UCLA
Recent Work

Title

Service Design, Competition and Market Segmentation in Business Information Services

Permalink

https://escholarship.org/uc/item/92c8m634

Authors

Anant, B.
Karmarkar, U. S.

Publication Date

2005

eScholarship.org Powered by the California Digital Library
University of California

Service Design, Competition and Market Segmentation in Business Information Services with Data Updates

Bashyam Anant
Macrovision Corporation
4301 Norwalk Drive, San Jose, CA 95129
Email: bashyam_anant@yahoo.com

Uday S. Karmarkar
UCLA Anderson Graduate School of Management
110 Westwood Plaza, Los Angeles, CA 90095-1481.
Email: uday.karmarkar@anderson.ucla.edu

Abstract

Business information services are intermediaries that collect, collate, package and distribute information of value to professional users. We consider two technologies that such intermediaries may use for delivering information. First, a packaged design that uses physical media like CD-ROMs to distribute information. Second, an online service that delivers such information via the Internet or other online networks. We model a market where subscribers may choose between "self-service", where they collect and collate information directly from sources, and a third party service provider who provides either a packaged design or an online service. Subscribers are indexed by their volume of usage for the service. In a duopoly, we show that providers with online or package technologies will serve different market segments. The package provider's limited ability to provide current information, combined with decreasing search costs in an online service will make a package provider increasingly vulnerable to being driven out of the market by the online provider.

Keywords: Industrial marketing: pricing, segmentation, competitive strategy. Computers: databases, Internet. Non-cooperative games: Competition in information goods

1. INTRODUCTION

Many professionals rely on current information relevant to their work. Some firms/subscribers may obtain information directly from source—what we refer to as the "self-service" option in the rest of the paper. In other cases, a third party intermediary, referred to as a Business Information Service provider, may undertake to collect and collate the information, and deliver it to subscribers either in the form of a database "packaged" as a CD-ROM (or other physical media) or as an online service. Thus, a Business Information Service provider is an intermediary between information sources and firms or subscribers that have a use for that information.

Examples of such services include: Reuters and Bloomberg for foreign currency price information delivered to currency traders; the Lexis/Nexis online service for legal, regulatory and company information; Jeppesen (a Times Mirror Company) for aeronautical maps in CD-ROM and paper formats targeting pilots and flight controllers and Component Information Services, described below, to motivate our research questions and modeling assumptions (Bashyam and Karmarkar 2000).

Component Information Services provide engineering information on semiconductor components for electronic design engineers. These services are essentially catalogs of electronic components with detailed technical descriptions of parts, engineering drawings and usage tips. Design engineers use this information to choose between alternative components for products they are developing. Even today, many such component catalogs are printed books or loose-leaf publications in binders.

Design engineers in this market always have the "self service" option, which refers to firms (and design engineers therein) collecting components information in whatever format (paper, CD-ROMs, websites, online downloads) their vendors make them available in and patching together an index or even a rudimentary database to this information, if possible. While this approach is cheap, it is more than likely to miss out the latest parts (because the search process is neither complete nor focused on current information) and does not simplify a design engineer's task of comparing components across vendors. Users will likely experience a large search cost to get to the component they are looking for, may give up and make sub-optimal component choices as a result.

Two Component Information Services were launched in the nineties to provide a components catalog for design engineers to alleviate the problem they faced searching for and comparing components from different vendors—i.e., lowering their search costs for component selection. The two services, Aspect Development and EnGenius (a Motorola division), took very different approaches to designing their service offering, and have evolved differently.

EnGenius, an online service (*Electronic Engineering Times* 1993) launched in 1993 was the first online catalog of semiconductor components. Catalog information was stored in standardized form, in the SGML (Standard Generalized Markup Language) format. Component vendors, who wished to have their product information contained in the database, transmitted data in the SGML format to EnGenius over an online network which EnGenius' software normalized to a common parts description hierarchy to allow comparisons. A subscriber of the service could connect to the database through the Internet. They could then search for, compare, view, edit and print (but not store) information about components that met requirements. Besides raw product data sheets, EnGenius provided images and technical comparisons of components. EnGenius charged subscribers a fixed fee ($39/month) plus a marginal price, which could be a fee per download or a fee per online session, or an hourly fee.

EnGenius ceased operations in 1994. Its failure was attributed to a catalog database that had too few components (only 5000) to be of use to engineers and high online search costs (because accessing the service's rich mix of graphical, textual and parametric information was too slow over available networks).

Aspect Development launched its service around the same time as EnGenius. Aspect provides a "packaged" service that consists of a database on CD-ROMs bundled with a software search engine to query and compare components information in the database. The CD-ROM and search software are shipped to a subscriber firm every 3-4 weeks—this update cycle is necessary to enter enough new data (between updates), manufacture CD-ROMs and send them to customers. Aspect originally created its database by collecting raw data from semiconductor companies in paper, CD-ROM and online files and entering data manually into its database. It uses the same process for capturing on going updates. Aspect's approach costs it more per record to update and maintain its database- a cost that was largely borne by component vendors in EnGenius' approach. Nonetheless, Aspect's database today has a stable size of 5 million components from 931 vendors that collectively account for 85% of such components sold worldwide. Its customers include most of the 200 largest electronic manufacturers in the world who on an average pay a fixed annual subscription fee of about $500,000.

