Pricing of Information Products on Online Servers: Issues, Models, and Analysis

Sanjay Jain • P. K. Kannan
The Robert H. Smith School of Business, University of Maryland, College Park, Maryland 20742
sjain@rhsmith.umd.edu • pkannan@rhsmith.umd.edu

Online information servers that provide access to diverse databases where users can search for, browse through, and download the information they need have been rapidly increasing in number in the past few years. Online vendors have traditionally charged users for information on the based on the length of the time they were connected to the databases. With hardware and software advances, many online servers have recently started changing their pricing strategies to search-based and/or subscription-fee pricing. This paper examines the various issues involved in pricing these information products, and presents an economic approach to analyze conditions under which the various pricing schemes may prove optimal for the online servers. Our results show that the variation in consumer expertise and valuation of information affects the choice of a pricing strategy by the server. We present general conditions under which subscription-fee pricing is optimal even when consumer demand is inelastic. We also find that, given the cost structures characterizing the market, undifferentiated online servers can compete and coexist in the market each making positive profits. We show that in a competitive setting an increase in costs of online servers can sometimes benefit them by enabling them to differentiate themselves. Our results offer insights into the trends in pricing strategies and may provide an explanation as to why many servers may persist with connect-time strategies.
(*Pricing; Digital Products; e-Service; Game Theory*)

1. Introduction

Recent years have seen a steady growth in the number of information services companies that offer electronic access to proprietary databases. With firms in a variety of different industries facing increased worldwide competition, there is a greater need for accurate and timely information. This trend has translated into an annual $33.5 billion market for the electronic information services industry for 1998 in the United States alone with a annual growth rate of 7.5% (Communications Industry Forecast 2000). There are two types of firms engaged in the information services business: information producers, who develop and update the databases, and online vendors, who provide online access to these databases. Together these firms meet customer information needs in diverse areas such as marketing information (market research, direct marketing, product movement, audience assessment), financial information (credit information, financial and economic information, stock quotes), chemical abstract services, legal abstract retrieval services, news retrieval services, and medical journal abstract retrieval services. Although such information is provided through various media such as print, CD-ROM, magnetic tape, floppy disk, etc., 70% of the revenue is generated from online delivery, where consumers access these databases through online servers. There were more than 14,000 online servers operational in late-1999 (Gale's Directory of Databases 2000). With the rapid development of telecommunication technologies and the Internet, this number is expected to grow at an increasing rate.

0025-1909/02/4809/1123$5.00
1526-5501 electronic ISSN

How the information product is priced is an issue of considerable importance to (1) the producers of the databases, (2) the online information servers who provide access to these databases, and (3) the customers who use the information. With the growth of the industry, online servers, as providers of consumer access to many proprietary databases, have been dominating the market and have had a significant impact on the information product pricing. Traditionally, the online servers charged a connect-time-based pricing in which consumers pay on the basis of the time they are connected to the database. The servers, in turn, shared a percentage of the connect-time charges with the database producer as "royalties" for using the database. However, with rapid improvements in information technology, better search algorithms, increased data transfer rates, increased connectivity, and capacity, this pricing strategy has become so cumbersome to implement and enforce that many online servers, consumers, and other market participants have started questioning the wisdom of the connect-time pricing strategy. Thus, while connect-time pricing still remains a common pricing strategy in the industry, a significant number of servers have started adopting different pricing strategies. Some are based on successful search—where successful search is defined as one which results in viewing/downloading of desired information and consumers pay only when they successfully retrieve (view and/or download) the desired information (per search pricing). Other methods of pricing include flat-rate pricing (or subscription-based pricing) and different variations of connect-time, search-based, and subscription pricing.

While there seems to be a general agreement among all industry participants that the dominance of connect-time pricing strategy may almost be over, there is no consensus among the online servers, who traditionally set the prices the customer pays, or among the database producers, as to what could be a viable alternative to the connect-time pricing strategy (see for example, Gregory 1992, Stricker 1995, Abels 1996). While some practitioner-oriented articles have addressed this issue (see Nelson 1995, Hawkins 1989, Rawley and Butcher 1996) and there has been significant research in a related area of network access pricing (Gupta et al. 1997, Konana et al. 1996, 2000, Dewan and Mendelson 1990, Shenker 1995, Stidham 1992), there is a clear need for formal research in understanding the pricing strategies for the sale of information products over the network environment.[1] The purpose of this paper is to provide a formal analysis of the pricing issues for information products. Specifically, we provide a unified framework for examining the pros and cons of the various pricing strategies from the perspective of the various market participants. We present an economic approach to analyze the conditions under which various pricing schemes may prove to be optimal for online servers. We also examine how competition among database producers and among online servers can affect the pricing strategies that are adopted. We also show how some of the recent market developments can be explained within the framework of our modeling approach.

The rest of the paper is organized as follows. In the next section, we provide an overview of the information services industry, the players in the market, a short history of the pricing practices, and the issues and implications of the various pricing strategies. We also position our research in the context of related work in OR, MIS, and economics. In §3, we develop a model of the information market with a monopolist online server. In §4, we extend the model to consider the impact of competition among online servers and in §5, we present other model extensions. We discuss the implications of our analysis for the information market, and conclude the paper in §6.

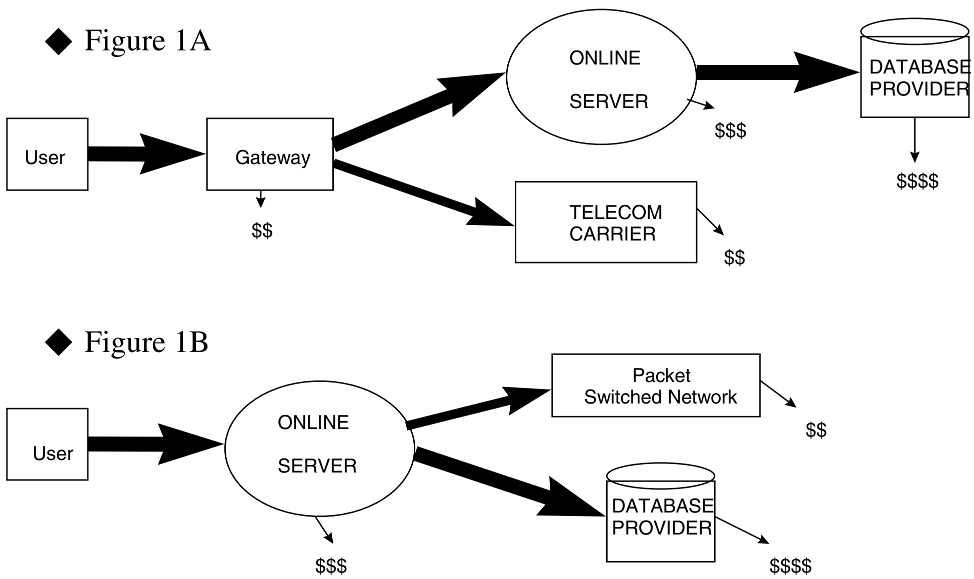
2. Online Information Products Market

2.1. The Market

There are three key players in the online information services market: the information producer, the online server/vendor, and the information consumer. The clients were and have been mainly organizations including large corporations in diverse industries, libraries, research centers, research firms, and

[1] In §2.4, we discuss this literature in greater detail and position our paper in relation to prior work.

Figure 1 Information Revenue Stream



(Source: *Online*, March 1989; David Hawkins).

small businesses. As the online services market grew and online technology improved, the power of online vendors (such as, for example, Mead and DIALOG) increased, and many database producers started using online vendors to gain access to the large markets the vendors owned. Over time, the revenue from print products have been steadily decreasing for the information producers while online revenues have been steadily increasing. Even though developments in technology have allowed information producers to gain access to some of the customers through the sales of data products on CD-ROMS, direct tape loads, and related technology, online delivery still accounts for more than 70% of their revenue, thus preserving the clout of online vendors. In recent years, there has also been vertical integration in the market with many information producers starting their own online servers, and online servers getting into database production through acquisitions.

2.2. Information Revenue Stream

It is necessary to have an overview of the information revenue stream in the online information market to understand the pricing issues. There are two ways a user can connect to a server: indirectly through a gateway or directly through a packet switched network (see Figure 1). In the first case (Figure 1A), the gateway provider charges the user for the link and arranges for the link between itself and the online server. The server then charges the user for the information and pays to the database producer. In the second case (Figure 1B), the online server collects all payments from the user and pays the database producer and the packet network provider. Among the online servers, the second approach is more common.

