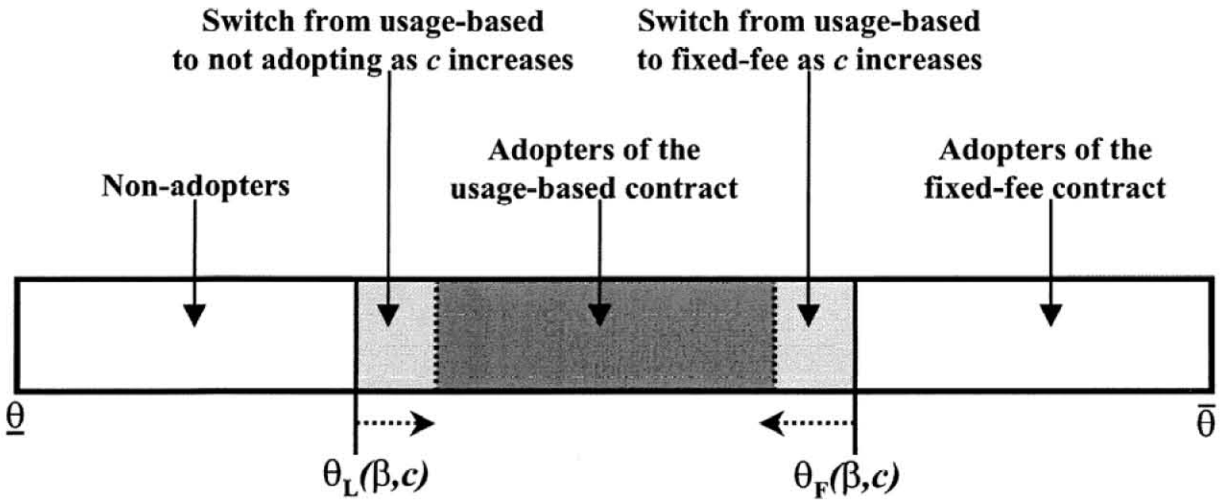


Figure 4a: Impact of increasing c on customer choice

