5-2010

Usage-based pricing of software services under competition

Ram Bala
Santa Clara University, rbala@scu.edu

Scott Carr

Recommended Citation

Bala, Ram, and Scott Carr. "Usage-based Pricing of Software Services under Competition." Journal of Revenue and Pricing Management J Revenue Pricing Manag 9.3 (2010): 204-16.

This is a post-peer-review, pre-copyedit version of an article published in Journal of Revenue and Pricing Management. The definitive publisher-authenticated version: Bala, Ram, and Scott Carr. "Usage-based Pricing of Software Services under Competition." Journal of Revenue and Pricing Management J Revenue Pricing Manag 9.3 (2010): 204-16. is available online at doi:10.1057/rpm.2010.12.

Usage-based Pricing of Software Services Under Competition

Ram Bala[1] Scott Carr[2]

February 11, 2010

[1]Ram Bala is an Assistant Professor of Operations Management at the Indian School of Business in Hyderabad, India. He holds a Ph.D. in Management Science from the UCLA Anderson School of Management. His main research areas are product line design, promotional effort allocation, global product development and pricing and contracting strategies for services. His research cuts across disciplinary lines, particularly operations management, marketing and information systems. Address: ISB Campus, Gachibowli, Hyderabad, India - 500032; Ph: 91 40 2318 7139, Fax: 91 40 2300 7038; e-mail: ram_bala@isb.edu

[2]Scott Carr is a Principal in LECG's Washington D.C. office specializing in antitrust and competition practice. He holds a Ph.D. in both business administration and in industrial engineering. He is an expert in numerous areas including competition economics, industrial organization, product design, project management, production and service systems, supply chains and distribution networks and information systems. Address: 1725 Eye Street, Suite 800, Washington D.C., USA - 20006; Ph: 202 973 6648, Fax: 202 466 4487, e-mail: scarr@lecg.com

Abstract

With the emergence of high speed networks, software firms have the ability to deploy "software as a service" and measure resource usage at the level of individual customers. This enables the implementation of usage-based pricing. We study both fixed and usage-based pricing schemes in a competitive setting where the firm incurs a transaction cost of monitoring usage if it implements usage-based pricing. Offering different pricing schemes helps to differentiate the firms and relax price competition, particularly at higher monitoring costs, even when competing firms offer the same service quality. However, the low usage customers acquired by offering usage-based pricing are unable to compensate for the monitoring costs incurred. This implies that managers should be cautious about implementing usage-based pricing in a competitive setting.

Keywords: Pricing, competitive strategy, game theory, software industry

1 Introduction

The rise of high speed networks such as the Internet has led to the rebirth of an old business model that was once prevalent in the software industry. Back in the days when all computers were huge and expensive, time-sharing on large IBM mainframes was common practice, so firms implemented pricing schemes based on resource usage. This business model became less important as computer hardware got a lot cheaper and faster. However, the recent spread of complex enterprise software has again increased the costs of software deployment and maintenance. Consequently, software firms are looking to leverage the existence of high speed networks to deliver "software as a service" (SaaS). This term is synonymous with past nomenclature such as "Application Service Providers" (ASPs), "on-demand computing" (popularized by IBM) and "utility computing". Each of these terms essentially refers to a business model where a software vendor "develops a web-native software application and hosts and operates (either independently or through a third-party) the application for use by its customers over the Internet"[1].

The "software as a service" business model lends itself to many pricing and contracting mechanisms. First of all, the length of a software license has significantly decreased. For example, many Application Service Providers such as Salesforce.com offer a monthly license. The ability to offer a short license for a software service is typically termed as "subscription pricing". This is different from selling the software product which typically involves a contract that is valid forever or at least a significantly long period of time. Subscription pricing may or may not be usage-based. In this paper, we restrict attention to evaluating the difference between usage-based pricing and fixed pricing over a single time period while assuming that the nature of the license is a subscription. Many ASPs have embraced subscription pricing but have not necessarily adopted usage-based pricing. A prominent example is Salesforce.com in the CRM space. In our paper, this is an example of a firm offering a fixed price (or flat fee) for unlimited usage.

The last point of the previous paragraph merits further scrutiny. What factors encourage a firm to adopt usage-based pricing? Past literature such as MacKie-Mason and Varian(1994) points out to several advantages. When consumers are heterogeneous in usage propensities, many low usage consumers may not purchase a fixed price service since the firm may optimally set a price to attract only the high usage consumers. Thus, offering usage-based pricing enables the acquisition of low usage consumers. Further, even low usage consumers who might have purchased the fixed price service would prefer usage-based pricing since these consumers would no longer be subsidizing high usage consumers. Firms would prefer usage-based pricing because they can potentially extract complete consumer surplus by charging customers based on their individual usage rather than set a common fixed price. Setting a common fixed price can either turn away some low usage customers or allow some high usage customers to gain significant positive surplus. Fixed pricing schemes can result in moral hazard whereby users can increase their usage after the contract terms are agreed upon and this could adversely affect the firm, particularly in a capacity constrained environment.

Given these advantages, one would expect to see significant adoption of usage-based pricing in many industry contexts where users differ in their usage propensities but a firm cannot ex-ante estimate the usage level of a particular user (even if it knows the usage distribution across consumers). Indeed, usage-based pricing is quite common among computing infrastructure service providers.