These case studies reveal several tradeoffs in the Business Information Services market that drive its evolution and help formulate the research agenda for this paper. First, online (e.g. EnGenius) and packaged (e.g. Aspect) services differ in their ability to deliver current information—specifically, longer update cycles for package versus online services. Second, search costs for users differ in the technologies. It is likely to be lower in the package versus online service because of online network delays in retrieving information. Third, pricing mechanisms for a package provider are less flexible than for an online provider. An online provider can charge a "pay per use" marginal charge, in addition to a fixed charge, as the EnGenius example demonstrates. A package provider is obviously more constrained in pricing because pay per use metering is not possible—in our model, we abstract this limitation into a simple fixed price charged by the package provider. Fourth, online providers like EnGenius can lower their database creation and update costs relative to package providers because they can have information sources submit information in a standard and normalized format and avoid or minimize costly data entry. Fifth, the database size is a critical determinant of the value of the service to subscribers as well as costs for building and maintaining a database (regardless of who bears the cost) as the failure of EnGenius demonstrates. Sixth, database size and pricing are two crucial service design parameters under the control of providers. Other services attributes such as the update frequency and minimum search costs are largely "frozen" by technology choice. Seventh, service providers have to ensure greater value for their service compared to the "self service" option, where subscribers collect and collate information directly from sources, in order for them to use a third party provider.

Given the above dynamics between online and package providers, our intent in this paper is to examine the following questions:

- How do technology choices impact service design and performance?
- How do service design characteristics, such as search costs, update frequency, pricing and database sizes affect the value of third party information services
- Given the above relationships, how would online versus package providers position their service if they compete?

- What technology (online versus packaged) should a service provider choose given the inherent tradeoffs in the technology? When would both approaches be used and how?
- Was a service like EnGenius doomed to fail given its high search costs and sparse database? If not, under what conditions could they have survived?

To answer these questions, we model the Business Information Services market as a duopoly where subscribers may choose between their "self service" option or one of the third party services—an online service or a package provider. Subscribers are indexed by their annual usage of such services. To capture a subscriber with a given usage volume, a service provider must exceed the value provided by that subscriber's self service option. The value provided by a service provider is a function of their service design, which is driven by their technology choice (online or packaged) and choice of service design parameters, specifically database size and pricing (fixed and marginal prices). A service provider's technology choice automatically "freezes" certain aspects of service design, such as (the minimum possible) update frequency and search costs, while aspects such as database sizes and pricing are under the control of the service provider. Our models for self service, online and package provider options relate a subscriber's value for each option to their usage volume, update cycles, search costs, database size and pricing. Subscribers pick whichever option (online, package or "self service") provides maximum value, allowing us to determine market segments for different options. In a duopoly, two third party providers, having chosen their technology (online or packaged), compete for subscribers who also have a "self service" option, by choosing their profit maximizing service design parameters, namely database sizes and fixed and marginal (for the online provider only) prices.

Our results are as follows. In a duopoly with an online and package provider, we show that providers will differentiate their service design by choosing different database sizes and prices, and serve different subscriber segments (i.e. different usage volume segments). Two cases are possible in a duopoly. First, when the update interval for the package provider is high enough and/or when the search costs in the online service are low enough, we show that the package provider will serve low usage volume subscribers, if they enter the market at all, while the online provider serves high usage volume subscribers. Second, when search costs are high enough in the online service and/or the package service's lower search costs compensates for its longer update interval, we show that the online provider will always enter (as long as they can be profitable), but serve low usage volume subscribers, while the package provider will server high usage volume subscribers. We also show that package providers are increasingly vulnerable to being driven out by online providers as online search costs improve. These results allow us to draw insightful conclusions about the questions we set out to answer.

2. LITERATURE REVIEW

As a representative work from the information sciences literature, Westland (1989) describes the determinants of market structure in online information services. Using a statistical model, he shows that high acquisition rates for data, coupled with low obsolescence rates, increase total marginal costs of maintaining an online database, resulting in a concentrated market with few firms. This observation, in part, supports the restricted market structure assumptions (i.e. duopoly) made in this paper. We also model how acquisition rates and obsolescence rates relate to database size.

The market segmentation effects described in this paper are most closely related to Moorthy (1988), which provides a duopoly model for understanding pricing and product design strategies. Potential consumers evaluate a product on its price and quality when choosing between the duopolists and a substitute producer. He shows that the duopolists differentiate their product, the higher quality being chosen by the higher margin firm. The marginal utility of a subscriber to an information service provider is similar in its economic effects to the product's quality attribute in Moorthy's work. Our paper however, explicitly relates the tradeoffs in technology (update cycles, search costs, pricing models, fixed and variable costs for providers) and choice of service design parameters (database sizes, pricing) to the value to customers for different information delivery options and to market segmentation and market structure.

The rest of the paper is organized as follows. Section 3 details the model. In Section 4, we analyze the self service option for a given subscriber with a certain usage volume which determines the threshold benefit level that a service provider must meet or exceed to capture that subscriber. In Section 5, we analyze a market where two service providers compete for subscribers who have a self service option by picking their database size and pricing, to describe market segmentation results. A final section concludes this work.

3. THE MODEL

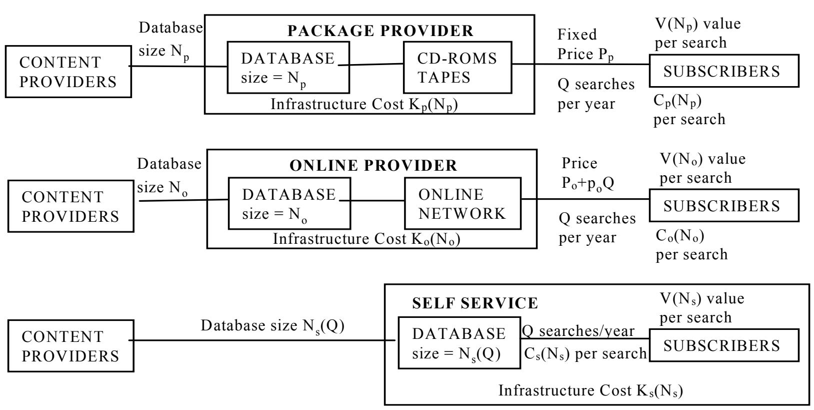

Figure 1. Packaged and online services compared with the self-service option.

The model structure and notation is depicted in Figure 1. Subscribers to business information services are indexed by their annual volume of usage of the service denoted by Q (searches/year). Subscribers may choose between an online or a package provider or may prefer "self service", where in they collect, collate and assemble a database of their own directly from source.