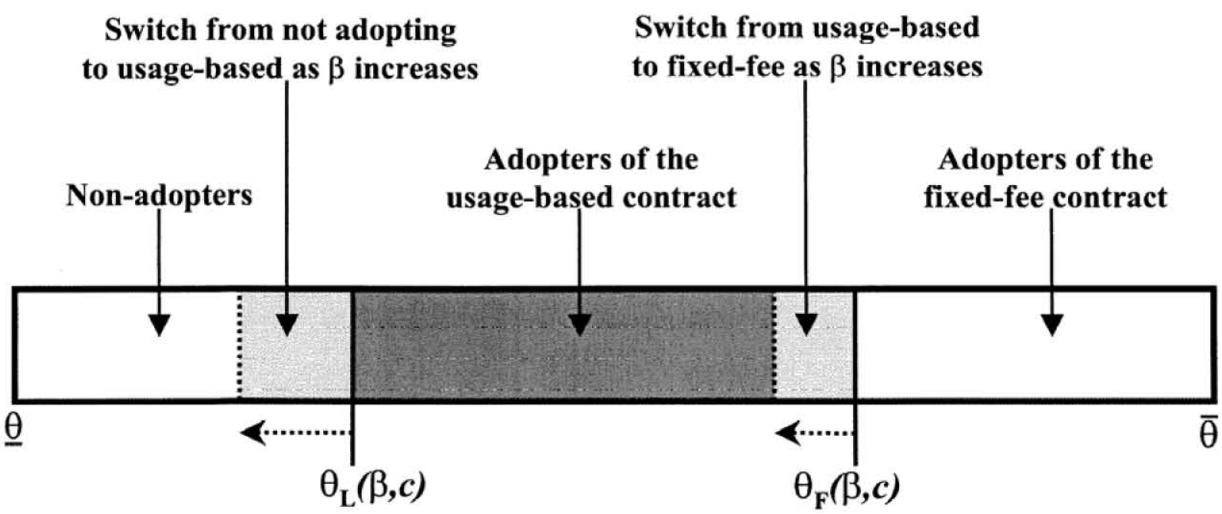


Figure 4b: Impact of increasing β on customer choice