[1] www.wikipedia.com

Some notable examples of firms that have pursued such a strategy in this space are IBM, HP and Jamcracker. However, in the software services space, usage-based pricing is nascent and coexists with fixed pricing schemes. For example, in the market for live conference software, Microsoft Live Meeting is offered as a monthly fixed price subscription service while its competitor, Adobe Acrobat Connect Professional is offered both at fixed prices and at per minute payment plans. In the project management software space, Microsoft Project Server (widely considered the high end vendor) is offered at a flat fee while ILOG Gantt for .NET is sold both as a fixed price unlimited use version as well as a run time based version. However, in the math programming / optimization space, ILOG Cplex is the high end provider and does not offer usage-based pricing for its web based service. At the same time, users operating in specific industries can opt for industry-specific web-based optimization software that is priced based on resource usage. One such example in the transportation industry is routesmith.com (used for optimizing fleet assignment and routing). In fact, routesmith.com charges users based on the number of vehicles in their fleet. Thus, in addition to the wide variation in pricing schemes, one also observes variation in the resource metric used. Run time is a commonly used metric as seen from the previous examples. InstantService.com, a web based tool for interacting with online customers, prices based on transactions and defines a transaction as one chat session or 3 email messages.

What could potentially explain the persistence of fixed price offerings in the software services market despite the much touted advantages of usage-based pricing? To seek an answer, we need to delve deeper into the specific characteristics of the software services market. Primarily, we would like to highlight two aspects of software firms and markets that may not be applicable to a wider variety of industries:

- The first aspect is that software products and services behave like *information goods.* Consequently, they display a unique characteristic of information goods as stated in MacKie-Mason and Varian(1994) and Sundararajan(2004). Administering a usage-based pricing schedule involves fixed and variable transaction costs. In many industries, these costs are low when compared to other variable production costs. However, in the context of information goods, variable costs of production are low. Consequently, when usage-based pricing is introduced, the monitoring costs appear significant and become especially relevant in designing usage-based pricing schemes for information goods.
- The second aspect is that the software service industry is a form of *access industry* (industries in which consumers pay a price based on resource utilization. Examples are gyms, theme parks, phone lines etc.). In many access industries, usage level alone is not an adequate measure of willingness to pay. Compare a retired senior citizen with a younger business executive, both of whom use an online word processing service. Clearly, the business executive would be willing to pay much more than the senior citizen for one unit of use. However, this does not reveal anything about the actual level of usage. The senior citizen may use the service as much as or even more than the business executive over a specific period of time. Thus, the willingness to pay calculation needs to take into account both the usage level and the per-use valuation of a consumer.

In this paper we analyze and compare *fixed* and *usage-based* pricing schemes in a competitive setting while incorporating certain factors that are specific to the software services industry. We restrict attention to usage-based schemes that are linear in usage. Although more complex non-linear usage-based schemes may be appropriate in many settings, linear schemes are easy to implement and hence commonly used in practice. We also make another assumption since software services are closely related to the *internet service industry*; in this industry, a first-order approximation of reality as stated in Essegaier et al(2002) is: consumer usage rate is more a function of individual usage propensity than a function of price. What this implies is that while the aggregate demand curve for the firm displays some price elasticity, an individual consumer's usage level is inelastic to changes in per-unit price. The aggregate demand curve is elastic because consumers vary in the value they attach to a single unit of use and they also vary in their individual usage propensities. Such an assumption is particularly true in the context of software services used by industrial customers. For example, the primary users of ILOG Cplex optimization software are either consultants or analytics departments of large corporations. Such customers of ILOG certainly have to make a decision on the kind of pricing scheme that they would accept. However, they are unlikely to change their usage level based on the per unit price because their usage level of the software is determined by their clients (whether external or internal) and not by them. In addition, the users of software in many large corporations are not the same people as the managers who make a decision on the pricing scheme to accept. Hence, actual usage levels are unlikely to change with changes in per unit price.

Our primary findings are as follows. In a duopoly setting with high monitoring cost, offering different pricing schemes can help to differentiate the firms even when the two firms are not differentiated in service quality. This occurs because the market segments targeted by the two firms do not completely overlap. However, this differentiation disappears as monitoring cost decreases to zero. At all values of the monitoring cost, the firm offering a fixed price makes higher profit than the firm offering a usage-based price. This implies that although low usage customers are acquired by offering usage-based pricing, the revenue accrued from them in a competitive environment does not compensate for the monitoring cost incurred. One possible remedy for the harmful competition that occurs at low monitoring cost is for the firms to differentiate their offerings such that the firm with higher service quality offers a fixed price while the other firm offers usage-based pricing. However, a higher monitoring cost may prove to be a "blessing in disguise" and such quality differentiation is not necessary.

The rest of the paper is organized as follows. Section 2 reviews related literature in the area. Section 3 sets up the basic model. Section 4 details the duopoly setting while Section 5 concludes the discussion and proposes directions for future research. All proofs are in the appendix.

2 Related Literature

The literature on non-linear pricing has roots in economics via Oi(1971) who shows that a truly discriminating two part tariff globally maximizes monopoly profits by extracting all consumer surpluses. The Oi model is extended by Schmalensee(1981) who analyzes the case of profit constrained welfare maximization. Phillips and Battalio(1983) investigate the situation where buyers can substitute between visits and also consumption between visits. Other forms of pricing can also be mapped into the nonlinear pricing framework, most notably quantity discounts (see Dolan(1987)).

The literature on nonlinear pricing described above makes some assumptions: there is no transaction cost to administering a usage-based price, and consumers are homogeneous in their preferences and usage propensities. The literature described below addresses these issues and studies nonlinear pricing for a monopoly when these assumptions are relaxed. Maskin and Riley (1984) and Wilson (1993) look at nonlinear pricing for a monopoly in the context of heterogeneous customer type but zero transaction costs of usage-based pricing. In this line of work, the firm only knows the customer type distribution but not the type of any particular individual. Hence, they find that the optimal pricing scheme is usage-based and is fully revealing. Nahata et al(1999) analyze the optimality of "buffet pricing" (similar to fixed pricing in our model) in a monopoly setting when offering usage-based pricing involves a transaction cost. However, in his model, consumers do not vary in their usage levels and they are all equally sensitive to price. Sundararajan (2004) considerably generalizes previous monopoly models in the context of information goods by using a utility function satisfying a broad set of properties that incorporates customer preferences and usage heterogeneity. He shows that offering a combination of usage-based and fixed pricing is optimal and consequently, this is never fully revealing when usage-monitoring costs are non-zero. However, customers are represented by only one type.