Subscribers pay an annual fixed charge P_p if they use a packaged provider. For the online service, subscribers pay both a fixed charge P_o and a marginal charge per search p_o, since usage can be monitored. We use the subscript "p" to denote a "packaged" and "o" to denote an online service in the rest of the paper. In reality, pricing for a package service is somewhat more complex than a simple fixed price described here. Aspect Development (Bashyam and Karmarkar 2000), for example, uses a tiered pricing arrangement based on the number of subscribers that simultaneously access their database (within their own environments, not online)—it is easy to see that Aspect is effectively using a menu of marginal prices from which subscribers can choose. We assume in our model that the package provider charges a simple fixed charge, primarily to reflect a package provider's limitation of not being able to use "pay

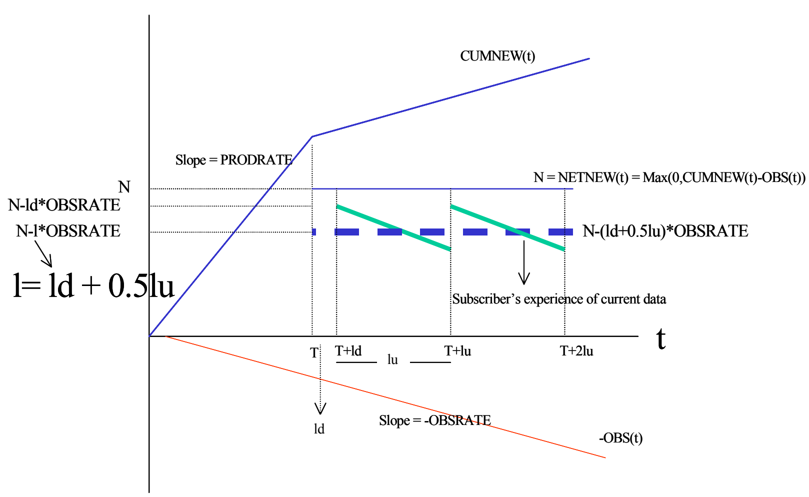

Figure 2: Growth pattern for new information in a service. Service providers grow their database to a size N, after which they "ship" it to subscribers at an update interval lu. Subscribers may experience an additional lag time ld, between when an update is ready and when thev receive the database.

per use” schemes available to online providers. Our models will still be applicable if we were to assume a marginal price for the package provider as well.

Service providers create a database and allow subscribers to access it, either in-house (for the package service) or online. The service provider’s process for building their database is shown in Figure 2. In the figure, the provider starts building a database at time t = 0. CUMNEW(t) represents the cumulative records in the service provider’s database at time t. The slope dCUMNEW(t)/dt is the service provider’s production rate for new records, denoted as “PRODRATE” in Figure 2 for the case where CUMNEW(t) is linear. OBS(t) represents obsolete records at time t. The slope dOBS(t))/dt is the rate at which records obsolesce in the database, denoted as “OBSRATE” in Figure 2 for the case where OBS(t) is linear. The difference between the two curves (CUMNEW(t)-OBS(t)) is equal to the net new records in the database at time t, denoted by NETNEW(t). At t = T, the service provider decides to alter their production rate to match the obsolescence rate. Consequently, the net new records in the database stabilizes at NETNEW(t) = N for all t ≥ T. N is referred to as database size in the rest of the paper.

This pattern of database growth by service providers, where they reach a “steady state” database size over time, is supported by current practice in the business information services industry. Aspect Development, for example, began by rapidly adding components to its database in the initial years. When a large percentage of new components in the electronics industry had already been captured in the database, Aspect was able to reduce its production rate to more or less the rate at which components obsolesced. Aspect’s “steady state” database size today stands at over 5 million parts from 921 manufacturers that collectively account for over 85% of the worldwide electronic components produced.

We assume that once service providers reach their steady state database size N, they allow subscribers to begin using the service. As in Figure 2, the provider "ships" a database of size N at an update interval lu, after entering new records since the last update, cleansed of obsolete records. Subscribers also experience a delay ld between when a service provider finishes an update and when the subscriber receives it—this is the time it takes, say for a package provider to manufacture CD-ROMs and ship them to customers. Effectively, the subscriber "experiences" a database with fewer new records because the data in the database becomes obsolete between updates and during the delay between when an update is completed and when the subscriber receives it. The subscriber's view of the net new records in the database at time t is represented by the serrated line in Figure 2. In the case where CUMNEW(t) and OBS(t) are both linear as in Figure 2, it is easy to show that the average database size "experienced" by a subscriber in steady state is N-l*OBSRATE, where OBSRATE = dOBS(t)/dt and l = ld + lu/2, denotes the effective update interval. In other words, the subscriber experiences a database that is smaller than the provider's steady state database size N by as many records as become obsolete during the provider's effective update interval. We assume for simplicity that the effective update interval (l = ld + lu/2) is not a service design parameter for providers, mainly because it is largely governed by the choice of technology (packaged versus online). For example, a package provider may find that shipping daily updates may be cumbersome for subscribers because of change over and version control problems, even if distribution costs were not an issue for the provider.