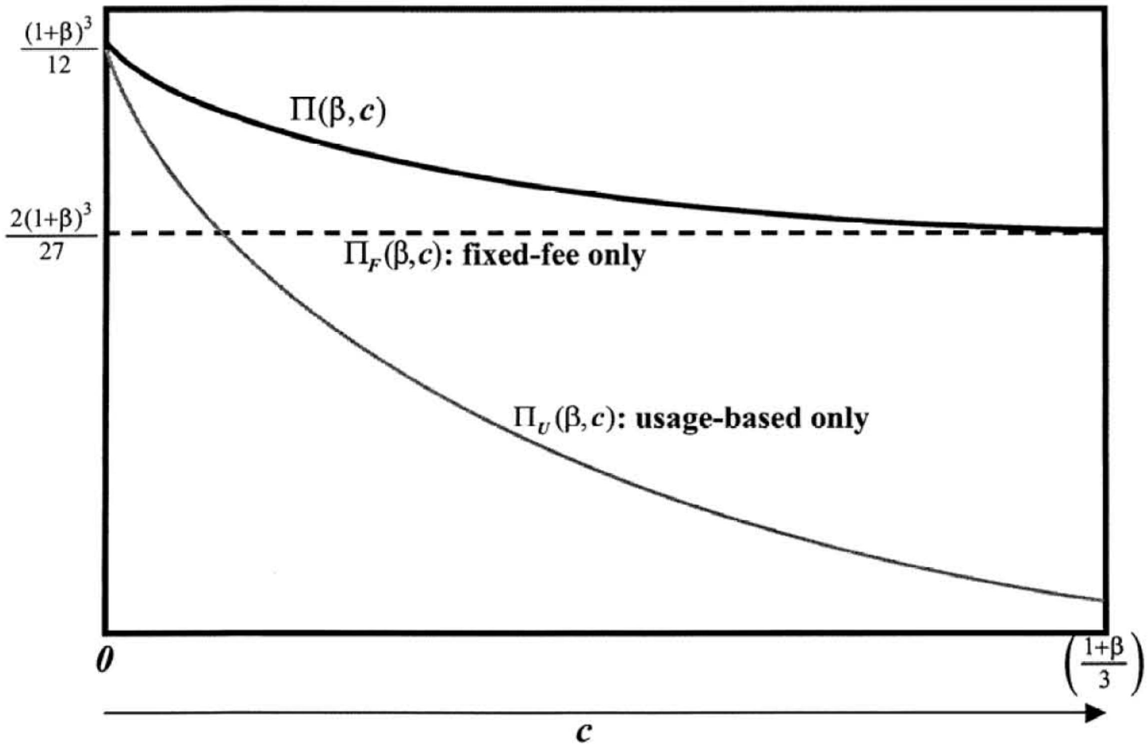


Figure 5a: Firm profits - fixed-fee, usage based and combination