Although the above literature focusing primarily on monopoly settings has added substantially to our knowledge, competition is a reality in many markets including the market for software products and services. Literature on non-linear pricing in the context of competition is sparse. Hayes(1987) shows that two part tariffs act as a form of insurance in environments with uncertainty and hence is offered by firms even in a competitive setting. Another extension to a competitive setting is by Oren & Smith(1983) who study nonlinear tariffs in the context of symmetric Cournot competition (all firms offer the same type of tariff). Jain & Kannan(2002) looks at competition between connect-time and search-based pricing. In their model, users differ in their ability to conduct successful searches. Consequently, the relationship between number of successful searches and connect time is different for different consumers. Thus, they examine competition between two different forms of usage-based pricing. In our paper, the definition of usage does not make a distinction between connect time and search effort since we wish to look at competition between usage-based pricing and fixed pricing. Some of our results look similar to their work, but that is simply coincidental given that our competitive setting is quite different. Another paper that looks at competition in this setting is Essegaier et al(2002), which compares two part tariffs with fixed pricing and usage-based pricing for access service industries under conditions of customer heterogeneity and limited capacity. Customer preference is modeled with a Hotelling model. Individual customer usage is inelastic in per unit price. They look at competition in addition to monopoly. However, in all of their settings, low usage customers may be either higher or lower in willingness to pay (WTP) as compared to high usage customers but not both. Consequently, their model is primarily one of horizontal differentiation with some reference to WTP. In our model, all types of customers coexist: 1) low usage, low WTP 2) low usage, high WTP 3) high usage, low WTP 4) high usage, high WTP. Thus, to the best of our knowledge, ours is the first paper to address the issue of competition between usage-based and fixed pricing in the context of vertically differentiated consumers. Other than Bashyam(1996), which focuses on a technology choice problem, we are unaware of any other papers that employ such a model.

In short, our paper adds to the literature on information goods pricing by evaluating the effectiveness of usage-based pricing in a competitive context for the software services industry. The specific industry context chosen requires the use of a richer model of customer heterogeneity than previously seen and also the assumption of a specific cost structure. The results of our analysis provide insights on some commonly observed pricing practices in the software services industry.

3 Model definition

This section defines the general model and notation. We begin with laying out the customer utility function.

Customers and service quality: To model the two types of heterogeneity discussed earlier, each customer is described by a two dimensional vector (α, β) in which α denotes the utility that the customer will derive from a single use of the service and β represents the frequency with which the customer will use the service. A scalar U parameterizes the service quality or functionality, so an "(α, β)-customer" enjoys utility of $\alpha\beta \cdot U$ from purchasing the service.

The set of potential customers is modelled as an atomless spread of (α, β) pairs distributed evenly over the $[0, 1] \times [0, 1]$ square. Thus, if M is the size of the potential market, then any market of size $x \cdot M$ corresponds to a fraction x of the area of this square. To simplify however, the market is normalized by setting M equal to one; this is just scaling and does not sacrifice generality. Purchasing decisions follow naturally; customers are assumed to self-select whatever purchasing option maximizes this utility net of price.

Fixed pricing: Buyers pay a common price P_f, and receive unlimited use of the service. Disregarding any other pricing options, purchasing is worthwhile for an (α, β)-customer if

$$\alpha\beta \cdot U - P_f > 0. \tag{1}$$

Figure 1 illustrates. Customers who share the same value of $\alpha \cdot \beta$ derive the same utility from the service, so hyperbolae on the unit square become lines of "iso-utility." These are the curves in figure 1(a); each iso-utility line is a locus of customers with identical purchasing behavior under fixed pricing. The P_f price chosen by the firm then segments the customers along one of these hyperbolae as shown in figure 1(b). In that figure, the boundary between the "do not purchase" segment and the "pay fixed-fee" segment is the curve $\alpha \cdot \beta = \frac{P_f}{U}$.

Usage-based pricing: In this pricing scheme, the customer pays P_u *for each use.* Since she uses the service with frequency β, her total payments are $\beta \cdot P_u$ versus utility of $\alpha\beta U$, so this scheme is worthwhile to her if

$$\beta\,(\alpha U - P_u) > 0.$$

Notice that we assume that β does not change with P_u. This is as stated in the introduction. The self-selection condition is analogous to equation (1) but results in a very different structure as illustrated by figure 2. Iso-utility lines are now vertical, and the firm's selection of a particular value of P_u segments the market along $\alpha = \frac{P_u}{U}$.