Service providers incur two types of costs for building and maintaining their service. First, they incur fixed costs relating to R&D costs for creating a normalized data hierarchy, software for cleansing and normalizing raw data, search software for the database and user interface software. Fixed costs also include cost for the hardware used for performing data entry and connecting the provider to an online service (in the online case). Second, they incur variable costs that depend on the size of the database created. These pertain to the actual cost of data acquisition (e.g. copyrighted information), reconciling format differences, data entry and indexing (Bashyam and Karmarkar 2000). Boeri and Hensel (1995) estimate a cost of $4.70 per record for data entry, formatting and indexing in the SGML (Standard Generalized Markup Language) format for the legal publishing industry. Up to the time the database reaches steady state ($t \leq T$) the service provider incurs these variable costs at the rate of dCUMNEW(t)/dt (see Figure 2). After that, these costs are incurred at the rate of dOBS(t)/dt. If CUMNEW(t) and OBS(t) are linear, then the total variable cost incurred to build and maintain a database can be expressed as kN + k*OBSRATE*t ($t \geq T$), where k is the variable cost per record. On an annualized basis, the fixed and variable costs can be abstracted into a cost function $K_i(N_i)$, where i indicates the underlying service delivery technology, the subscripts s, p, and o denoting "self service", packaged, and online information service options respectively. When CUMNEW(t) and OBS(t) are linear, $K_i(N_i)$ = 1 year * ((Fixed costs+ k_iN_i)/useful life of service + k_i*OBSRATE). Let $K_i(0)$ = Fixed costs/useful life of service + k_i*OBSRATE. Thus, $K_i(N_i)$ = $K_i(0)$ + k_iN_i/(useful life of service). As formulated, the steady state database size, N, is a crucial variable in the design of the service.

We assume that $K_o(N) < K_p(N)$ because an online provider can drive down their marginal production costs per record (ie $k_o < k_p$) by having content sources submit updates in a standard format. The online provider lowers their cost by transferring the cost of cleansing, normalizing and entering data to the source, practiced, for example, by EnGenius. The package provider takes on these costs which is why we assume $K_o(N) < K_p(N)$. Of course, this also implies that the fixed costs are more or less comparable, which is a reasonable assumption in practice because most of the fixed costs are common to both types of providers.

For the package service, there may be additional costs for duplicating the database on CD-ROMs for each subscriber. However, these costs do not depend on the database size and are negligible compared to the costs for data entry. Specifically, manufacturing costs for CD-ROMs are independent of the amount of data they contain. Unit manufacturing costs for standard CD-ROMs are about $1 in lots of 1000, with mastering charges adding $700 to the total, bringing the manufacturing cost per CD to less than $2 (Wiedemer and Boelio 1995).

An online provider may incur costs for data transmission to subscribers if they are on a private wide area network. For ease of exposition, we assume that these costs are subtracted from the marginal price per search (p_O) that the online provider charges. An online provider's fixed costs may also be impacted by their market share because database queuing delays in a successful service may result in additional infrastructure costs to support an acceptable level of service for a large subscriber base. For simplicity, we assume that these costs are factored into the fixed costs by determining the costs associated with the upper bound on market share for the online provider (for example, by determining the fixed costs for the provider to support their

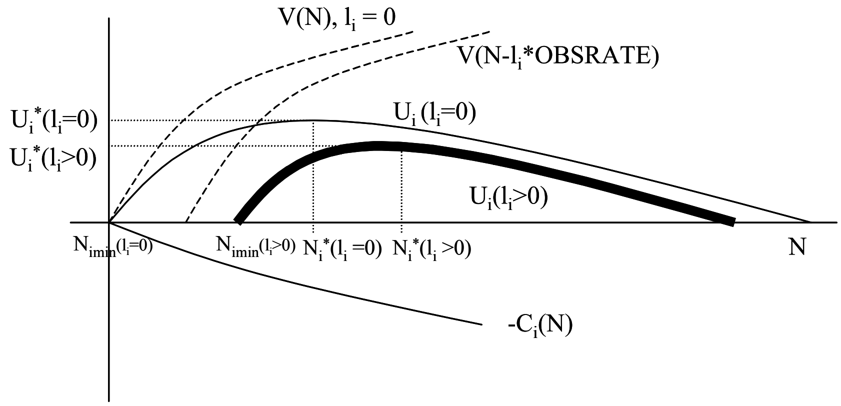

Figure 3: Net benefit U_i for a subscriber in technology i as a function of database size, and effective update interval.

monopoly market share). Note that database queuing delays also impact the package provider, except that the resulting fixed cost increases would be borne by the subscriber.

Subscribers incur a search cost for each record accessed. The search costs depend on several characteristics of the database including its size (in number of records), depth of information capture per record (e.g. how many fields describe a record), access technology (whether packaged or online), quality of indexing, degree of specialization of the service (e.g. a database customized for particular segment of users) and queuing delays (particularly, in an online service). Online data transfer rates can be 100 times slower than the transfer rate for a CD-ROM. Even today, many small and mid-sized businesses do not have high speed online access. Queuing delays in an online service further add to the search costs for online subscribers. Queuing delays in the online case depend on the number of subscribers connecting to the service simultaneously, which is driven in part by the service's market share. For simplicity, we assume that the online provider spends additional fixed costs to deliver an acceptable response time (i.e. acceptable search costs) based on an upper bound of their market share. We assume that the essence of the search cost for a query may be captured in a cost function $C_i(N_i)$, $C_i(0)=0$, which accounts only for its technology and database size dimensions. Factors such as the quality of indexing, degree of specialization and depth of information capture per record are not modeled into the search costs because they apply equally to online and packaged services, while our interest in this paper is study the impact of their differences. Naturally, we expect the search cost to increase with database size for a given technology. Because of data transfer and queuing delay limitations, it is reasonable to assume that $C_p(N) < C_o(N)$ for a database size N. In addition, we should expect (though we do not assume) that $C_o(N) < C_s(N)$ because typically a self-service database is often less functional than a third party service. For example, in the components information service example, the self-service database used by many firms is a collection of free paper databooks and CD-ROMs provided by component manufacturers (Bashyam and Karmarkar 2000). Such databases have little or no search capabilities, as a result, self-service searches are likely to be more costly than an online (third party) service for a given database size N. In rare cases, it is conceivable that a subscriber firm may actually be able to build a “self-service” database that is as sophisticated as a third party service—in such cases, it is conceivable that such firms may enter the service provider market themselves. This was the case with Motorola with its EnGenius Components Information Service—as a large user of components information, they were able to build their own database and provide third party access to it as a service. In summary, we assume that $C_p(N) < C_o(N)$, but make no assumptions about where $C_s(N)$ falls in this ordering.