One of the unique aspects of purchasing information online is that it is usually paid for *as* it is retrieved or *after* it is retrieved as against printed information, which is paid for *in advance*. Since there is uncertainty involved in searching for the information needed, the primary concern of early online servers such as DIALOG Information Services was to develop a pricing scheme that was easily understood by all and reduced the uncertainty in what the consumer would pay. This was the start (in 1971) of the connect-time-based pricing where consumer pays on the basis of the time he or she is connected to the server. The online server, in turn, pays royalties to the database producer and connect charges to the packet network provider based on the connect time. When the partnership between information producers and online servers first started, the royalty split was around 50%–50% between the two. However as competition increased, and as online servers' cost of sales went up, the royalty split changed with more revenue going to the online vendors. The standard royalty split currently is in the range of 25%–35% for an information producer (Abels 1996, p. 236), depending on how exclusive and differentiated the information product is.

2.3. Issues in Pricing

Connect-time pricing is still a very commonly used strategy among the large online servers (e.g., DIALOG and BRS). However, there are many problems with this strategy. First, connect time does not relate well to the amount of information retrieved, and thus, sophisticated searchers get more information for their money than naïve users. Second, many consumers tend to structure their searches in advance, use PC-based search software (for example, ProSearch), and minimize their time online, thus minimizing the use of the interactive nature of the system. Also, advances in modem technology and speeds have led to shorter connect times and the need for frequent changes in connect-time rates. In addition, clients find the connect-time pricing difficult to budget in advance as the more one searches the more one pays. Also, it is paradoxical that the more successful an online server is in terms of providing information, the less profitable it becomes for the server to provide the information.

Given the disadvantages of connect-time strategy, in 1985 EasyNet announced a flat fee per-search pricing in which the user pays a single fee for a successful search, and it was soon followed by ESA-IRS, a Europe-based online server. In this pricing scheme, users pay only for the information that is successfully viewed and/or downloaded. Thus, "success" is defined on the basis of viewing/downloading the desired information. In addition to the "success" fee a small fixed session fee is charged regardless of whether information is purchased or not.

Finally, in early 1993, some online servers (OCLC FirstSearch, H. W. Wilson Company, etc.) started offering subscription-fee pricing, that is, a single fee for an entire year's worth of searching, no matter how many searches are done or how many items are retrieved

Table 1 Pricing Strategy of Online Servers

	1997		2000	
Pricing strategy	Number of database servers	%	Number of database servers	%
Connect-time-based pricing	483	67.4%	447	62.3%*
Search-based pricing	149	20.8%	167	23.2%
Subscription-fee pricing	70	9.7%	89	12.4%
Others (free, contracts, contribution)	15	2.1%	15	2.1%

Note. The 1997 data is based on a survey of 508 databases randomly chosen from Gales 1997 Directory of Databases and Online Servers. The 2000 data is based on the survey of almost the same databases (96% overlap) from the Gales Online Directory. A * indicates significantly (at 0.05) different ratios over the two time periods based on a two independent sample t-test.

(Tenopir 1993). The subscription fee that is contracted with a customer generally depends on how much a customer has searched in the past with the other pricing schemes, plus a percentage to account for the anticipated increased usage once the subscription fee is in place (There are exceptions, such as H. W. Wilson Co. whose prices are the same for all customers).[2] Though the subscription-fee per year strategy is not really cheap for all clients, it does take away the uncertainty in budgeting and could be seen as a viable option for high-volume consumers. The subscription fee generally varies for the same customer from year to year depending on the past year's usage so as to ensure that the online servers get adequate return.

The current status of pricing in the industry is as follows: While more servers have started adopting per-successful-search pricing or subscription-fee per year pricing (O'Leary 1993), many servers continue to use the connect-time-based pricing (Nelson 1995, Abels 1996). Table 1 shows the incidence of various pricing strategies adopted by servers for accessing databases based on a survey of 510 databases (718 database-server combinations) drawn randomly from the Gale's Directory of Databases and Online Servers

[2] While the subscription fee strategy could be seen as a move to take on the CD-ROM market, the fact is that CD-ROMs were never in direct competition with online servers, which offered a lot more services, search algorithms, and flexibility, that CD-ROMs could never match. Thus, the impact of CD-ROMs on online information market has never been very significant (see Tenopir 1996).

(Gale's Directory 2000) (approximate 5% sample). As seen from Table 1, over the two time periods shown (1997 and 2000), an average of 65% of the instances a connect-time based pricing strategy is used as compared to an average of 22% of the instances involving search-based pricing strategy. Some servers are also adopting different pricing strategies for accessing different databases through their online services. Connect-time-based strategy is also quite popular with the database providers for calculating royalties as they understand the implications of revenue better than with the other newer strategies (Nelson 1995). Online service providers and online practitioners are still debating the pros and cons of the different pricing strategies, and no consensus has emerged as yet.

Against this backdrop, there are several questions that the industry is confronted with. Are the days of connect-time-based pricing over? Will the technological advances and the significantly lower marginal costs for computing, storage, and telecom make the strategy obsolete? If this is the case, why is connect-time-based pricing still the most widely observed strategy (Table 1)? What will be the impact of alternative pricing strategies on the continued viability of the database producers? How will competition between database providers affect the pricing strategies? How will competition between online servers shape the market and pricing strategies? Our objective in this paper is to provide a simple modeling framework that will provide a starting point for gaining insights into the research questions posed above. Given the power structure in the industry, we examine these issues from the perspective of an online server.

2.4. Related Literature and Research Positioning

Pricing of information services poses some issues that are not studied in the traditional pricing literature (for reviews see, for example, Lilien et al. 1992). The typical pricing problem involves a situation where a consumer knows the price of a product before he or she buys the product and the seller knows the cost that he or she incurs for the product for a given consumer. However, this assumption is not necessarily true in the pricing problem that we study. This is because the time that a consumer takes to find the desired information is usually a random variable. Since most database providers charge online servers on the basis of connect time, the cost of searches for the servers is, therefore, also a random variable. Furthermore, if the online server also charges a connect-time-based price, the price that each consumer pays is also a random variable. Thus, a consumer makes a purchase decision without perfect knowledge of the price he or she is going to pay.

There are two streams of literature that are related to the work we present in this paper. First, there is a large body of literature which has dealt with pricing of network services when network congestion is an issue. The second streams deals with pricing of information goods which are sold over a network. We discuss each of these streams of literature.

Pricing of Network Services. There is substantial amount of research which has examined how network services should be priced when networks have limited capacity and network congestion is an issue. Mendelson (1985) examines optimal pricing for computer services. He considers a situation in which usage by a consumer exerts a negative externality on other consumers by increasing delay costs. He finds that in order to maximize the efficiency of the system, users should be charged on the basis of the externality that they impose on other users. Thus, his results suggest that optimal pricing is related to usage amount. Mendelson and Whang (1990) extend this analysis by considering various classes of users and derive an incentive-compatible pricing scheme which can maximize the net value of the system. Dewan and Mendelson (1990) extend Mendelson's (1985) analysis by considering both user's delay cost and capacity cost. Stidham (1992) examines optimal pricing and capacity decisions. Westland (1992) extends these models by considering both positive and negative externalities.

Some authors have specifically focused on designing optimal pricing schemes for Internet services so as to ensure proper allocation of Internet infrastructure. Cochi et al. (1993) suggest that optimal pricing of network services require that consumers are charged on the basis of service that they desire. For example, consumers desiring faster transmission need to pay a higher price. They show that such a pricing scheme can lead to optimal network performance.

The pricing scheme that the authors propose is static in that the prices do not change with the degree of congestion. MacKie-Mason and Varian (1995) propose a dynamic pricing scheme in which users of Internet services should be charged on the basis of usage when the network is congested. They propose a smart market mechanism in which each user would bid a price for processing their job. The smart mechanism will determine the prices that will be paid by the users on the basis of congestion and the users' bids at the time. However, the proposed smart market mechanism may be difficult to implement (Gupta et al. 1996). Gupta et al. (1997) also study the pricing of Internet services when network congestion is an issue. They propose a pricing scheme in which a user of Internet services can choose from a menu of options which include the price, quality level of service, expected time at which the service will be provided, etc. Using a computer simulation model, they show that approximately socially optimal prices can be implemented in a practical decentralized system.

In the context of pricing of real-time databases, Stonebracker et al. (1996) suggest an economic approach for managing a wide area distributed network. They propose a new distributed database management system called MARIPOSA, in which jobs are processed by a broker on the basis of bids (which in turn are based on cost-delay curves). Konana et al. (2000) propose priority pricing policy (which is similar to MARIPOSA) for managing databases. However, their approach accounts for the dynamic environment by considering not only the current but also the future arrival rates. They propose an implementation methodology for their theoretical model. Using simulation they show that their approach can increase efficiency.

In summary, the primary purpose of the research on network pricing is on relating the performance of the pricing schemes on network loads and congestion, and understanding how price discrimination strategies can affect optimal access times/costs and increase overall consumer welfare. Our model, instead, focuses on the pricing of information product that uses the networks as an infrastructure. We do not consider the factors of congestion externality or the capacity constraints in the supply side. Our formulation considers the optimality of various pricing schemes that firms use to offer information services. Similar to the literature in network pricing, our model also considers the implications of delay on optimal pricing schemes. However, the reason for the delay is entirely different in our model from the research in network pricing. In particular, we do not focus on delay arising out of network congestion but instead focus on the potential delay that a consumer faces when s/he is searching for information in a database. This delay may be related to the size of the database, the complexity of the query, the sophistication of the user in formulating queries, etc. In this context, unlike prior literature on network pricing, we consider the impact of risk averseness of consumers and its impact on optimal pricing.