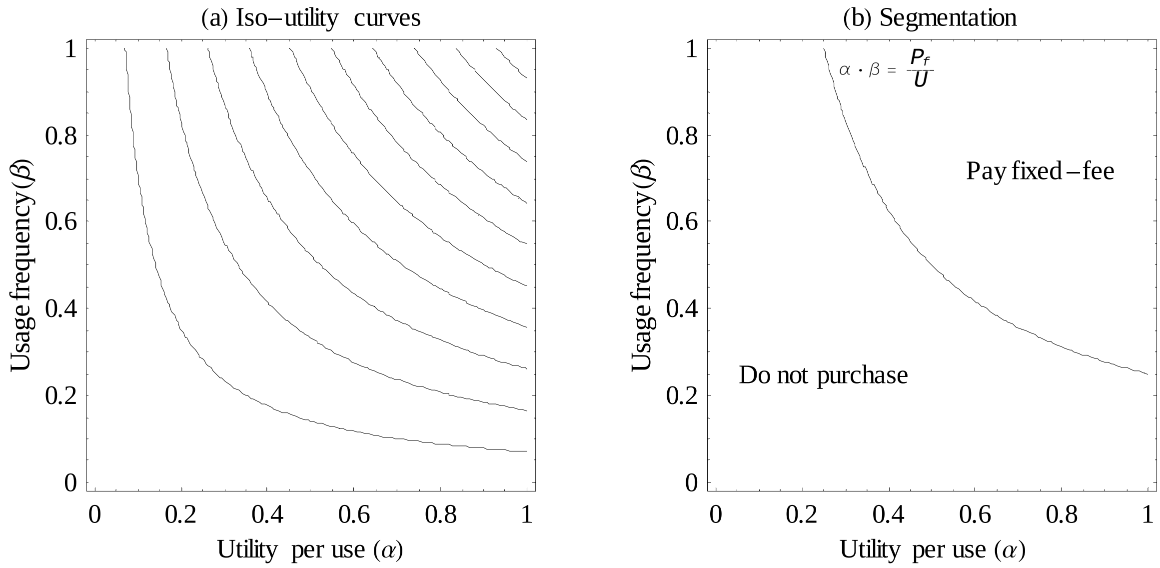

Figure 1: Fixed pricing: (a) iso-utility curves, and (b) segmentation

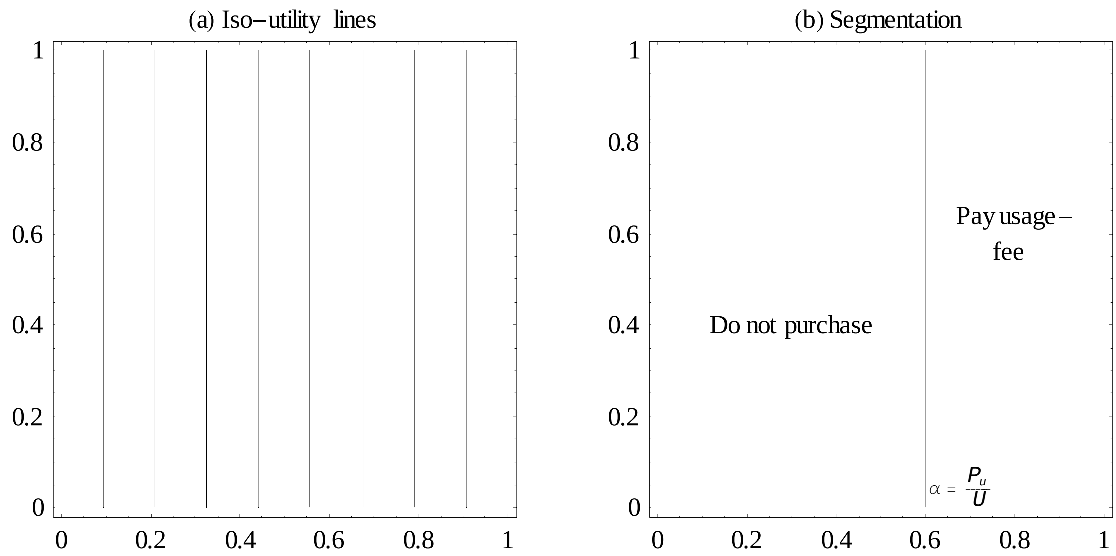

Figure 2: Iso-utility lines and segmentation: usage-based pricing

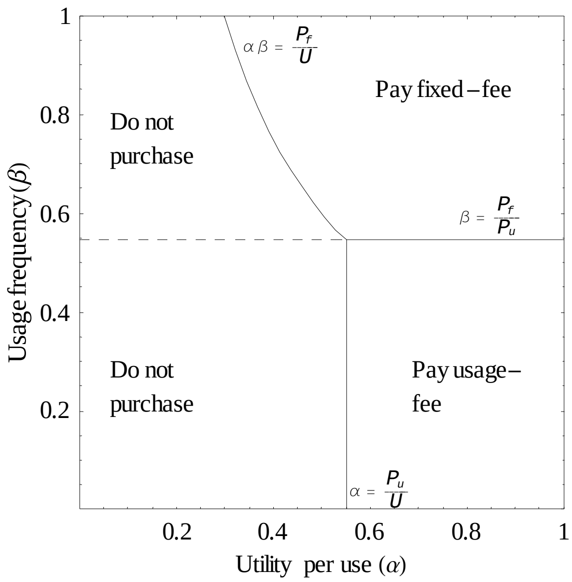

Figure 3: Segmentation for dual-scheme duopoly

Costs and revenues: With fixed pricing, a firm's profits are P_f times the area above the segmenting hyperbola. With usage-based pricing, revenues accrue on a per-use basis and the firm additionally incurs costs of c_u dollars per-use to cover the costs of metering and monitoring.

Equilibrium: The first stage of this "game" is the pricing decision; customer purchases then follow. The solution concept employed is a standard Stackelberg Nash equilibrium between the set of potential customers and the firm or firms supplying the software.

4 Duopoly analysis

We now consider scenarios with competition between two firms whose service offerings are not differentiated. We also assume that they are not differentiated in terms of cost structure. Each firm incurs a similar transaction cost of monitoring usage. Each firm decides on the appropriate pricing schedule (whether usage-based or fixed) and then both firms compete by setting prices for their respective pricing schemes. If both firms offer the same type of pricing scheme, it can be easily shown that this would result in Bertrand competition where each firm prices at marginal cost[2]. The only remaining case involves one firm that offers fixed pricing and the other that offers usage-based pricing. As before, both firms offer the same service quality (i.e., both offer the same U). Figure 3 illustrates the segmentation that results when both options are offered. The two pricing mechanisms (firms) compete by truncating the other's segment along the horizontal line $\beta = \frac{P_f}{P_u}$ – this is the locus of indifference between the fixed- and usage-based schemes.

[2]The proof of this result is straightforward and available with the authors.