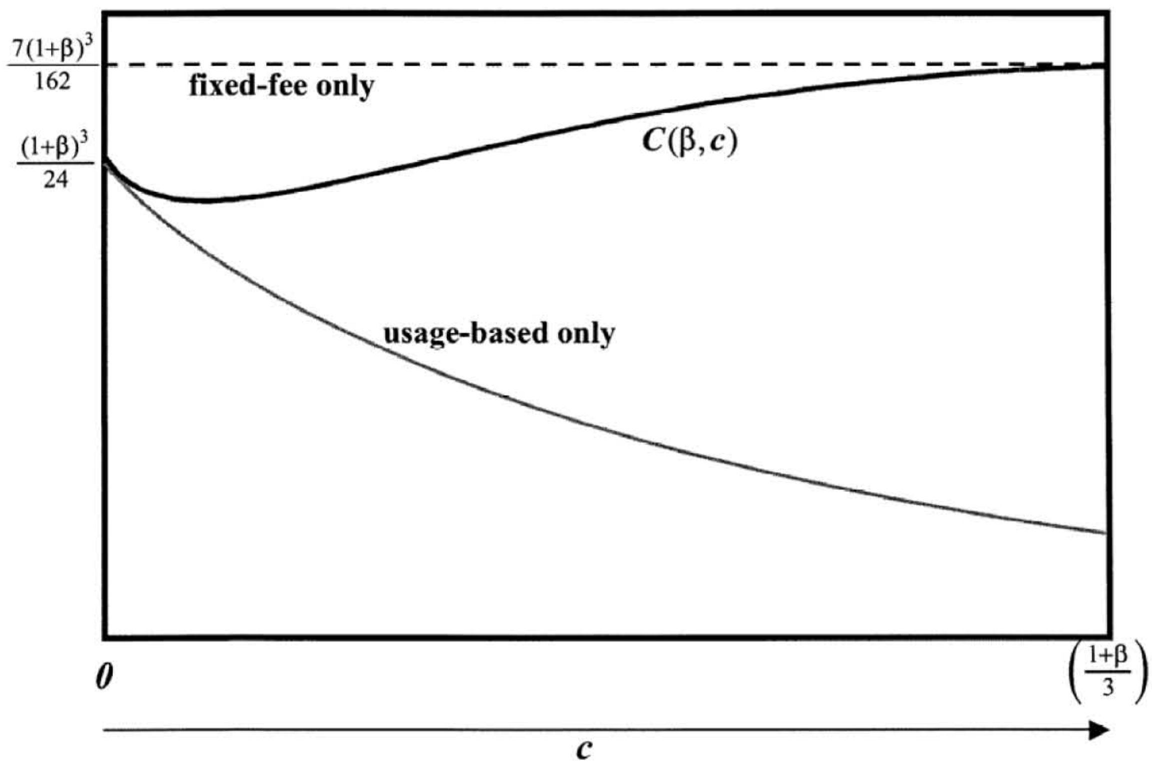


Figure 5b: Customer surplus - fixed-fee, usage based and combination