Pricing of Information Goods. Our research is also related to the recent literature which focuses on the pricing and distribution of information goods. This research also does not consider network congestion issues and primarily focuses on the selling of the information good itself. Dewan et al. (2000) examine whether producers of proprietary databases should distribute their information products only through proprietary channels or also through the Internet. They conclude that if the market is sufficiently competitive then database manufacturers may find it profitable to use both channels. Bashyam (2000) examines whether information producers should offer their products online so that they can purchase the products on the basis of usage or on a CD-ROM which can be purchased by paying a fixed fee. In a duopoly setting, he shows that in equilibrium one firm will offer its products online while the other firm will offer CD-ROMS. The focus of these papers is however on channel selection and not on pricing issues. Therefore, they do not consider various pricing mechanisms that are observed in practice and their optimality in different scenarios. These models do not account for the uncertainty consumers face in searching and with respect to prices.

Finally, Bakos and Brynjolfsson (1999, 2000) also examine how firms should distribute and price different information goods. In particular, they examine whether firms should bundle different information goods and charge one bundle price. Their results show that firms can benefit by bundling if the marginal cost of production of goods is low. Their focus is on the optimality of the bundling strategy when a firm offers a variety of information goods which is quite different from ours. They do not consider the issues of consumer uncertainty in the search process. Also, they restrict the firm to charge either on the basis of per unit good or on the basis of a fixed price for multiple units.

In summary, in contrast to the earlier research, our paper focuses on the optimality of various pricing mechanisms that are observed in practice. Since our primary focus is on the pricing of the database, we ignore network congestion issues. We consider a situation in which consumers are searching for information from a database. We assume that there is some uncertainty about the time that it will take for the information to be retrieved. In this context, we evaluate the conditions under which a firm may wish to charge on the basis of connect time, successful search, or offer subscription plans. Our research is also different from the prior literature in that it explicitly incorporates the impact of uncertainty on consumers' purchase decision.

3. Model

Consumers typically enter the database with the intent of finding specific information. However, consumers only have an estimate of the time it will take them to successfully retrieve the desired information. Let us assume that a consumer has a valuation v for a unit of information which takes him or her time t to retrieve where t is a normally distributed random variable with mean μ and variance σ^2. The unit of information could be a page or a report—a discrete, logical unit that provides the desired information. Note that μ and σ are related to many consumer-specific and firm-specific characteristics. For example, a firm may have an inefficient search engine which increases the expected time for successful retrieval of information. On the other hand, the ability to search and the complexity of the search query may vary across consumers and can determine the time that it takes a consumer to retrieve information. Although our model does not explicitly account for each of these separately, it does so implicitly by acknowledging that the parameters μ and σ varies across the population. We assume that the consumers are risk averse with a utility function of the form $-\exp(-ky)$, where k is a risk parameter and y is the surplus (which may be a random variable). We assume, for now, that an individual consumer's demand is inelastic. This is reasonable in the context of information goods and within the price ranges encountered in practice. We relax this assumption later and show that most of our results continue to hold even if demand is elastic.

We assume that the online server (which is an intermediary) is risk neutral and pays c to the database producer for each unit of time that the consumer uses the database.[3] This assumption is in line with the actual practice as discussed before (see Hawkins 1989, p. 21, Abels 1996). We assume that the online server is a monopolist. This assumption is true for a large proportion of the market (see Table 2). Later, we also reexamine these issues in a duopoly setting. We will consider the three most common pricing schemes discussed earlier: connect-time-based pricing, flat rate per successful search pricing (which we will call search-based pricing), and subscription pricing. We will also examine the possibility when the firm can offer consumers the option of either going with a combination plan or with a subscription plan which allows unlimited use of the database. We denote the subscription fee by F, the connect-time fee by p_c and search-based fee by p_s. If the online server charges consumers a price p_c on the basis of connect time, then a consumer who uses time t pays a total price tp_c. However, if the server uses a search based pricing strategy the consumer pays a fixed price p_s per "successful" search. We define successful search as one which results in viewing/downloading of desired information from the database.[4]

[3] We are not explicitly considering the payments made to the packet switch network/telecom carrier by the server. Since these payments are also based on the connect time, they may be included in c without any loss in generality.

[4] Of course, the value that a consumer receives depends on the "fit" between the downloaded information and the consumers' tastes. This "fit" can often be only assessed after viewing the information. For the moment, we assume that this fit is either perfect, i.e., has a valuation v or has zero valuation. Later, in §5, we will discuss how relaxing this assumption will affect our analysis.

Table 2 Online Server Pricing and Channel Structure

	Channel structure and server pricing	1997 %	2000 %	No. of Databases
1	Databases available only through a single server			358
	Pure connect-time pricing	71.0%	62.8%	
	Pure subscription-fee pricing	2.2%	5.9%	
	Combination of search based/ subscription-fee	10.1%	14.5%	
	Independent contracts and other pricing basis	16.6%	16.6%	
2	Databases available through two servers			73
	Both servers connect-time pricing	13.6%	8.2%	
	One connect-time; other search/ subscription-fee	45.2%	46.6%	
	Other combinations	38.4%	42.5%	
	Both pure subscription-fee pricing	2.7%	2.7%	
3	Databases available through three or more servers			66
	All connect-time pricing	1.5%	1.5%	
	One connect-time; others different	25.8%	24.2%	
	Other combinations and mixed schemes	72.7%	74.2%	

Note. Data based on survey of Gales' Online Directory of Databases and Online Servers (1997 and 2000). Databases with free access are excluded from the above table.

To analyze the optimal pricing scheme, we must recognize that the various parameters in the model, i.e., μ, and v, are likely to vary across consumers. To keep the analysis tractable and to highlight the differential impact of variation in each parameter, we will conduct the analysis by allowing one parameter to vary across consumers and keep the other parameters fixed. Later, we examine the case when both μ and v vary.[5]

3.1. Valuation v Varies Across Consumers

We first consider the case where σ and μ are the same across the population but v is distributed across the population according to a continuous and twice differentiable probability distribution function $f(v)$. To ensure that the profit functions are quasi-concave and an optimal strategy exists, we assume that the distribution function is such that the hazard rate for the function is nondecreasing.[6] This technical condition is satisfied by a variety of distribution functions including uniform, chi-square, exponential, Laplace, logistic, normal, and the extreme-value distribution.[7] Furthermore, under some parameter restrictions the condition is also satisfied for the Weibull distribution, Beta distribution, and the log-logistic distribution.

The online servers can offer (a) a connect-time-based price, (b) a flat rate per successful search price, or (c) a combination of (a) and (b). We also consider the case when the firm can offer consumers a menu of options in which they can either choose to pay a subscription fee for unlimited usage or pay a combination of search-based or connect-based price. Since the pricing trends in the industry have transformed from a pure connect time to other forms, we first analyze pure connect-time price versus flat rate per search price.

If the online server charges a price p_c per unit time (i.e., a connect-time strategy), and the consumer enters the market then the consumer receives utility v. The consumer's costs are $p_c t$, which is normally distributed with mean μp_c and variance $p_c^2\sigma^2$, where μ and σ are the mean and standard deviation of time to retrieve t. In this case, the consumer's certainty equivalent from entering the market is

$$CE(p_c) = v - p_c\mu - kp_c^2\sigma^2/2, \tag{1}$$

where k is the risk parameter. The consumer will purchase only if his/her certainty equivalent is greater than zero.

Assuming that all consumers either buy one unit or do not purchase and normalizing the market size

[5] We have also examined the case when σ varies across the population. In terms of pricing implications, the results are the same as those for the case when v varies. To preserve space, we do not present this analysis in the text but instead present it in Technical Appendix C. All appendices can be viewed at http://mansci.pubs.informs.org/ecompanions.html or requested from authors.

[6] Lemma 1 in Technical Appendix A shows existence and uniqueness of optimal pricing schemes under this condition.

[7] In the context of profit maximization, this condition translates into the requirement that for each unit increase in price, the fraction of the remaining customers that the firm loses is increasing. This ensures that the profit function is quasi-concave and an optimal price exists.

to one, the total demand for the online server under connect-time strategy is

$$D_c = \int_{p_c\mu + kp_c^2\sigma^2/2}^{\infty} f(v)\,dv. \tag{2}$$

The online server's expected profits are

$$\Pi_c = \mu(p_c - c)\int_{p_c\mu + kp_c^2\sigma^2/2}^{\infty} f(v)\,dv. \tag{3}$$

Now consider the case when consumers are charged on the basis of successful searches of information. In this case, consumers bear no risk and therefore the demand for the server at price p_s per unit of successful search is given by

$$D_s = \int_{p_s}^{\infty} f(v)\,dv, \tag{4}$$

and therefore total profits are given by

$$\Pi_s = (p_s - \mu c)\int_{p_s}^{\infty} f(v)\,dv. \tag{5}$$

We have the following proposition.