P_f is chosen by the "pay fixed-fee" segment of figure 3, and q_f is the size of this segment.

$$q_f = \int_{\frac{P_f}{U}}^{1} \left(1 - \frac{P_f}{\alpha U}\right) d\alpha - \int_{\frac{P_u}{U}}^{1} \left(\frac{P_f}{P_u} - \frac{P_f}{\alpha U}\right) d\alpha = 1 - \frac{P_f}{P_u} + \frac{P_f}{U} \ln\left(\frac{P_f}{P_u}\right).$$

P_u is chosen by the "pay usage-fee" segment, and the number of times this segment uses the software service is:

$$\int_{\frac{P_u}{U}}^{1} \int_{0}^{\frac{P_f}{P_u}} \beta \, d\beta \, d\alpha \quad = \quad \left(\frac{P_f}{P_u}\right)^2 \left(\frac{1}{2} - \frac{P_u}{2U}\right)$$

The profit functions for the two firms are:

Fixed price firm:

$$\pi_f = \begin{cases} P_f \cdot (1 - \frac{P_f}{P_u}) + \frac{P_f^2}{U} \ln\left(\frac{P_f}{P_u}\right) & \text{if } P_f < P_u \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

Usage-based pricing firm:

$$\pi_u = \begin{cases} \frac{1}{2} \cdot (P_u - c_u) \cdot (1 - \frac{P_u}{U}) \cdot \left(\frac{P_f}{P_u}\right)^2 & \text{if } P_u > P_f \\ \frac{1}{2} \cdot (P_u - c_u) \cdot (1 - \frac{P_u}{U})\} & \text{otherwise} \end{cases} \tag{3}$$

An equilibrium is a (P_u, P_f) pair that simultaneously satisfies the firms' respective objectives:

$$\max_{P_u}\{\pi_u\} \text{ and } \max_{P_f}\{\pi_f\}$$

The following two lemmas outline some technical results and characterize the "best response" prices and are used to develop the equilibrium results given in the immediately following proposition.

Lemma 1 *If a non-negative function g has $g(a) = g(b) = 0$ and $g''' > 0$. Then, on the interval (a, b), the function g has a maximum and no (other) local maxima.*

Lemma 2 *At equilibrium:*

(i) For any usage-based price P_u, the fixed price firm selects a strictly lower price

(ii) For any fixed price P_f, the usage-fee firm selects:

$$P_u = \begin{cases} \frac{2cU}{1+c} & \text{if } P_f < \frac{2cU}{1+c} \\ P_f & \text{if } \frac{2cU}{1+c} \le P_f \le \frac{(1+c)U}{2} \\ \frac{(1+c)U}{2} & \text{if } P_f > \frac{(1+c)U}{2} \end{cases}$$

Result i) ensures that the fixed-pricing firm gets at least one customer. A combination of results i) and ii) along with lemma 1 help derive the equilibrium pricing strategy. This is detailed in the next proposition.

Proposition 1 *(i) There exists a unique equilibrium in pure strategies.*

At equilibrium,

(ii) $P_u^* = \frac{2cU}{1+c}$

(iii) The fixed-price firm always makes higher profit than the usage-based pricing firm for all non-zero values of the monitoring cost.

(iv) Fixed-price profit increases with the monitoring cost.

(v) There exists an interval bounded below by $c = 0$ *for which the usage-fee profit increases with the monitoring cost.*

(vi) There exists an interval bounded above by $c = 1$ *for which the usage-fee profit decreases with the monitoring cost.*

(vii) In the limit as monitoring cost approaches zero: the market is fully covered, each firm gets exactly half of the market, and $\pi_f = \pi_u = 0$.

Noting that the firms compete through price-setting and that the firms' service qualities are identical, one might expect to see the "Bertrand result" that the firms are unprofitable at equilibrium. It is thus interesting to observe that both firms' equilibrium profits are actually strictly positive in this model (for all $c \in (0, 1)$). The key here is that the Bertrand result depends on a complete lack of differentiation. In this duopoly however, the fact that the two firms offer different pricing schemes provides a form of differentiation that results in a profitable equilibrium. This occurs because the target customer segments of the two different pricing schemes do not completely overlap. This fact is illustrated visually by the different shapes of the market segments that result from the two pricing schemes. Result (*iii*) reveals an additional insight however. Although the firms are not differentiated in service offerings, the fixed pricing firm makes higher profit. This result is not ex-ante obvious because although the usage-based pricing scheme incurs a cost not incurred by the fixed pricing scheme, it also has the ability to attract several low usage consumers. It is this very fact that enables a monopoly firm to prefer usage-based pricing over fixed pricing at low monitoring costs[3]. However, in a competitive setting, the revenue generated by the low usage consumers who are acquired cannot compensate for the monitoring cost incurred. Consequently, the usage-based pricing firm always makes lower profit. This formalizes the insight that usage-based pricing, though valuable in a monopoly setting with low monitoring costs, is highly sensitive to competition from a fixed pricing scheme.

Equilibrium results for the dual-fee duopoly are illustrated in figure 4. A notable feature is that the usage-fee firm's profit is significantly lower than that of the fixed-fee firm. Further, usage-fee profits (π_u) are non-monotonic in the monitoring cost c (as opposed to what would be expected for a usage-fee monopolist). Rather, that firm's profits first increase and then decrease with c (as anticipated by proposition 1(*v*) and (*vi*)). The fact that π_u can increase as its costs increase is

[3] The proof of this result is straightforward and is available with the authors.