Subscribers derive an intrinsic value of information, regardless of how it is delivered, which is abstracted into V(N), V(0) = 0, common to all subscribers. Because of the update interval in delivering the service to a subscriber, they effectively experience a database of size N-l*OBSRATE, where l is the effective update interval and OBSRATE is the rate of obsolescence of information, as discussed in Figure 2. Thus, a subscriber's net value for each search in an information service is $U_i(N) = V(N\text{-}l_i OBSRATE) - C_i(N)$; let $U_i(0) = 0$, $U_i(N) \geq 0$ for $N \geq N_{imin}$, where the subscripts i (s, o, or p) relate value to the underlying service technology. We assume that for $N \geq N_{imin}$, $U_i(N)$ is a concave function of N, with a unique maximum at N_i^* (see Figure 3). This assumption captures the notion that, for a given search, more information is valuable, but after a certain number of alternatives, the search costs become significant and lower the net benefit from the search. To summarize, we assume:

Assumption 1. V(N) *and* C_i(N) *are both monotonically increasing functions of the database size, for an information delivery technology* i. *In addition, the net benefit function*, U_i (N) = V(N)-C_i (N) *is concave for* $N_i > N_{imin}$ *with a unique maximum* N_i^* *as in Figure 3. If* V(N) *is more "concave" than* C(N), *this assumption would hold.*

N_i^* is the maximum database size that a provider would build—anything more than N_i^* will not add value to subscribers given their search costs. Note that N_i^* depends on the effective update interval l_i (i = s,p,o) for the service provider as a shown in Figure 3. The effect of the update interval is to increase the maximum database size ($N_i^*(l_i>0) > N_i^*(l_i=0)$) and decrease the corresponding maximum net benefit ($U_i^*(l_i>0) > U_i^*(l_i>0)$) for subscribers. It is reasonable to expect that $l_s < l_o < l_p$, because of the longer update cycles in a packaged service compared to online or self service options. Clearly, the package provider will not be viable in certain industries (e.g. in industries with OBSRATE > $N_p{}^*/l_p$), which is why package providers are absent in markets such as financial information services.

In sum, this section provided us with the essential ingredients of the cost structure of information service providers and cost and benefits for their subscribers. In the next section, we explore the subscriber's self service option. By characterizing a subscriber's self-service benefit, we establish the lower bound on the benefit that a third party provider (online or package) must exceed to capture that subscriber. In Section 5, we first derive the subscriber's benefit function for online and packaged options. Then, we analyze a duopoly with two providers that compete for subscribers indexed by their usage volume Q and have the option of self service. Though the market structure assumptions appear restrictive, they are actually quite representative of many information services markets. For example, Aspect is a monopolist in the Components Information Services (though it may seem that component distributors with their websites may compete with Aspect, in reality such websites are of little use to design engineers that Aspect targets. Distributor web sites are oriented towards procurement professionals who need to check price and availability of specific components). Reuters, Bloomberg and Jeppesen monopolize their markets. We believe that such concentration arises because of the specialized nature of content, high liability costs associated with incorrect information, high data entry and update costs, service branding and long lead times to create a valuable service—factors that lead to high entry barriers.

4. THE SELF SERVICE OPTION

Potential subscribers to information services are characterized by the number of searches they undertake per year, denoted by Q; $0 \le Q \le Q_{max}$. Given their level of usage, a firm opting for self service must decide its preferred database size N_S, by trading the benefits of a large database against the costs of building and maintaining it. For simplicity, we assume the lag time for self service $l_S = 0$. The net benefit per search is given by $U_S(N) = V(N) - C_S(N)$, for a database of size N.

The problem solved by a subscriber firm with annual usage of Q searches is:

$$N_S(Q) = \arg\max_{N_S \ge 0} \, (V(N_S) - C_S(N))Q - K_S(N_S)$$

$$W_S^*(Q) = \max_{N_S \ge 0} \, (V(N_S) - C_S(N_S))Q - K_S(N_S)$$

where

N_S	is the size of the self service database; we will treat N as continuous for convenience
$V(N)$	is the value to the firm from searching a database of size N; $V(0)=0$
$C_S(N)$	is the firm's cost of searching in its in-house database of size N; $C_S(0)=0$
$U_S(N)$	is the firm's net marginal benefit
$K_S(N)$	is the firm's annual cost of creating and maintaining a database of size N
$W_S^*(Q)$	is the net benefit from self service for a subscriber with annual usage Q

The self service benefit function $W_S^*(Q)$ is characterized below. Proofs are given in Appendix A.

Proposition 1. (a) $W_S^*(Q)$ *is convex, non-decreasing in* Q, $W_S^*(Q) \geq 0$. *See Figure 4.*
(b) $W_S^*(Q) \geq U_S^* Q - K_S(N_S^*)$, where $N_S^* = \arg\sup_{N \geq 0} U_S(N)$, $U_S^* = \sup_{N \geq 0} U_S(N)$.
(c) *The linear bound of (b) is approached asymptotically as* Q *becomes large*. *If* U_S^* *is positive then* $W_S^*(Q)$ *is eventually increasing*.

$W_S^*(Q)$ represents the minimum benefit that a third party information service provider must ensure in order for a potential subscriber firm with annual usage Q to join the third party service. $U_S^*$is the maximum possible benefit under self service (see Figure 3), and is driven by the self service search costs traded against the value of a self service database.

5. MODELS OF COMPETITION

Here, we develop duopoly models where two information service providers compete with the self-service option. Service providers choose their profit maximizing service design parameters, their database size N_i, and price parameters (fixed price for package, fixed and marginal price for online). As a first step, we characterize the benefit functions for subscribers under online and packaged technology below.