Proposition 1. *If consumers vary in terms of v and if* $\sigma, k > 0$ *then the server makes higher profits by charging on the basis of successful searches rather than on the basis of connect time, i.e.,* $\Pi_s^* > \Pi_c^*$.[8]

The reason for the proposition can be understood by noting that a search-based strategy essentially removes any risk for the risk-averse consumers. Thus, the result says that a risk-neutral firm can always benefit by insuring the risk-averse consumers. This insurance allows the server either to charge higher prices or to get a substantially larger market (or both). Thus, it is not surprising that the major need cited in the push towards search-based strategy was increasing the market size and the need to move the risk-averse consumer from *saving time* to *using more information* (Garman 1990).

Proposition 2. *If the server uses a search-based pricing, its prices increase with c and* μ. *If the server charges a connect-time-based price, its price increases if c increases. Further, the server's* (*connect-time-based*) *price increases as* μ *decreases*.

[8] Proofs of all propositions are in Technical Appendix A.

The proposition shows that search-based prices decrease as the intensity of competition among database providers increases (and therefore c decreases). The results also show that as μ decreases, a server using a search-based strategy will decrease its prices. However, if the server charges on the basis of connect time, its prices will increase if μ decreases. Note that μ decreases steadily as technology advances: when consumers have faster access to the database, e.g., via a faster network or modem, or when consumers become more efficient searchers with the availability of PC-based search algorithms. With rapid technological advances and consumer sophistication, the results show that one will observe vastly different impact of such developments on pricing of online servers with different pricing schemes. The servers using connect-time-based strategies have been forced to increase their connect-time prices to keep up with the technological changes (Garman 1995).

Proposition 3. *If consumers vary in terms of v and the server uses a search-based pricing, its profits always increase if* μ *decreases. However, if the server uses a connect-time-based pricing, then its profits increase as* μ *decreases if and only if* $(p_c^* - c)p_c^*\sigma^2 k < \mu c$. *If a server uses a connect-time-based pricing, then for low values of* μ, *its profits always decrease as* μ *decreases*.

The results with respect to Π_s are intuitive since a server which uses a search-based strategy can only benefit when the average search time decreases, since this implies that the costs for the server are lower. However, this is not always true when the server is charging consumers on the basis of connect time, since a decrease in the search time not only decreases the costs but also the revenues. Using the inequality in the proposition, we can see that the result implies that if the consumers are highly risk averse (k is large) and/or if there is large uncertainty in the search process (σ^2 is large), then the firm charging a connect-time price would be worse off as μ decreases.

The result therefore suggests that servers using a search-based pricing will find increasing consumer sophistication and higher connection speeds, as positive developments. In fact, these servers may find it profitable to invest in technologies to make search faster, allow faster access, and train consumers in making efficient database searches. They will also have the option of passing on the savings with technological advances to the database providers or consumers. However, servers using a connect-time-based strategy may find it counterproductive to invest in such technologies. Such developments will mostly be consumer initiated as history indicates (Hawkins 1989).

The online server also has the option of providing a consumer a menu of options. We consider the case in which the consumers can choose from a combination of connect-time-based and search-based pricing, and a subscription-based pricing. Assume that each consumer has demand for x_i unit of information during the given period. We assume homogeneity of value of each unit to the consumer. To keep the analysis tractable, assume that there are two types of consumers: high-demand consumers with demand x_h and low-demand consumers with demand x_l where $x_h > x_l$. Further, without any loss in generality, we assume that each segment is of equal size. The certainty equivalent of a consumer under the combination of connect-time-based and search-based pricing will be

$$CE(p_c, p_s) = x_i(v - p_c\mu - kp_c^2\sigma^2/2 - p_s), \tag{6}$$

where x_i is the demand of the consumer. Now the consumer will only consider the combination of a search and connect-time scheme only if

$$v \geq p_c\mu + kp_c^2\sigma^2/2 + p_s. \tag{7}$$

Let the subscription fee charged by the online server be F for the given period. The consumer will consider a subscription fee only if

$$v \geq F/x_i. \tag{8}$$

Further, a consumer will prefer to buy a subscription rather than a connect-time/search-based scheme if his/her certainty equivalent from buying a subscription is greater than by paying using a combination of search and connect time. This condition reduces to

$$F \leq (\mu p_c + p_s + kp_c^2\sigma^2/2)x_i. \tag{9}$$

Note that since $x_h > x_l$, it follows that if low-demand consumers prefer the subscription then so will the high-demand consumers, but not vice versa. In this case, we can establish the following result.

Proposition 4a. *If consumers vary only in terms of v and the distribution $f(\cdot)$ is the same across both the high-demand and the low-demand consumers, then the server makes optimal profits by only charging on the basis of successful searches.*

Proposition 4b. *If the distribution of v for the low-demand consumers f_l is dominated by the distribution f_h for the high-demand consumers using the hazard rate ordering criterion, then the firm makes weakly higher profits by offering a subscription plan along with a search-based price.*[9] *Further, for sufficiently large x_h offering a subscription plan leads to strictly higher profits.*

The results suggest that if only v varies across the population, then in general, a firm would offer a combination of subscription price and a search-based pricing scheme. However, a subscription is a useful addition to offering a search-based price only under certain circumstances. Proposition 4b provides the condition under which selling a subscription, in addition to search-based pricing, can increase profits. The condition specified in the proposition implies that the market consists of low-demand consumers who have relatively higher valuations than the high-demand consumers. In such a case, the firm can increase its profits by selling a subscription plan to the high-demand consumers and a search-based price to the low-demand consumers. If the high-demand consumers have low valuations, then this can be profitable while being incentive compatible (so that the high-demand consumers do not find the combination plan more attractive). The condition specified in the proposition can hold under many buying situations. For example, some customers may require access to a database frequently, although the value of an individual piece of information is relatively low.[10] However, some firms only infrequently need to

[9] The hazard rate ordering criterion is a little stronger than the condition that f_l stochastically dominates f_h (see, for example, Ross 1983).

[10] This is possibly true for universities.

access the database but may have a higher value for each piece of information. In such circumstances, the online server can benefit from offering a subscription plan to the heavy volume customers while offering a search-based pricing to the low-volume consumer.

3.2. μ Varies Across Consumers

Now we consider the case when consumers are differentiated in terms of μ. Again, we assume that μ is distributed according to a distribution $f(\mu)$ and each consumer has the same v, σ, and k. If the server uses a search-based strategy and $p_s \leq v$, then its profits are

$$\Pi_s = p_s - c\int_0^{\infty} \mu f(\mu)\, d\mu = p_s - c\bar{\mu}. \tag{10}$$

Profits in the case of connect-time-based pricing would be

$$\Pi_c = (p_c - c)\int_0^{\psi_2(p_c)} \mu f(\mu)\, d\mu, \tag{11}$$

where $\psi_2(p_c) = v/p_c - \sigma^2 p_c k/2$ and represents the value of μ for the customer who is indifferent between buying and not buying. From the profit equations, it is clear that if $\bar{\mu}$ is large, a search-based strategy can be dominated by a connect-time-based strategy. The reason is that a connect-time-based strategy reduces servers' costs by making it infeasible for high-cost customers (those with large μ) to buy the information product. Since a large μ could be a result of lower access speeds and/or inefficient search algorithms, one might be tempted to recommend a connect-time strategy in an environment of low access speeds and poor search technologies. However, the argument that it reduces consumers' risks still favors a search-based strategy. Furthermore, a search-based strategy allows cross-subsidization of the high-cost customers (large μ) by the low-cost customers (small μ). This cross-subsidization is not feasible when a server uses a connect-time-based pricing. Thus, there are different aspects favoring the use of connect-time and a search-based strategy. The combined effect of these aspects is ambiguous and therefore neither type of pricing dominates in all cases. Nevertheless, lowering values of μ with increasing access speeds because of technological advances would favor the use of a search-based strategy over a connect-time-based pricing.

Notice that as long as the consumer with average waiting time is served under a connect-time-based strategy, a search-based strategy is optimal. To see this, note that if a consumer with $\mu = \bar{\mu}$ is served, then $v > \bar{\mu} p_c^*$. In this case, we have

$$\Pi_c^* = (p_c^* - c)\int_0^{\Psi_2(p_c^*)} \mu f(\mu)\, d\mu, \tag{12}$$

$$< (p_c^* - c)\int_0^{\infty} \mu f(\mu)\, d\mu = (p_c^* - c)\bar{\mu} = \Pi_s(\bar{\mu} p_c^*), \tag{13}$$

$$\leq \Pi_s^*. \tag{14}$$

Thus, if c is not too large, then this suggests that a search-based strategy will dominate. The next proposition shows that the cost c does play an important role in determining which pricing scheme is optimal.