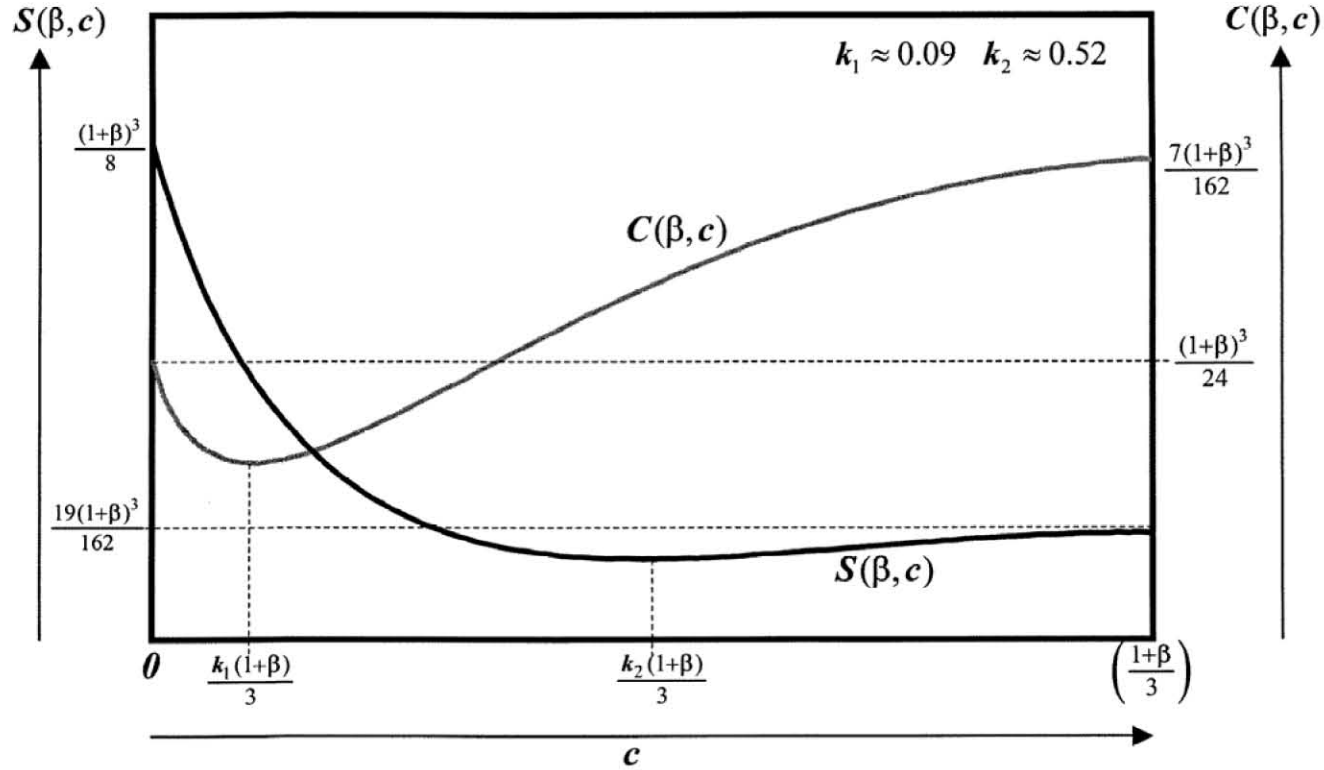


Figure 5c: Customer and total surplus as *c* varies