5. 1. Subscriber Benefit Functions for Package and Online Services

5.1.1. Package Provider's Benefit Function

In the package service, the number of searches made by a subscriber cannot be monitored, and the subscriber can only be charged a fixed price $P_p \geq 0$ for the service. The value from the package service for a subscriber making Q searches/year can be written as:

$$(V(N_p - OBSRATE^* \, l_p) - C_p(N_p))Q - P_p = U_p(N_p - OBSRATE^* \, l_p)Q - P_p$$

where N_p, OBSRATE, l_p are the packaged database size, obsolescence rate and effective update interval for the package provider. Notice that the benefit function is linear.

5.1.2. Online Provider's Benefit Function

In the case of an online service, a subscriber's search activity can be monitored, and a charge p_O made for each search. The value from the online service for a subscriber making Q searches/year can be written as:

$$(V(N_O - OBSRATE^* l_o) - C_O(N_O) - p_O)Q - P_p = (U_O(N_O - OBSRATE^* l_o) - p_O)Q - P_O$$

where $U_O(N_O\text{-}OBSRATE^*l_o) = V(N_O\text{-}OBSRATE^*l_o) - C_O(N_O)$ and N_O, OBSRATE, l_O are the online database size, obsolescence rate and effective update interval for the online provider.

Figure 4 displays these benefit functions along with the self-service benefit function $W_S^*(Q)$. In the figure, it is easy to visualize market segments for online and package providers by having subscribers choose whichever option (self service, online or package) maximizes their benefit. Importantly, Figure 4 reveals that the market segment for package and online providers is a contiguous interval of the form (Q_1, Q_2), since the subscriber's benefit function is convex for self service and linear with an information service provider. These market segments can be described by:

$$M_p = \{Q|\max(W_S^*(Q), (U_O(N_O\text{-}OBSRATE^*l_o)\text{-}p_O)Q - P_O) < U_p(N_p\text{-}OBSRATE^*l_p)Q\text{-}P_p\}$$

$$M_O = \{Q|\max(W_S^*(Q), U_p(N_p\text{-}OBSRATE^*l_p)Q\text{-}P_p) < (U_O(N_O\text{-}OBSRATE^*l_o)\text{-}p_O)Q - P_O)\}$$

where M_p and M_O are to the market segments for the package and online providers, respectively.

5.2. Models of Duopoly

In this section, we consider a duopoly where potential subscribers choose between self-service and one of the two third party providers, whichever maximizes their net annual benefit. An illustrative segmentation is shown in Figure 4. Two cases can occur. In one case, the segments are adjacent; in the other, the segments are separated by a group of subscribers that select self service. For a given set of price and size decisions, these cases can be detected by finding the breakeven point between the two information service providers, and checking to see whether the value of self service at this point is higher or lower than that for the information services. This observation is summarized in Proposition 2, and is used to determine the market shares for the duopolists.

Proposition 2. *Define* $Q_{po} = (P_p\text{-}P_o)/(U_p\text{-}U_o\text{+}p_o)$. *If* $W_s^*(Q) \leq U_p\, Q_{po} - P_p$, *then the market segments of the duopolist are adjacent, otherwise a group of self service subscribers separate the two segments. Visualize this by manipulating the benefit lines in Figure 4.*

When Proposition 2 holds, the market share for the packaged provider is given by $MS_p(N_p,P_p,N_o,P_o,p_o) = G(Q_{po})\text{-}G(Q_{p2})$ if $U_p > U_o\text{-}p_o$ and by $G(Q_{p1})\text{-}G(Q_{po})$ if $U_p < Uo\text{-}p_o$, where Q_{pi}, are as in Figure 4, and G(Q) is the proportion of subscribers whose annual usage volume exceeds Q.

Consider the case where one firm provides a packaged database, and the other an online one. Define **A** and **B** as the strategy spaces for the packaged and online information service providers respectively. We characterize them below:

$\mathbf{A} = \{(N_p, P_p) |\, N_{pmin} \leq N_p \leq N_p^*, 0 \leq P_p \leq P_p^{max}\} \subseteq \mathbf{R}_+^{\mathbf{2}}$

$\mathbf{B} = \{(N_o, P_o, p_o) |\, N_{omin} \leq N_o \leq N_o^*, 0 \leq P_o \leq P_o^{max}, 0 \leq p_o \leq U_o^*\} \subseteq \mathbf{R}_+^{\mathbf{3}}$

where for i = s,p,o

$U_i^* = \max_{N \geq 0} U_i(N\text{- OBSRATE}^*l_i)$ is the maximum benefit per search for provider i

$N_i^* = \arg\sup_{N \geq 0} U_i(N\text{- OBSRATE}^*l_i)$, is the benefit maximizing database size for provider i

Let **a** and **b** denote strategies of the two players, with $PR_i(\mathbf{a},\mathbf{b})$ denoting the profits for firm i, and should be interpreted as the profit that the package provider makes given strategy vector b for the online provider and vice-versa. The upper bound P_i^{max} on fixed price is clearly provider i's fixed price as a monopolist. The upper bound on database size N_i^* arises because U is concave (Figure 3), and exceeding N^*, costs more without capturing a larger market, as can be seen by manipulating the slope of either provider's benefit line in Figure 4.

Firms play a one shot game in which they choose the service design parameters in their control (for package provider (N_p, P_p), for online service (N_o,P_o,p_o) simultaneously. We assume that the technology choice decision (online or package) is made before the game is played.

We define a Nash equilibrium for that game as follows (Friedman 1977):

Definition 1. *A Nash equilibrium is a combination* $(\mathbf{a}^*,\mathbf{b}^*) \in \mathbf{A}\mathrm{x}\mathbf{B}$ *that satisfies* $PR_a(\mathbf{a}^*,\mathbf{b}^*) \geq PR_a(\mathbf{a},\mathbf{b}^*)$ *for* $\mathbf{a} \in \mathbf{A}$ *and* $PR_b(\mathbf{a}^*,\mathbf{b}^*) \geq PR_b(\mathbf{a}^*,\mathbf{b})$ *for* $\mathbf{b} \in \mathbf{B}$.