Proposition 5. *As c increases $\Pi_c^* - \Pi_s^*$ increases. Furthermore, there exists a c^* such that if $c \leq c^*$, a search-based strategy dominates a connect-time-based pricing strategy and a connect-time strategy dominates if $c > c^*$.*

Note that the proposition implies that as the competition among database providers increases, a monopolist online server will find it more attractive to switch to a search-based pricing. However, if the database provider is a monopolist, then it is likely that c will be high. This provides one possible explanation as to why we observe predominantly connect-time based strategies when monopolist database providers sell through single servers (see Table 2).

Now let us consider the case when the server also offers a subscription plan. We examine whether offering such a plan could increase profits. If the server offers a subscription plan and also a combination of connect-time and search-based pricing schemes, then each consumer must choose between the two options. A consumer will prefer the subscription-fee pricing rather than a combination pricing only if

$$F \leq x_i(p_s + \mu p_c + \sigma^2 k p_c^2/2). \tag{15}$$

In other words, if

$$\mu \geq \frac{1}{p_c}(F/x_i - p_s - \sigma^2 k p_c^2/2) \equiv \mu_i^*. \tag{16}$$

Note that the equation suggests that if the server offers a subscription-fee pricing it is more likely to attract high-demand consumers *and* high-cost consumers. If there are a large number of high-cost consumers, then a server must be careful in offering a subscription-fee pricing.

The consumer will buy a subscription only when such a purchase gives him a positive utility. In other words,

$$v \geq F/x_i. \tag{17}$$

This equation again suggests that high-demand consumers are more likely to find subscription-fee pricing profitable.

A consumer will prefer the combination pricing scheme if[11]

$$\mu < \mu_i^*. \tag{18}$$

The consumer will buy the combination scheme if he finds it to be profitable. That is, if

$$v \geq p_s + \mu p_c + \sigma^2 k p_c^2/2. \tag{19}$$

In other words, if

$$\mu \leq \frac{v - p_s - \sigma^2 k p_c^2/2}{p_c} = \hat{\mu}. \tag{20}$$

We have the following result:

PROPOSITION 6A. *If the distribution of μ is the same across both the high- and the low-demand consumer, then the firm does not make any additional profits by offering a subscription plan.*

PROPOSITION 6B. *Suppose the distribution $f_l(\cdot)$ for the low-demand consumers stochastically dominates (in the first-order sense) the distribution $f_h(\cdot)$ for the high-demand consumers. Further, if $\mu_H c < v < \bar{\mu}_l c$, where μ_H is the upper limit of the support of the distribution $f_h(\cdot)$ and $\bar{\mu}_l$ is the mean of the distribution $f_l(\cdot)$, then if x_h/x_l is sufficiently large, the firm is strictly better off by offering a subscription plan.*

The proposition shows that if the high-demand and the low-demand consumers have similar costs then the firm will not benefit from offering a subscription. However, if the high-demand consumers are either expert searchers or have less complicated search needs and consequently have low costs (i.e., low μ), then the firm can benefit by offering a subscription plan provided there is a sufficient separation among the demand levels of the two groups (i.e., x_h/x_l is large). In such a situation, the firm can offer a subscription to the high-demand and relatively low-cost consumers and offer a combination plan to the low-demand, high-cost consumers.[12]

What are the implications of the above results? If the online server costs c were to come down, either through competition upstream or through the databases adopting a profit-sharing or fixed-fee mechanism (making cost per connect time, $c = 0$), then the servers will not find it profitable to charge connect-time-based prices. Further, since in this case the costs of the high-demand and the low-demand consumers will be similar (since $c \sim 0$), using Proposition 6a, it follows that the firm will not benefit by offering a subscription plan. Thus, in the longer run, search-based pricing schemes are likely to be more attractive for the servers in the future. This result also has implications for pricing for database producers. For database producers, the major portion of their costs are fixed costs and are independent of the time that the database is used. Thus, the total channel profits when the server charges a combination price is given by

$$\Pi^T = \int_0^{\psi(p_c, p_s)} p_c f(\mu)\, d\mu - C, \tag{21}$$

where $\psi(p_c, p_s) = v/p_c - p_s/p_c - k\sigma^2 p_c/2$ and C is the fixed cost of producing the database. It is clear that the total channel profits are maximized when the server charges $p_c^* = 0$ and only charges a search-based price. Thus, if the database producers were to move to a profit-sharing arrangement and charge $c = 0$,

[11] Since we will assume that the distribution functions for μ are continuous with no mass points, it does not matter how we break the ties.

[12] There are also other situations where a subscription plan can dominate both the search-based and connect-time pricing. This can happen when the firm offers multiple information products to a consumer who may have different valuation for each product. In this case, as is well known in the bundling literature, bundling of goods by offering a subscription plan can be optimal (see, for example, Bakos and Brynjolfsson 1999, 2000). In this paper, however, we do not focus on these bundling issues.

Table 3 Impact of Model Parameters

Parameter conditions	Optimal pricing	Profit implications	Market/segment strategy implications
v varies across consumers			
μ decreases	Search based	$\Pi_s^* \uparrow \forall \mu$ $\Pi_c^* \downarrow$(low values of μ)	Favors search-based strategy
c increases	Search based	$\Pi_s^* \downarrow$; $\Pi_c^* \downarrow$ $\Pi_c^* \downarrow$(low values of μ)	Favors connect-time strategy
f_l dominates f_h using the hazard rate criterion	Subscription and search based	Offering subscription can increase profits	Subscription to high-demand consumers and search-based to low-demand consumers
μ varies across consumers			
c increases	Combination plan	$(\Pi_c^* - \Pi_s^*) \uparrow$	High-cost consumers are forced out using connect-time pricing
v increases	Combination plan	$\Pi_s^* \uparrow$; $\Pi_c^* \uparrow$	Total market expands
$F_h(\cdot) \geq F_l(\cdot)$	Subscription and combination plan	Offering subscription can increase profits	Subscription to high-demand consumers and combination to low-demand consumers

then search-based pricing will be more commonly observed. (Table 3 summarizes the results and their implications).

3.3. Both μ and v Vary Across Consumers

In this section, we briefly examine the case when the parameter v and μ across consumers. In general, the time that it takes the consumer to search may be related to his or her valuation. For example, high-valuation consumers have more complicated search needs and therefore may have higher costs. In other words, μ and v may be positively correlated. Alternatively, high-valuation consumers may also be more knowledgeable and therefore may require lower search times. In this case, μ and v would be negatively correlated. In general, therefore it is difficult to specify the relationship between μ and v a priori.

Figures 2, 3, and 4 illustrate how different relationships between the two variables can affect the optimality of search-based versus connect-time pricing. The analysis in the figures is for the special case when μ and v are uniformly distributed from $(0, 1)$ and when $k = 0$. Figure 2a shows that when v and μ are independent, then search-based pricing is optimal while, for higher costs, connect-time pricing becomes more profitable. This is consistent with our results in Proposition 5. Interestingly, Figure 2b shows that although the server might prefer search-based strategies, a connect-time-based strategy leads to higher consumer and social welfare. This difference however reduces as c increases.

Figure 3a considers the case when μ and v are perfectly positively correlated. In other words, a consumer with high valuation also has correspondingly higher costs. In this case, we find that a connect-time strategy dominates the use of a search-based strategy. This is intuitive since the firm is able to

Figure 2a Impact of c on Profits Under Search-Based and Connect-Time Pricing: v and μ Are Independent

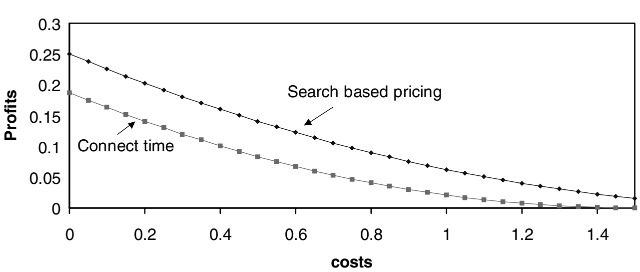


Figure 2b Impact of c on Consumer Surplus and Social Welfare Under Search-Based and Connect-Time Pricing: v and μ Are Independent

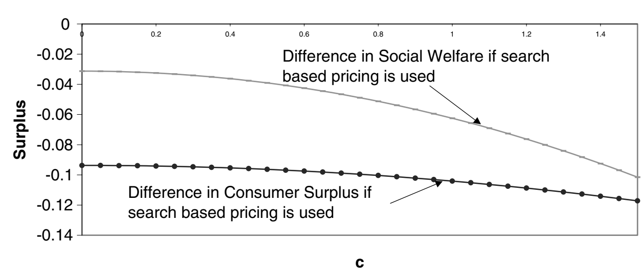


Figure 3a **Impact of c on Profits Under Search-Based and Connect-Time Prices: v and μ Are Positively Correlated**