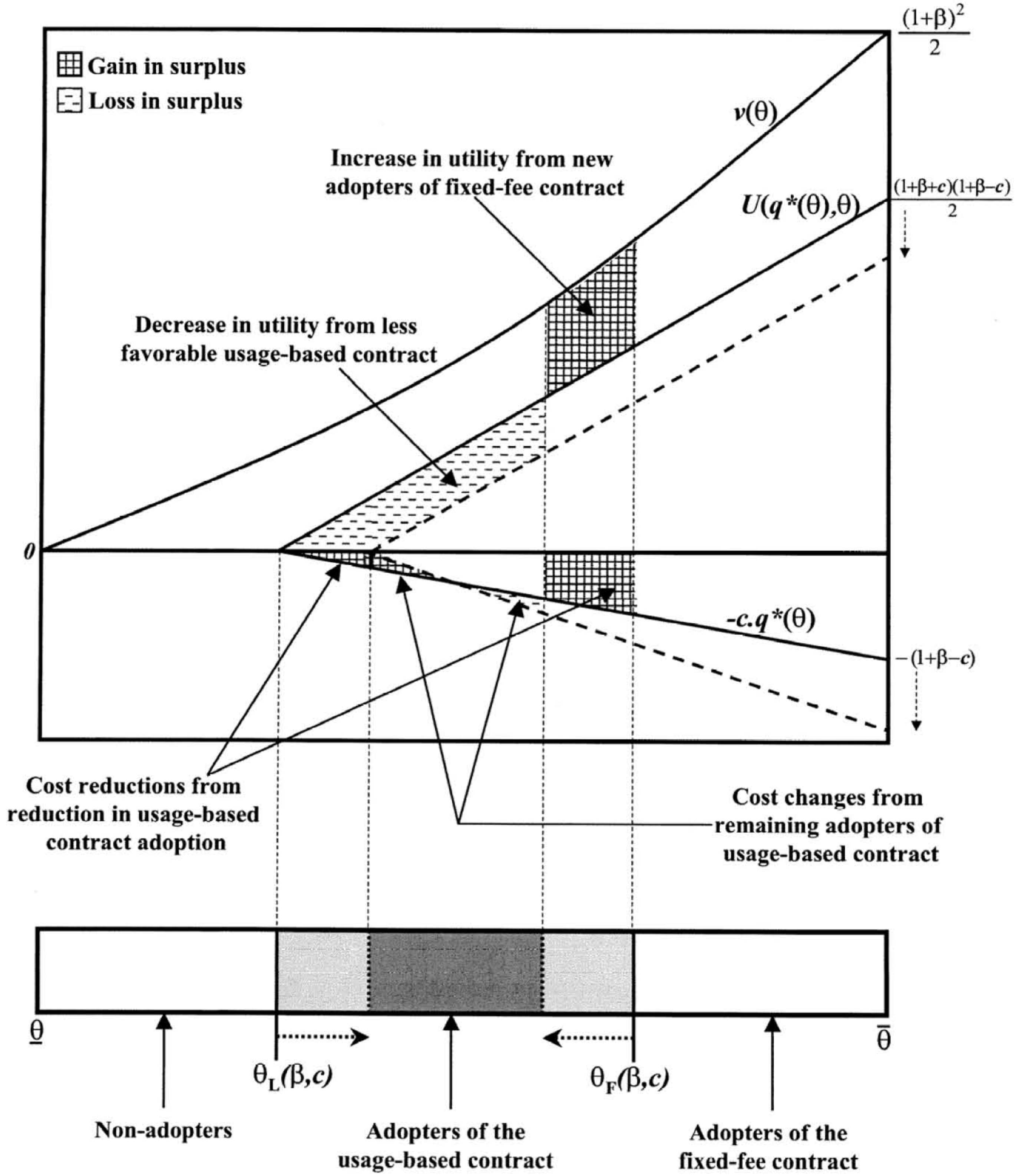
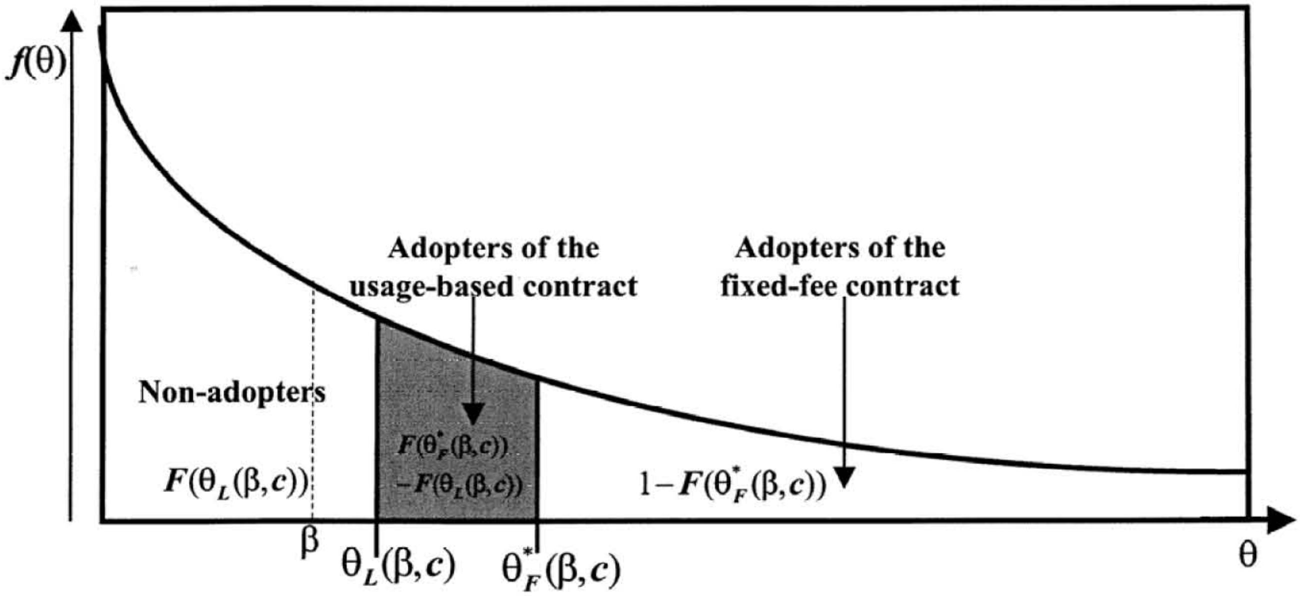
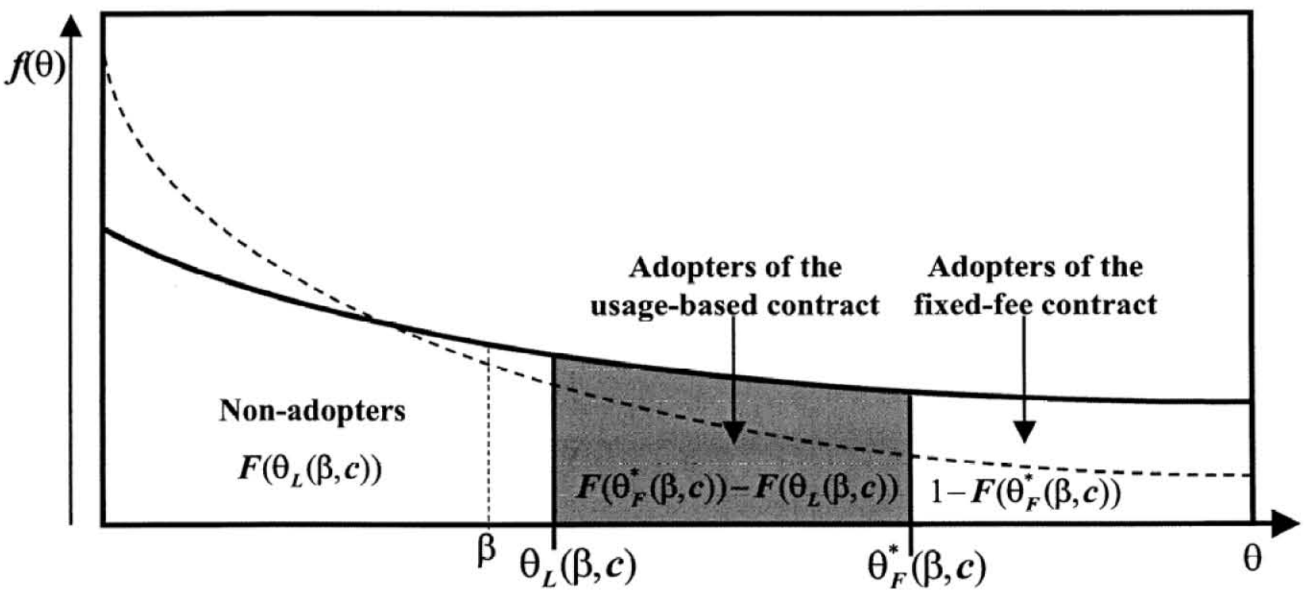