A Nash equilibrium is essentially a pair of strategy vectors (database size and prices) for the package and online providers such that the package provider, taking the online provider's equilibrium strategy for granted, has no incentive to change their equilibrium strategy and vice versa for the online provider. Equivalently, each player's strategy is really their combination of slope (U_i, which is governed by database size and marginal charge) and intercept (P_i, fixed price) of their benefit functions as in Figure 4.

A necessary and sufficient condition for an equilibrium to exist is for the profit function $PR_i(N_p,P_p,N_o,P_o,p_o)$ i = p,o to be quasi-concave (Friedman 1977) and for the strategy spaces (**A** and **B**) for the two players to be compact. In general, it is not possible to derive analytical expressions for the profit functions formulated earlier for the package and online providers, primarily because we cannot derive analytical expressions for the breakeven points for the providers (see Figure 4). As a result, concavity of the profit functions cannot be established. However, we are able to derive conditions that characterize the equilibrium in a duopoly, which allows us to drive interesting conclusions about market structure and segmentation effects. This is discussed next.

5.2.1. Duopoly with Homogenous Service Providers

Suppose that there are two providers of packaged (or online) databases, with similar technologies and capabilities, so that they have the same cost structure for creating and maintaining databases, and provide customers with the same search capabilities, implying the same search cost functions. Index the two firms as 1 and 2. We assume that customers will pick the service that delivers the maximum benefit to them. Then the following holds:

Proposition 3. *Consider a market with two identical service providers, indexed 1 and 2. If both providers survive in equilibrium, they will employ different database sizes and prices, such that* $P_1 > P_2$ *and* $U(N_1)-p_1 > U(N_2)-p_2$, *make identical equilibrium profits and have disjoint market segments-- an example of disjoint market segments is depicted in Figure 4. Note that if there are two online providers, the result holds with* $p_{1,2} \equiv 0$

It is useful to visualize the disjoint market segments result for the case when F(Q) is a Binomial distribution. Suppose the usage volume is either q1 or q2 (q2 > q1), with x% of the subscriber base having usage volume q1 and 1-x% with volume q2. By definition, the market segments will be disjoint. It is easy to see that if both providers survive, one provider will opt for subscribers with volume q1 and the other will serve subscribers with volume q2. In addition, both have to make equal profits, otherwise, the lower profit firm has an incentive to mimic and undercut the higher profit firm.

Corollary 3.1. *Consider a market with more than two identical service providers. If all providers survive in equilibrium, they will employ different database sizes and prices, such that* $P_i > P_j$ *and* $N_i > N_j$, for any i and j, $i \neq j$, *make identical equilibrium profits and have disjoint market segments.*

5.2.2. Duopoly with Heterogeneous Service Providers

Consider a market with one package and one online provider.

Let $M_p = \{Q|\ U_p Q - P_p > \max(W_s^*(Q), (U_o - p_o)Q - P_o)\}$ denote the market segment for the package provider. Then, the package service provider's problem may be stated as:

$$\text{Max } PR_p(N_p, P_p, N_o, P_o, p_o) = P_p\, m\, MS_p(N_p, P_p, N_o, P_o, p_o) - K_p(N_p)$$
$$a \in \{P_p, N_p\} \subseteq \mathbf{A}$$

where

m	is the total number of potential subscribers
P_p	is fixed charge (annualized) for purchasing the packaged information service
MS_p	$= G(\min\{M_p\}) - G(\max\{M_p\})$ is the provider's market share
PR_p	is the net profit for a given database size and fixed price
F(Q)	is the proportion of subscribers with less than Q searches per year
G(Q)	= 1-F(Q), is the proportion of subscribers with more than Q searches per year
A	$= \{(N_p, P_p)|\ N_{pmin} \le N_p \le N_p^*,\ 0 \le P_p \le P_p^{max}\} \subseteq \mathbf{R}_+^{\mathbf{2}}$
U_p^*	$= \max_{N \ge 0} U_i(N - OBSRATE*lp)$;
N_p^*	$= \arg\sup_{N \ge 0} U_i(N - OBSRATE*lp)$

The online provider's problem can be stated as:

$$\text{Max } PR_o(N_p, P_p, N_o, P_o, p_o) = P_o m MS_o(N_p, P_p, N_o, P_o, p_o) + m p_o S(N_o, P_o, p_o) - K_o(N_o)$$
$$b \in \{P_o, p_o, N_o\} \subseteq \mathbf{B}$$

where

$MS_o(N_p, P_p, N_o, P_o, p_o) = G(\min\{M_o\}) - G(\max\{M_o\})$ is the provider' market share

$M_o = \{Q|\ (U_o - p_o)Q - P_o > \max(W_s^*(Q), U_p Q - P_p)\}$

$$S(N_o, P_o, p_o) = E(Q_{o1}) - E(Q_{o2}), \text{ where } E(Q) = \int_Q^{Q_{max}} Q f(Q) dQ$$

$\mathbf{B} = \{(N_o, P_o, p_o)|\ N_{omin} \le N_o \le N_o^*,\ 0 \le P_o \le P_o^{max},\ 0 \le p_o \le U_o^*\}$

$U_o^* = \max_{N \ge 0} U_o(N - OBSRATE*l_o)$

$N_o^* = \arg\sup_{N \ge 0} U_o(N - OBSRATE*l_o)$

Two possibilities can arise in this market. First, when $U_o^* > U_p^*$, which may happen if higher search costs for the online service are overcome by that providers shorter update cycles (because $l_o << l_p$) or if search costs in the online service drop below that of the packaged service. Second, when $U_o^* \le U_p^*$, which may happen if search costs for the package service are so low that they overcome that provider's longer update intervals (because $l_p >> l_o$).