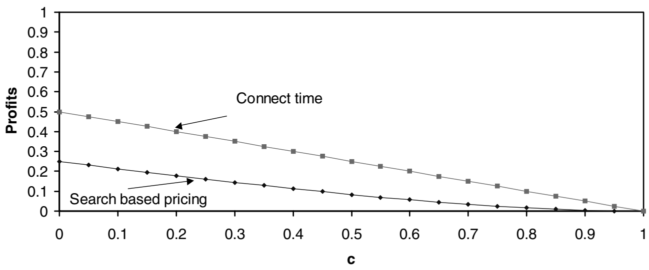
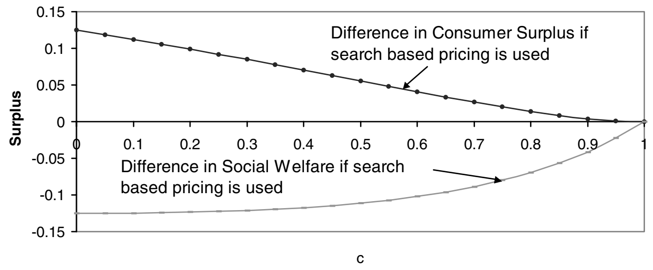


Figure 3b **Impact of c on Consumer Surplus and Social Welfare Under Search-Based and Connect-time Pricing: v and μ Are Positively Correlated**



better price discriminate among the high-valuation and the low-valuation consumers by using connect-time prices which are directly related to consumers' valuations. *This result therefore suggests that in markets where the high-valuation customers have more complicated needs and require more online time to search, the firm should charge on the basis of connect time*. An example of this could be DIALOG's DataStarWeb pricing strategy. The database caters to the needs of professionals such as market researchers and consultants. These users are likely to have higher valuation for the information and also have complicated needs such as verifying information from several sources which lead to larger search times as compared to novice users. DataStarWeb's pricing is similar to a connect-time-based pricing. Figure 3b shows that while connect-time-based pricing results in higher welfare, it leads to lower consumer surplus. This is also intuitive, since in this case, the firm is able to extract more consumer surplus when it uses a connect-time-based pricing.

Figure 4a shows that when μ and v are negatively correlated, the firm should charge a search-based price. This is because in this case the firm can afford to extract higher revenues from the high-valuation customers by charging on the basis of successful searches while at the same time it can eliminate the high-cost consumers by charging a high p_s. *In other words, if the market is such that the high-valuation customers are expert searchers who require less time than the low-valuation novice consumers then the firm should charge on the basis of successful searches.* This could well the reason for LEXIS law database's move to a search-based pricing strategy in the late 1980s. Law firms have higher valuation for the information and are typically experienced searchers (as they employ professional searchers) as compared to the low-valuation novices who use the database on a sporadic basis. Figure 4b shows that the use of search-based strategy in this case leads to lower consumer surplus. This is consistent with intuition, since a search-based strategy in this case enables the firm to extract more consumer surplus.

Figure 4a **Impact of c on Profits Under Connect-Time and Search-Based Pricing: v and μ Are Negatively Correlated**

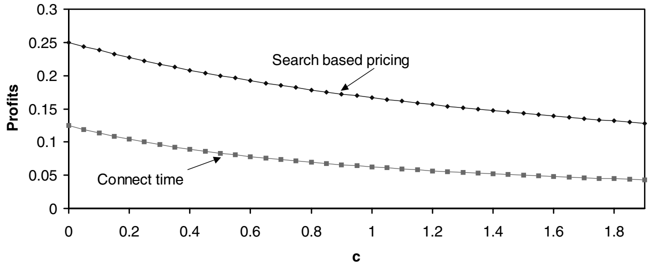
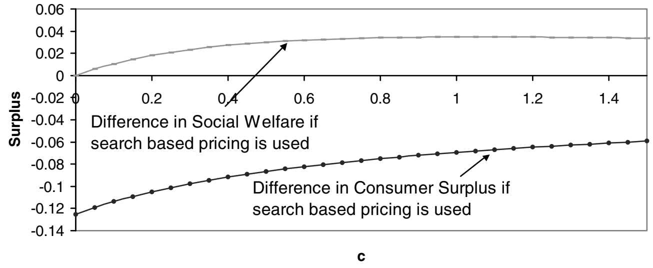


Figure 4b **Impact of c on Consumer Surplus and Social Welfare Under Search-Based and Connect-Time Pricing: v and μ Are Negatively Correlated**



4. Competition Among Online Servers

In this section, we examine how competition among two online servers affects the pricing strategy that they choose. We assume the following decision structure. In Stage 0, servers choose a pricing scheme. We will restrict the types of pricing schemes to pure connect-time-based pricing or a pure search-based pricing. In Stage 1, after observing each other's pricing strategy, each server decides on the specific prices. This decision structure is reasonable since servers find it much easier to change specific prices than changing the basis on which a server charges prices, i.e., its pricing strategy. For example, to change its pricing policy from connect time to flat fee, a firm needs to educate its consumers about this change (as their behavior depends on the pricing structure), forecast the impact of such a change, and also need to change its accounting systems (Hawkins 1989, p. 21; Garman 1990, pp. 30–31; Basch 1998). Also, while the change of specific price may be a tactical decision, changing price policies is a higher-level strategic decision.[13]

In Stage 2, consumers make their purchase decisions. We analyze the impact of competition by first considering the case when consumers are differentiated in terms of v and then considering the case when consumers are differentiated in terms of μ. We will assume that the two online servers are undifferentiated since they provide the same information to the consumers. This assumption is realistic as in today's market multiple servers can provide access to the same databases. We will use the concept of subgame perfectness to derive the equilibria.

First, we consider the case when consumers have different valuations v for the product. As before, we assume that these valuations are distributed across the population according to a differentiable probability density function $f(v)$. We have:

Proposition 7. *Assume that firms choose pricing policy and then choose specific prices, $c > 0$ and consumers are differentiated only in terms of v. Then there are three pure strategy equilibria of the game—a symmetric equilibrium in which both servers offer a search-based policy and two asymmetric equilibria in which one server offers search-based pricing and makes positive profits while the other server offers connect-time-based pricing and makes zero profits.*

The proposition shows that an asymmetric equilibrium can emerge with identical servers choosing different pricing schemes. Further, the asymmetric equilibria Pareto-dominates the symmetric equilibrium and is likely to be the focal equilibria which is observed (see Harsanyi 1964, Fudenberg and Tirole 1983, 1985, Harsanyi and Selten 1992 for justification in the use of pareto-dominance as a reasonable criterion for equilibrium selection). Further, even when servers are undifferentiated, at least one server can make positive profits. Notice that the proposition shows that unlike the monopoly case where a connect-time strategy will not be observed if consumers only differed in term so of v, once we consider competition, we may observe that some firms will continue to offer connect-time-based pricing.

Next, we consider the case when consumers are differentiated in terms of μ. It is clear that when both servers choose the same pricing strategy, they make zero profits. We therefore consider the case when one server chooses a search-based pricing scheme and charges p_s while the other server charges a price p_c on the basis of connect time. Clearly, we need $p_s \leq v$. It can be shown that a consumer of type μ prefers a search-based scheme over a connect-time-based scheme if

$$\mu \geq \frac{p_s}{p_c} - \frac{\sigma^2 p_c^2 k}{2p_c} = \psi_3(p_s, p_c). \tag{22}$$

Note that by selecting different pricing schemes, the servers segment the market into two types. Consumers with high search times prefer the search-based scheme and consumers with low search times prefer

[13] What is important for our analysis is that the firm should be able to change specific prices more frequently than its pricing policies. Hawkins (1989) notes that when a database producer wanted to change price structures, online servers "protested that the lead time of only five months would not give them time to reprogram their complex accounting systems." A director of online server ESA-IRS reveals that it took them a year to change their pricing policy (Garman 1990). Thus, our assumption is justifiable in this context. Nevertheless, this is an important assumption without which an equilibrium may not exist (see for example Gupta et al. 1996).

to pay connect-time-based prices. The profits of the two servers are given by

$$\Pi_s = \int_{\psi_3(p_s,\, p_c)}^{\infty} (p_s - \mu c) f(\mu)\, d\mu, \tag{23}$$

$$\Pi_c = (p_c - c) \int_0^{\psi_3(p_s,\, p_c)} \mu f(\mu)\, d\mu. \tag{24}$$

The first-order conditions for finding the Nash equilibrium prices are:

$$\int_{\psi_3(p_s,\, p_c)}^{\infty} f(\mu)\, d\mu - f(\psi_3(p_s,\, p_c)) p_s/p_c + c\psi_3(p_s,\, p_c) f(\psi_3)/p_c = 0, \tag{25}$$

$$\int_0^{\psi_3(p_s,\, p_c)} \mu f(\mu)\, d\mu - (p_c - c)\psi_3(p_s,\, p_c) \times f(\psi_3)(p_s + \sigma^2 k p_c^2/2)/p_c^2 = 0. \tag{26}$$

Proposition 8. *Assume that firms first choose pricing policies and then choose specific prices and consumers are differentiated in terms of μ then if $\sigma > 0$, $k > 0$, there exists an equilibrium in which both firms offer different pricing policies. Further, there exist parameters such that both firms make positive profits by choosing different pricing strategies, and consequently in equilibrium, both firms choose different pricing policies.*

The proposition shows that once we consider variation in μ, there are situations in which symmetric firms will choose asymmetric strategies and will make different profits. The reason for this result is that identical servers can differentiate themselves by using different pricing schemes which segments the market on the basis of μ. In other words, the different pricing schemes can serve as mechanisms for the servers to differentiate themselves. This argument will also hold if we allow firms to offer subscriptions. As long as the products are undifferentiated, firms will want to offer different pricing schemes. Thus, pricing schemes can be used by online servers as a means of differentiation. This provides some explanation for the variety of schemes that abound in the market.