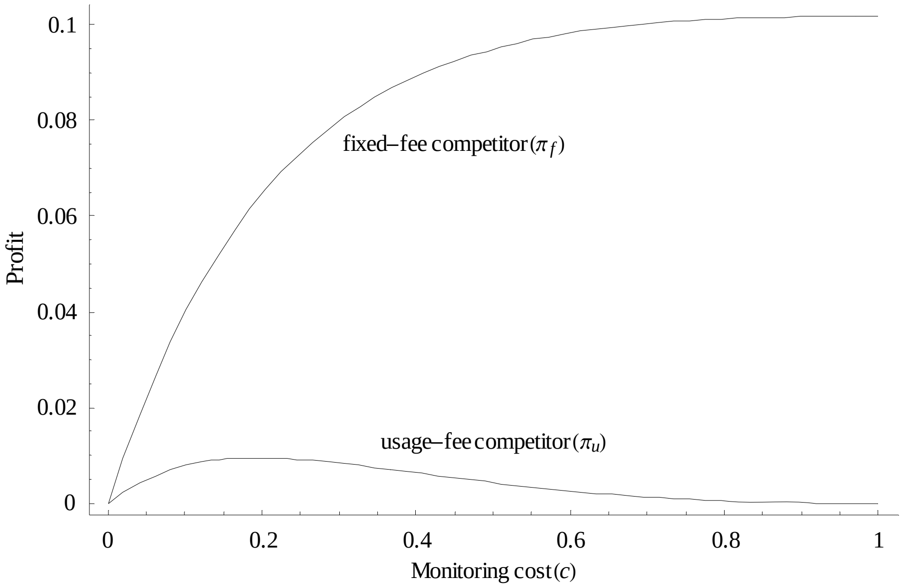

Figure 4: Competitor's profits: dual scheme duopoly

somewhat surprising, but it has a straightforward explanation – the increase in c has the effect of reducing the degree of competition between the two firms, and this is beneficial to both competitors. This does not continue indefinitely however; for c greater than about 0.2 the negative effects of increasing c dominates and π_u begins to fall. As c gets large, π_u disappears and the fixed pricing firm becomes essentially a monopolist. An analogous result is developed in Jain & Kannan(2002) but in a competitive setting where the firms offer either connect-time based or search-based pricing. To the best of our knowledge, this is the first paper to highlight this result in the context of competition between a firm offering a fixed price and a firm offering a usage-fee.

While this numerical study reveals important details, a broader conclusion emerges regarding the usage-based pricing scheme. While the usage-based pricing scheme is optimal for a monopoly firm at low monitoring cost, it is highly sensitive to competition against a fixed pricing competitor. At the same time, a diversity of pricing schemes can help to differentiate firms if the service offerings are not significantly differentiated. However, such differentiation disappears when the monitoring cost decreases to zero. Consequently, firms would have to differentiate their offerings in the context of low monitoring costs[4]. However, a higher monitoring cost may prove to be a "blessing in disguise" and such quality differentiation is not necessary.

5 Concluding Remarks

The rise of the "software as a service" business model has led to the rebirth of usage-based pricing. However, usage alone is not an adequate measure of the willingness to pay for consumers of software services. Consumers vary in the value they derive with the same amount of use. In ad-

[4] An analysis of competition between differentiated software service providers is available with the authors.

dition, firms that implement usage-based pricing incur a monitoring cost. Given a choice between fixed and usage-based pricing schemes, a monopolist would find usage-based pricing to be optimal at low monitoring costs and a fixed price at higher monitoring costs. We study a competitive setting where consumers offer the same service offering but can potentially offer different pricing mechanisms(usage-based versus fixed pricing). When monitoring costs are high, we find that offering different pricing schemes enables firms to differentiate themselves purely on the basis of pricing mechanisms even when their service quality values are not different. This occurs because the two pricing schemes target customers who do not completely overlap. However, when firms compete with different pricing schemes, lower monitoring cost leads to intense price competition. An increase in monitoring cost relieves price competition and benefits both firms. This result overturns the monopoly result where lower monitoring cost is always better. Another insight is that the fixed pricing firm always makes higher profits than the usage-based pricing firm despite the fact that both firms offer the same service quality. This occurs because the revenue accrued from low usage customers by deploying usage-based pricing in a competitive setting does not compensate for the monitoring cost incurred. This reveals that the usage-based pricing scheme is highly sensitive to competition, specifically when the competitor offers a fixed price. These results serve to explain the persistence of fixed pricing in the software services industry. In a duopoly scenario with high monitoring costs, firms do better when they differentiate their pricing schemes even in the absence of service differentiation, thus allowing some firms to persist with fixed pricing. Hence the much touted notion of "usage-based pricing" may not be a panacea for software service providers.

As always, our results are limited by our assumptions. One of our assumptions is that both usage level and utility per use of customers is uniformly distributed. To evaluate the impact of a non-uniform distribution, we observe that the usage-based pricing firm acquires customers who have lower levels of usage but higher levels of utility per unit of usage. Thus, if the distribution of customers over the unit square has higher density over the low usage - high utility per use quadrant, then the number of customers (and consequent revenue) acquired by the usage-based pricing firm in a competitive setting could offset the monitoring cost incurred. Another possibility that we do not consider is that users may be uncertain about their own levels of use and the choice of pricing schemes is frequently affected by such uncertainty. However, uncertainty in terms of usage frequency merely reinforces the vulnerability of usage-based pricing. When usage level is uncertain and customers are risk averse, they are more likely to opt for the flat/fixed pricing schemes, thus negatively impacting usage-based pricing. Thus our key result is likely to be reinforced with such a setting. Further research might seek to address such issues and also focus on empirical validation of some of our key research findings.

References

[1] Bashyam, T. C. A., 1996. Service design and price competition in business information services. Unpublished Ph.D. dissertation, UCLA

[2] Dolan, R. 1987. Quantity discounts: managerial issues and research opportunities. *Marketing Science 6(1) 1-22*

[3] Essegaier, S., Gupta, S. and Zhang, Z. J. 2002. Pricing access services. *Marketing Science 21(2) 139-159*

[4] Oi, W. 1971. A disneyland dilemma: two-part tariffs for a mickey mouse monopoly. *Quarterly Journal of Economics 85(1) 77-96*

[5] Hayes, B. 1987. Competition and two-part tariffs. *Journal of Business 60(1) 41-54*

[6] Jain, S., and P. K. Kannan. 2002. Pricing of Information Products on Online Servers: Issues, Models, and Analysis. *Management Science. 48(9) 1123-1142*