Figure 6: Changes in total surplus as *c* varies: a closer look



Lower value of β



Higher value of β

Figure 7: Impact of increasing β on customer choice

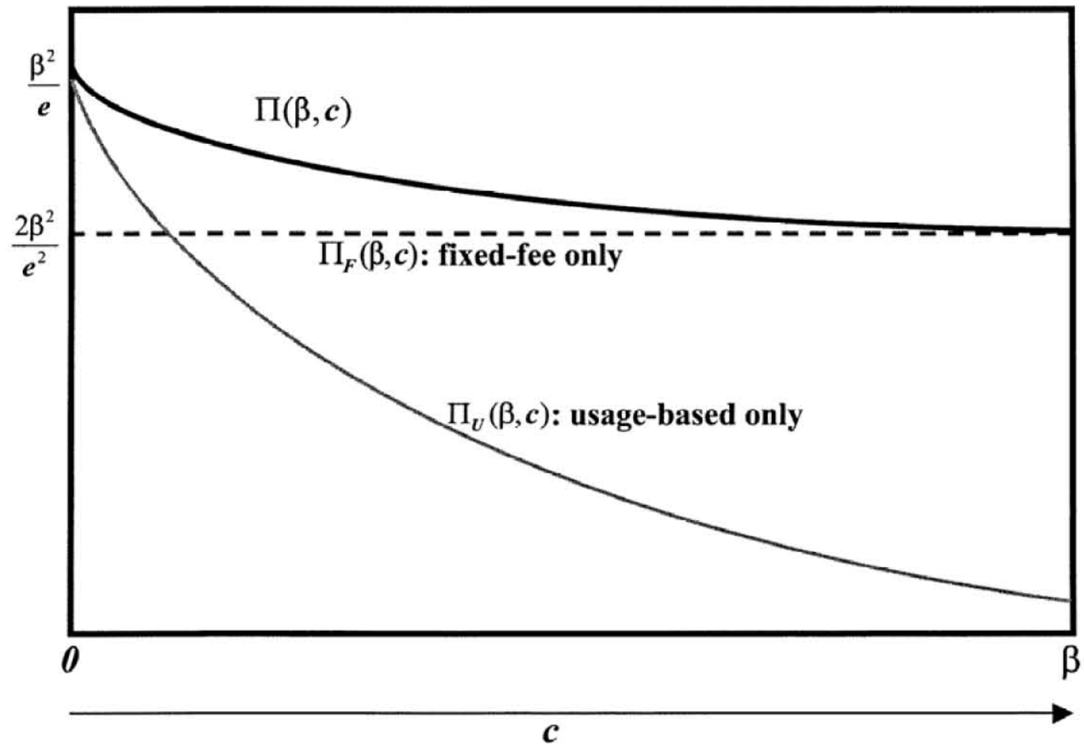


Figure 8a: Firm profits - fixed-fee, usage based and combination