The above result also explains the entry of West Publishing Company's online venture as a direct competitor to Mead Data Central's LEXIS (Berss 1993). Mead moved away from connect-time-based pricing in the late 1980s, and their pricing strategy of resource unit pricing is quite similar to search-based pricing. West entered the market offering a purely connect-time-based pricing and targeted sophisticated users who can perform a computer search quickly and then go offline to use West's books. West was able to capture a significant share of the sophisticated users (The market share of Lexis dropped from 95% in early 1980s to 60% in 1993; see Berss 1993). Proposition 8 suggests that both these competitors can coexist and make profits while offering different pricing schemes.

Now we consider the impact of c on the profits of undifferentiated firms.

Proposition 9. *If both servers use different pricing policies and $c = 0$ then, in one equilibrium of the game, $p_c^* = 0$, $p_s^* = 0$ and both firms make zero profits. However, if $c \in (2\bar{\mu}/(\sigma^2 k), \sqrt{2v/(\sigma^2 k)})$, then in any pure strategy equilibrium of the pricing game, at least one firm makes positive profits.*

The proposition thus shows that there are situations in which an increase in the costs of the servers can lead to higher profits. This result is in contrast to standard results in the channels literature. The reason for this result, however, can be understood by noting that when $c = 0$, firms cannot differentiate themselves by offering different pricing schemes. This is because they both charge zero prices regardless of the pricing scheme and consequently make zero profits. However, when $c > 0$, the firms are able to differentiate themselves by offering different pricing schemes. In particular, the firm offering connect-time-based price is more attractive to consumers who have low costs while a search-based pricing policy is more attractive to high-cost consumers. Consequently, the firms are able to effectively segment the market using different pricing schemes and therefore make positive profits when $c > 0$.

The above proposition however does not show that firms' profits can *monotonically* decrease with c. We illustrate this possibility with a numerical example. Consider the case when $\sigma = 0$ and μ is uniformly distributed with range $(0, 1)$. In this case, we find that if $v \geq 4c/3$ then there is a unique stable equilibrium with $p_c^* = 2c$, $p_s^* = 4c/3$. The firm offering connect-time pricing serves $2/3$ of the market, and makes a profit $2c/9$ and the firm offering search-based pricing makes a profit of $c/6$.[14] Thus, this example shows that servers' profits rise monotonically as their marginal costs increase. Note that $c \to 0$ if competition among the upstream database producers becomes intense. Thus, the results suggest that competition among the upstream database providers can actually hurt the online servers. The intuition for the result is as follows: As $c \to 0$, the efficacy of the two pricing approaches in differentiating the servers decreases. Consequently, an increased competition among database producers can lead to more intense price competition among online servers and reduce profits. This result suggests that online servers also have a significant economic interest in ensuring strong copyright protection for databases and retain the barriers against easy competition among database providers. If this is not ensured, it can lead to lower profits not only for the database producers but also for the online servers. This indicates that the online server industry's general concern for their suppliers' (database providers') welfare and that their profit margins not squeezed, is right on track.

This result also has important implications for the pricing scheme that the database producers should use. Recall that we found that in the case where the database producer sells through a single server, it is optimal for it to enter into a profit sharing arrangement with the server and charge $c = 0$. However, when the database producer sells through multiple servers who compete with each other such a pricing scheme would lead to negative total profits. Thus, the producer should either sell through only one server or charge $c > 0$. Note that this result provides at least one reason why we will continue to observe connect-time pricing in the future even when network congestion is not an issue.

5. Model Extensions

In our preceding analyses, we have made many simplifying assumptions. In this section, we explore how our results will be affected if we relax some of these assumptions.

[14] The derivation of the numerical results are presented in Technical Appendix A available at http://mansci.pubs.informs.org/ecompanions.htm, or from the authors.

5.1. Demand Is Elastic

In our analysis, we have assumed that while the demand at the aggregate market level is elastic, each consumer's demand is inelastic. In particular, in our model a consumer either purchases one unit or does not buy. This assumption is similar to that made in many economic models (see for example, Bakos and Brynjolfsson 1999, Dewan et al. 2000). Nevertheless, in a more general setting, consumer demand is likely to be a function of prices. In Technical Appendix B, we examine how our results will be affected if we assume that consumer demand is elastic. We find that most of the results continue to hold in this case. However, the assumption does impact the results on the optimality of subscription plans. In particular, subscription plans are more likely to be optimal if consumer demand is elastic. This is because, when consumers have elastic demands then charging a per unit price leads them to decrease their consumption. In contrast, offering them a subscription plan enables them to consume the maximum amount. Therefore, a subscription plan can be beneficial in such circumstances since it enables the firm to extract more consumer surplus by increasing their demand. Note that this result is similar to the results on two-part tariffs in the literature.

5.2. Consumers Incur Waiting Costs

In our analysis, we have assumed that consumers do not incur any waiting costs. In reality, if the search takes a long time, consumers not only incur the costs because of the prices charged by the server but also costs in terms of their opportunity cost of time. To model this, we can assume that each consumer has a waiting cost w per unit time. In this case, note that the uncertainty for consumers is not totally removed even when the firm offers a search-based price. In general, the effective price that a consumer pays is changed by a term $\mu w + k\sigma^2 w^2/2$. Notice that this increase in effective prices is the same regardless of the pricing policy used by the firm. Therefore, qualitatively our results remain unchanged.

5.3. Quality of Search Affects the Valuation

In our model, we have assumed that the value of the information that a consumer retrieves is a priori known. In practice however, the valuation is likely to be a random variable. In other words, the consumer is better able to determine the worth of the information after observing it. The degree of uncertainty is likely to be related to both the consumer's expertise and also the quality of the search engine that the firm has. We can incorporate this in our formulation by assuming that v is a random variable which is normally distributed with mean $\bar{v}$ and variance σ_v^2. In this case, the effective price that a consumer pays will be increased by a term $k\sigma_v^2/2$. This increase will occur regardless of the pricing strategy used. Consequently, the qualitative nature of our results will remain unchanged.

6. Implications and Discussion

In this paper, we have examined the conditions under which the most commonly used pricing schemes—connect-time-based pricing, search-based pricing, and subscription-fee pricing—are optimal. It must be pointed out that our model provides a simple framework, as a start, to understand and analyze the complexities of pricing in the online market. We do not consider the aspects of negative externality such as network congestion, and supply side capacity constraints, which is an obvious limitation of the model. These factors indeed play an important role in furthering our understanding of the impact of different pricing policies. Nevertheless, our results show that variation in consumer expertise and valuation of information are also important factors in the choice of optimal pricing strategy. Thus, our analysis provides clear insights into some of the issues the online industry is grappling with. The rapid technological changes, the need to increase the size of the markets, the need for increased profits all favor the use of search-based pricing strategies. This said, the days of connect-time pricing are not over yet. In fact, we have shown that competition among information servers may make it optimal for them to choose different pricing schemes and therefore in equilibrium one may observe different pricing schemes being offered in the market as we see in the case of Lexis and West Publishing (Berss 1993). The important factor driving this phenomenon is the variation in consumer expertise, and given the risk-averse nature of consumers, the less sophisticated will always find it profitable to choose search-based/subscription-fee pricing and the sophisticated will find it profitable to choose connect-time pricing schemes. Our results also show that a monopolist data producer may have the clout to force a connect-time-based pricing strategy on the server by keeping the connect-time cost c high, this however could be a myopic strategy that does not auger well for the expansion of the market.