[7] MacKie-Mason, J. and Varian, H. Some FAQs about Usage-based pricing. *Computer Networks and ISDN Systems, vol. 28 (1995): 257-65*

[8] Maskin, E. and Riley, J. 1984. Monopoly with incomplete information. *The RAND Journal of Economics 15(2) 171-196*

[9] Nahata, B., Ostaszewski, K., and Sahoo, P. 1999. Buffet pricing. *Journal of Business 72(2) 215-228*

[10] Oren, S. and Smith, S. 1983. Competitive Nonlinear Tariffs. *Journal of Economic Theory 29 49-71*

[11] Phillips, O. and Battalio, R. 1983. Two-part tariffs and monopoly profits when visits are variable. *The Bell Journal of Economics 14(2) 601-604*

[12] Schmalensee, R. 1981. Monopolistic two-part pricing arrangements. *The Bell Journal of Economics 12(2) 445-466*

[13] Sundararajan, A. 2004. Nonlinear pricing of information goods. *Management Science* 50(12) 1660-1673

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

Appendix

Proof of lemma 1

First, it must be that g is concave-convex on $[a, b]$.[5] That is, there is a value c such that g is strictly concave over subinterval $[a, c)$ and strictly convex over $(c, b]$. To see this: (1) g must be strictly concave at a – otherwise, $g''' > 0$ would imply that g is strictly *convex* everywhere, and this would make $g(a) = g(b) = 0$ impossible. (2) $g''' > 0$ implies that if g is convex at c then g is strictly convex over all $(c, b]$. Next, g must have exactly one point in (a, b) at which $g' = 0$. This is seen by contradiction: assume $g'(x) = g'(y) = 0$ for some $x < y$ in (a, b). Then, (because g is concave-convex) g must be concave at x, convex at y, and increasing between y and b. This then implies that $g(y) < g(b)$ which together with $g(b) = 0$ contradicts the premise of nonnegative g. Finally, the strict concavity of g over (a, c) together with $g = 0$ at endpoints a and b implies that the unique inflection point must be a maximum.

Proof of lemma 2:

[5] Although the convex portion may be empty.

(i) (by contradiction): If $P_f \geq P_u$ then: (1) For every (α, β)-customer, the utility $\alpha\beta U - P_f$ derived from the fixed price option is $\leq \beta(\alpha U - P_u)$, the utility derived from the usage-fee option (because $\beta \leq 1$). (2) Thus, q_f, sales by the fixed price firm are 0. (3) But, the continuity of π_f guarantees that the fixed price firm can always find a price that will supply strictly positive profits. (4) Thus, a (P_u, P_f) pair with $P_f \geq P_u$ violates this fixed firm's optimality criterion. Hence at optimal pricing for the fixed price firm, we have $P_f < P_u$

(ii) To simplify the analysis, let: $\pi'_u = \frac{\pi_u}{U}$, $p_f = \frac{P_f}{U}$, $p_u = \frac{P_u}{U}$ and $c = \frac{c_u}{U}$:

The profit function of the usage-fee firm in equation (3) becomes:

$$\pi'_u = \begin{cases} \frac{(p_u - c)\cdot(1-p_u)\cdot p_f^2}{2\cdot p_u^2} & \text{if } p_u > p_f \\ \frac{(p_u-c)\cdot(1-p_u)}{2} & \text{otherwise} \end{cases} \tag{4}$$

Differentiating the first line of the profit function in equation (4) with respect to p_u :

$$\frac{\partial \pi'_u}{\partial p_u} = \left(-\frac{1}{p_u^2} + \frac{2c}{p_u^3} - \frac{c}{p_u^2}\right)\frac{p_f^2}{2}$$

$$\frac{\partial^2 \pi'_u}{\partial p_u^2} = \left(2 - \frac{6c}{p_u} + 2c\right)\frac{p_f^2}{2p_u^3}$$

Setting the first derivative to zero, we get $p_u^*(p_f) = \frac{2c}{1+c}$. Now $\frac{\partial^2 \pi'_u}{\partial p_u^2} \leq 0$ for $p_u \in [0, \frac{3c}{1+c}]$ and > 0 for $p_u \in (\frac{3c}{1+c}, 1]$ implying that π'_u is concave-convex in p_u. Also, at $p_u = c$, $\pi'_u = 0$ and $\frac{\partial \pi'_u}{\partial p_u} = -(1-c)\frac{p_f^2}{2} \leq 0$ at $p_u = 1$ implying that the function is strictly concave with zero value at the lower limit of the domain and strictly convex decreasing at the upper limit of the domain. Combining the above facts implies strict quasi-concavity of the objective function over the given domain. Setting the first derivative equal to zero provides a unique maximum. If $p_f < \frac{2c}{1+c}$, the best response $p_u^*(p_f) = \frac{2c}{1+c} > p_f$. If $p_f \geq \frac{2c}{1+c}$, then the quasi-concavity of the profit function dictates that $p_u^*(p_f) = p_f$. Differentiating the second line of the profit function in equation 3, we find the optimal usage-fee to be $p_u^*(p_f) = \frac{1+c}{2}$. If $p_f \leq \frac{1+c}{2}$, then $p_u^*(p_f) = \frac{1+c}{2}$ else $p_u^*(p_f) = p_f$. Combining the results above, the best response usage-fee for the entire range of fixed prices can be constructed as stated in the lemma.