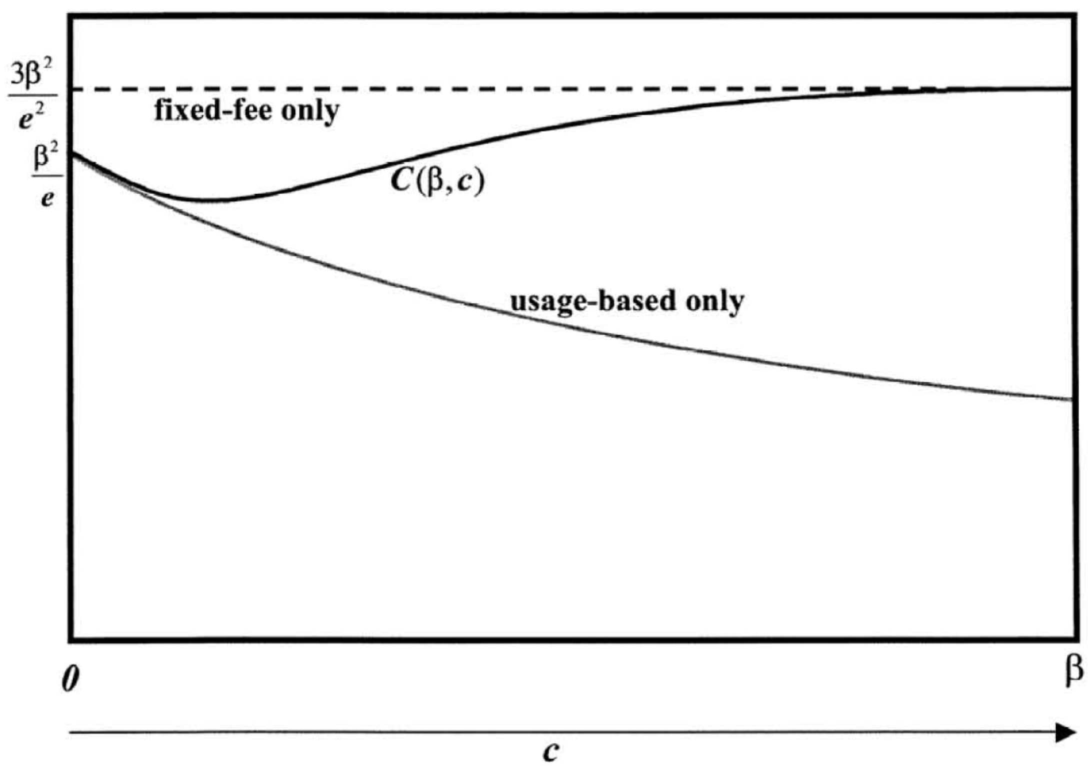


Figure 8b: Customer surplus - fixed-fee, usage based and combination

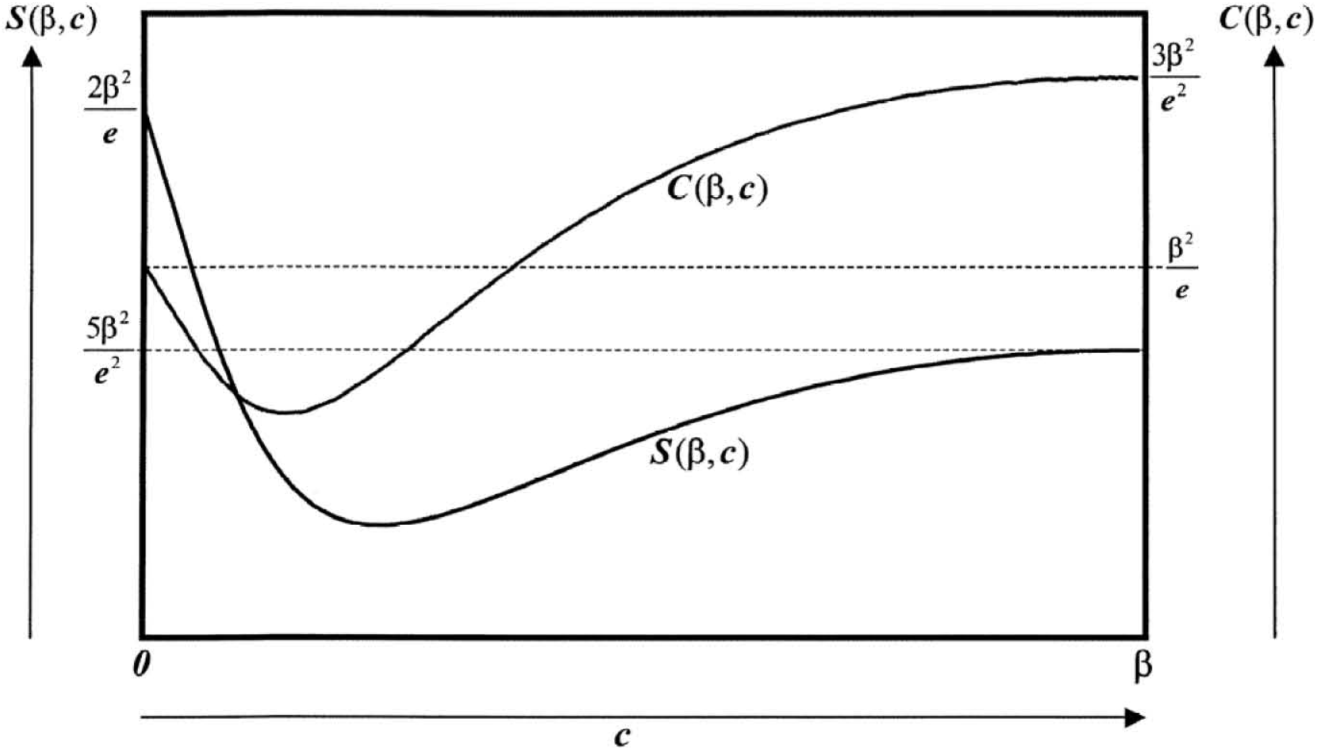


Figure 8c: Customer and total surplus as *c* varies