It is very useful to review our results in light of the existing market structure in the electronic information market. Table 2 provides the results of the survey (Table 1) conducted on the database industry based on the data available in *Gale's Directory of Databases and Online Servers* (Gale's Directory 1997, 2000) in a different form by including information on channel structure. Table 2 shows that 73% of the time databases are made available exclusively through one online server and 14% of the time through two online servers. This is not surprising given that the databases cover many diverse disciplines and operate within niche markets with almost no competition. Since the niche markets also tend to be small markets, there is not much incentive for competitors to enter the market. Thus, many of the databases are monopolists who choose to deal with one server exclusively. It is also evident from Table 2 when a database is made available exclusively through one online server, 71% of the time the pricing strategy tended to be connect time based in 1997. In 2000, this figure has dropped to 63% showing a tendency for moving away from connect-time strategy. Yet many monopoly database producers still insist on royalties based on connect time as they tend to view any other scheme, especially flat-rate pricing, with skepticism as they could involve "giving away valuable data for free" (Abels 1996). Our results provide some justification to the connect-time strategy used by these servers. Since connect-time charges c for these servers are relatively high and there are a large number of unsophisticated, high-cost consumers, a connect-time strategy enables these firms to keep the unprofitable customers out of the market. However, with increased customer sophistication, decreasing transmission costs and increasing power of the servers, it is likely that connect-time strategy will be less profitable to the monopolist servers who may want to use a search-based or subscription-based pricing. The decreasing trend in the usage of connect-time strategy may indicate that online servers are realizing this quickly.

Table 2 also shows that when a database is made available through two online vendors, only 10.9% of the time, both servers either adopt connect-time-based pricing or subscription-fee strategies (based on the 2000 data), the others then involve line servers using different pricing strategies. When the number of channels increase to three of more, pure pricing strategies are even less common. This supports our results that online vendors tend to use pricing strategies to differentiate themselves.

What will be the impact of technological advancement on the pricing strategies? Our analysis suggests that it will prove conducive to the offering of multiple strategies in the market. Increased modem speeds, sophisticated and fast searching algorithms can contribute to increased profits for servers using per search/subscription pricing, while advances that help consumers use offline search techniques will ensure that there is always a demand for connect-time-based information products among the sophisticated consumers. The net impact will be a rapid growth in the online information market, with more new users getting trained on servers using per search pricing schemes and, possibly, graduating over time to servers offering connect-time prices as they become more sophisticated.

Another important issue involves the continued viability of database providers. All market participants realize that the database providers are providing valuable products and must make profit to stay in business. Technological advances may drive users' expectations of prices to be lower, while data production costs may increase as they tend to be labor intensive. Increased resistance to price increases by database providers may jeopardize their continued viability. Increased competition among database providers may also bring down the costs to online servers but may be very detrimental to database providers. Our analysis show that increased competition among database providers may in fact hurt the online servers. Thus, online servers have a vested interest in ensuring strong copyright protection and keeping the barriers for competition among database providers high. This may be the key to the healthy growth of the industry.

Acknowledgments

The authors' names are listed in alphabetical order. Both have contributed equally to the paper. The authors thank the editor, the associate editor, and three anonymous reviewers for their detailed and insightful comments. P. K. Kannan acknowledges the financial support of Robert H. Smith Foundation for this work.

References

Abels, Eileen G. 1996. Pricing of electronic resources: Interviews with three vendors. *J. Amer. Soc. Inform. Sci.* **47**(3) 235–246.

Basch, Reva. 1998. DIALOG's new pricing structure and the DialUnit debacle. *Online* **22**(5).

Bashyam, T. C. A. 2000. Service design and price competition in business information services. *Oper. Res.* **48**(3) 362–375.

Berss, Marcia. 1993. Logging off Lexis (Mead Data Central Inc's legal database faces competition from West Publishing Co.). *Forbes* **151**(January 4) 46.

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

——, ——. 2000. Bundling and competition on the internet: Aggregation strategies for information goods. *Marketing Sci.* **19**(1) 63–82.

Cocchi, R., S. Shenker, E. Ester, L. Zhang. 1993. Pricing in computer networks: Motivation, formulation and example. *IEEE/ACM Trans. Networking* **1**(6) 614–627.

Communications Industry Report. 2000. 14th Edition, Veronis, Suhler & Associates, New York.

Dasgupta, Partha, Eric Maskin. 1986. The existence of equilibrium in discontinuous economic games, I: Theory. *Rev. Econom. Stud.* **172** 1–26.

Dewan, S., H. Mendelson. 1990. User delay costs and internal pricing for a service facility. *Management Sci.* **36**(12) 1502–1517.

——, M. Freimer, A. Seidmann. 2000. Organizing distribution channels for information goods on the internet." *Management Sci.* **46**(4) 483–495.

Fudenberg, D., J. Tirole. 1983. Capital as commitment: Strategic investment to deter mobility. *J. Econom. Theory* **20** 340–359.

——, ——. 1985. Preemption and rent equalization in the adoption of new technology. *Rev. Econom. Stud.* **52** 383–401.

Gale's Directory of Databases and Online Servers. 2000. http://galenet.gale.com/.

Garman, Nancy. 1990. Online pricing: An interview with Marino F. Saksida of ESA-IRS. *Online* **14**(1) 30–36.

——. 1995. The pricing conundrum. *Online*, (January-February) 6–7.

Gregory, Andrew. 1992. The inverted file: Maxwell online and the future. *Online* **16**(6) 10–11.

Gupta, Alok, Dale Stahl, Andrew B. Whinston. 1996. Economic issues in electronic commerce. *Readings in Electronic Commerce*.

Ravi Kalakota, Andrew B. Whinston eds. Addison Wesley, Reading MA. 197–227.

——, ——, ——. 1997. A stochastic equilibrium model of internet pricing. *J. Econom. Dynamics Control* **21** 697–722.

Harsanyi, J. 1964. A general solution for finite non cooperative games, based on risk dominance. M. Dresher, A. W. Tucker, eds. *Advances in Game Theory*. Annals of Mathematics, Study **52** Princeton, NJ. 627–650.

——, R. Selten. 1992. *A General Theory of Equilibrium Selection in Games*. MIT Press, Cambridge, MA.

Hawkins, David. 1989. In search of ideal information pricing. *Online* **13**(2) 15–30.

Konana, P., A. Gupta, A. B. Whinston. 1996. Pricing of information services using real-time databases: A framework for integrating user preferences and real-time workload. *Proceedings of the Seventeenth International Conference on Information Systems*. Cleveland, OH.

——, ——, ——. 2000. Integrating user preferences and real-time workload in information services. *Inform. Sys. Res.* **11**(2) 177–196.

Lilien, Gary, Philip Kotler, K. Sridhar Moorthy. 1992. *Marketing Models*. Prentice Hall, NJ.

Mackie-Mason, J., H. Varian. 1995. Pricing the internet. B. Kahin, J. Keller, eds. Public Access to the Internet. Prentice Hall, NY.

Mendelson, H. 1985. Pricing computer services: Queueing effects. *Comm. ACM* **28**(3) 312–321.

——, S. Whang. 1990. Priority pricing for the M/M/1 queue. *Oper. Res.* **38**(5) 870–883.

Nelson, Micheal L. 1995. Database pricing: A survey of practices, predictions, and opinions. *Online* **19**(6) 76–86.

O'Leary, Mick. 1993. Flat-rate Online: A new online era begins. *Online* **17**(1) 34–38.

Pack, Thomas. 2000. The changing role of the information vendor. *Online* **24**(2).

Ross, Sheldon. 1983. *Introduction to Stochastic Dynamic Programming*. Academic Press, Burlington, MA.

Rawley, Jennifer, David Butcher. 1996. Pricing strategies for business information on CD-ROM. *J. Inform. Sci.* **22**(1) 39–46.

Shenker, S. 1995. Service models and pricing policies for an integrated services internet. B. Kahlin, J. Keller, eds. Public Access to the Internet. Prentice-Hall, Englewood Cliffs, NJ.

Stidham, S. 1992. Pricing and capacity decisions for a service facility: Stability and multiple local optima. *Management Sci.* **38**(8) 1121–1139.

Stricker, Ulla de. 1995. Pricing information: The publisher's puzzle. *Searcher* **3**(June) 26–27.

Stonebraker, M., P. Aoki, A. Pfeffer, A. Sah, J. Sidell, C. Staelin, A. Yu. 1996. MARIPOSA: A wide-area distributed database system. *The VLDB J.* **5**(1) 48–63.

Takayama, A. 1974. *Mathematical Economics*. Cambridge University Press, Cambridge, UK.

Tenopir, Carol. 1991. The impact of CD-ROM on online. *Library J.* **116**(February) 61–62.

——. 1993. Flat-fee pricing and other choices. *Library J.* **118**(February) 58–59.

——. 1996. Has online made CD-ROM obsolete. *Library J.* **121**(October) 34–36.

Westland, J. C. 1992. Congestion and network externalities in the short run pricing of information systems services. *Management Sci.* **38**(7) 992–1009.

Accepted by Chris Kemerer; received February 4, 1999. This paper was with the authors 8 months for 2 revisions.