Proof of proposition 1: (i) Using lemma 2, we can restrict analysis to the case where $p_f < p_u$. Similar to previous part of the proof, we set $\pi'_f = \frac{\pi_f}{U}$, $p_f = \frac{P_f}{U}$, and $p_u = \frac{P_u}{U}$. The first line of the profit function of the fixed price firm from equation (2) is simplified to give:

$$\pi'_f = p_f \cdot (1 - \frac{p_f}{p_u}) + p_f^2 \ln\left(\frac{p_f}{p_u}\right) \tag{5}$$

with:

$$\frac{\partial \pi'_f}{\partial p_f} = 1 + p_f - \frac{2p_f}{p_u} + 2p_f \ln\left(\frac{p_f}{p_u}\right) \tag{6}$$

and:

$$\frac{\partial^2 \pi'_f}{\partial p_f^2} = 3 - \frac{2}{p_u} + 2\ln\left(\frac{p_f}{p_u}\right)$$

$$\frac{\partial^3 \pi'_f}{\partial p_f^3} = \frac{2}{p_f} > 0$$

Using this fact about the third derivative and that $\pi_f = 0$ at the endpoints $p_f = 0$ and $p_f = p_u$, lemma 1 provides the result that there exists a unique fixed price response to any usage-fee. From the earlier part of this proof, we know that when $p_f < p_u$, the usage-fee firm has a unique response $p_u = \frac{2cU}{1+c}$. Since there is a unique fixed price response to this price, the resulting unique set of prices maximizes the profit of both firms given the strategy of the competitor and hence constitutes an equilibrium in pure strategies. Let this price pair be denoted by $(p_f^e\ ,\ p_u^e)$.We show uniqueness of this equilibrium by contradiction: suppose that $(p'_f\ , p'_u)$ represents an equilibrium price pair in addition to the price pair $(p_f^e\ ,\ p_u^e)$. By lemma 2, $p'_f < p'_u$ resulting in $p'_u = \frac{2c}{1+c}$ which equals p_u^e. From the earlier part of this proposition, there is a unique fixed price response to this price. Hence $p'_f = p_f^e$. Hence $(p'_f\ , p'_u) = (p_f^e\ , p_u^e)$ and the equilibrium is unique.

(ii) Follows from $i)$

(iii) At equilibrium we know that the first derivative of the profit function for the fixed price firm should be zero. Using the expression for this first derivative from equation (6) and applying this condition:

$$1 + p_f - \frac{2p_f}{p_u} + 2p_f \ln\left(\frac{p_f}{p_u}\right) = 0$$

Rearranging this:

$$p_f \ln\left(\frac{p_f}{p_u}\right) = \left(\frac{1+p_f}{2}\right) - \frac{p_f}{p_u}$$

Substituting for the lhs of the above equation in the profit function given by equation (5), we get:

$$\begin{aligned}\pi_f &= p_f\left(1 - \frac{p_f}{p_u}\right) + p_f\left(\frac{p_f}{p_u} - \left(\frac{1+p_f}{2}\right)\right) \\ &= \frac{p_f(1-p_f)}{2}\end{aligned}$$

Comparing firm profits at equilibrium, we get $\pi_f > \pi_u$ only if:

$$\frac{p_f(1-p_f)}{2} > \frac{1}{2}(p_u - c)(1-p_u)\left(\frac{p_f}{p_u}\right)^2$$

Simplifying the above equation and using the fact that $p_u = \frac{2c}{1+c}$:

$$p_f < \frac{4c}{(1+c)^2} = \left(\frac{2}{1+c}\right)\left(\frac{2c}{1+c}\right)$$

Observe that, at equilibrium:

$$p_f < p_u \leq \left(\frac{2}{1+c}\right)p_u = \left(\frac{2}{1+c}\right)\left(\frac{2c}{1+c}\right)$$

Thus, we must have $\pi_f > \pi_u$ at equilibrium.

(*iv*) The fixed-price firm's modified profit function is given by equation (5). This profit is an increasing function of the usage-fee as shown by direct differentiation:

$$\frac{\partial \pi'_f}{\partial p_u} = \frac{p_f^2}{p_u}\left(\frac{1}{p_u} - 1\right) > 0$$

This is true for every p_u, so it is also true for the equilibrium p_u. Also, c does not appear in the fixed-price firm's profit function. This implies that the fixed-price firm's equilibrium profit increases if and only if the equilibrium p_u increases with c. By inspection of lemma 2, we find that it does.

(*v*) & (*vi*) Let $\pi_u^*(c)$ represent the equilibrium profit for the firm offering the usage-fee at optimal prices as a function of the monitoring cost. $p_u^* < 1$ implies that customers will purchase a strictly positive number of usage-fee transactions, and $p_u^* > c$ implies that these transactions are profitable. Thus, $\pi_u^*(c) > 0$ whenever $0 < c < 1$. The stated results are then implied by the continuity of π_u^* in c (which can be verified by the implicit function theorem).

(*vii*) At equilibrium prices, the first derivative of the fixed price firm's profit (c.f. equation 5) as a function of price must be zero. That is:

$$\frac{\partial \pi'_f}{\partial p_f} = 1 + p_f - 2\frac{p_f}{p_u} + 2p_f \ln(\frac{p_f}{p_u}) = 0$$

where p_f and p_u are the equilibrium prices. Taking the limit of the above equation as $c \to 0$ gives:

$$1 + \lim_{c \to 0} p_f - 2 \cdot L + 2 \cdot \lim_{c \to 0} p_f \cdot \ln(L) = 0$$

where $L = \lim_{c \to 0}\left(\frac{p_f}{p_u}\right)$. We know that at equilibrium, $0 < p_f < p_u$. However, $\lim_{c \to 0} p_u = 0$ (using the closed form expression of p_u). Taking limits on both sides of the inequality, we have $\lim_{c \to 0} p_f = 0$. Using this in the above equation:

$$1 - 2 \cdot L = 0$$
$$\lim_{c \to 0}\left(\frac{p_f}{p_u}\right) = \frac{1}{2}$$

which is the same as:

$$\lim_{c \to 0}\left(\frac{P_f}{P_u}\right) = \frac{1}{2}$$

At $c \to 0$, we have both prices: $p_u \to 0$ and $p_f \to 0$. At zero prices, the market is fully covered. $\frac{p_f}{p_u}$ is the line of indifference between the two segments and $\frac{p_f}{p_u} \to \frac{1}{2}$ as $c \to 0$. Hence each firm gets exactly half of the potential market.