In the first case, we can show that if the package provider enters the market at all, they will have to serve low end (low Q) subscribers and make smaller profits in equilibrium than the online provider. In the second case, the situation is reversed with the package provider serving high volume subscribers while the online provider serves low volume subscribers. These are summarized below:

Proposition 4. *Consider a market with one online and one package provider. When* $U_o^* > U_p^*$, *if the package provider enters the market, they will serve low volume subscribers and make less profits than the online provider. The online provider will make higher profits than the package provider and serve high volume subscribers with a database size and marginal price such that* $U_o > U_p^*$. *The online provider's profits by imitating the package provider's low end strategy (slope of the benefit line and fixed price), has to be lower than what the online provider makes by serving high volume subscribers. If not, the online provider will monopolize the market.*

Proposition 5. *Consider a market with one online and one package provider. When* $U_o^* \leq U_p^*$, *the online provider enters the market as long as they make a profit, but will serve low volume subscribers. If the package provider enters the market, they will serve high volume subscribers with a database size such that* $U_p > U_o^*$. *The online provider will pick a low end strategy (slope of the benefit line and fixed price), which if imitated by the package provider would result in lower profits for the package provider compared to their high end strategy.*

A key insight from Proposition 4 and 5 is that the online provider always enters the market as long they can make a profit and will try to drive out the package provider if possible. Using a numerical example to simulate the above models, we are able to show that a package provider would be driven from the market as search costs in the online service decrease. The simulation starts with $U_o^* \leq U_p^*$, then U_o^* is increased by reducing search costs. We are able to show that profits for both providers decrease because of price competition. Eventually, when the search costs are equal, the package provider stays out of the market, and the online provider becomes a monopolist. As more subscriber firms connect to high bandwidth online access, package providers will struggle to survive given their larger costs for database creation and updates. Providing online access may be the only strategic option for the package providers in the future. By going online, such a firm would capture high end users through the packaged design and low end users through an online service.

5.2.3. The Two Stage Duopoly Game

What technology should a service provider choose, based on their knowledge of competition and profits in the market? Consider a market where there are two identical providers, indexed 1 and 2. They play a two stage game as follows. In Period 1, each provider picks a technology (either online or package). In Period 2, knowing each other's technology choice they pick their service design parameters (database sizes and prices) in a one shot game as described in Sections 5.2.1-2. The Period 2 game is a sub-game of the overall game—its equilibrium has been characterized in Propositions 3-5. The payoff matrix for the Period 1 game is shown in Table 1:

	PLAYER 1 CHOICE IN PERIOD 1 (PROFIT= first term in each pair)		
PLAYER 2 CHOICE IN PERIOD 1		PACKAGE	ONLINE
	PACKAGE	(-PR,-PR) or (PR_{p0}, PR_{p0})	(PR_{o1}, PR_{p1})
	ONLINE	(PR_{p1}, PR_{o1})	(-PR,-PR) or (PR_{o0}, PR_{o0})

Table 1: Payoff matrix for Period 1 game

The profit pairs in Table 1 correspond to what each firm would make in equilibrium in the Period 2 sub-game. When both firms pick identical technologies, we know from Proposition 3 that either they will drive each other out and make negative profits (denoted by –PR in Table 1) or if they both survive, they will make equal profits (denoted by PR_{p0} when both firms pick a package design and survive and by PR_{o0} when both pick an online service and survive). When one picks online and the other package, the equilibrium profits for the package firm is denoted as PR_{p1} and by PR_{o1} for the online provider—these values will have to calculated using the heterogeneous duopoly results of Propositions 4 and 5.

If there is no equilibrium in the Period 2 sub-game where both providers survive with identical technologies (i.e. according to Proposition 3, there is no choice of strategies where profits are equal and segments disjoint), we can show that the overall equilibrium in Period 1 will result in one firm picking the package design and the other will pick the online service. If there is an equilibrium in the Period 2 sub-game where both firms survive (i.e. according to Proposition 3, there exists a choice of strategies where profits are equal and segments disjoint), then three possibilities can arise in the overall game in Period 1. First, both providers choose a package design in Period 1—this will happen if $PR_{p0} > PR_{o1}$ in Table 1. Second, both providers choose an online service in Period 1—this happens if the Period 2 sub-game results in $PR_{p0} < PR_{o1}$ and $PR_{p1} > PR_{o0}$. Third, one provider chooses an online service and the other picks package design in Period 1—this happens if the Period 2 sub-game results in $PR_{p0} < PR_{o1}$ but $PR_{p1} < PR_{o0}$.

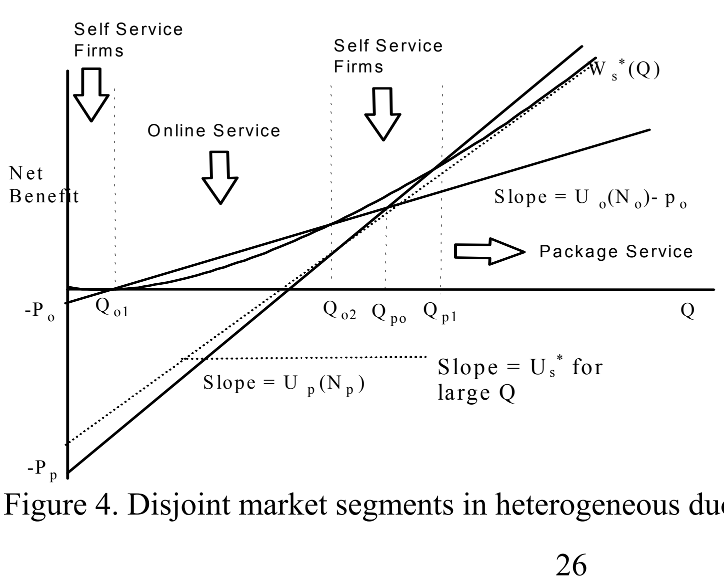

Figure 4. Disjoint market segments in heterogeneous duopoly, see Proposition 4.

The preceding discussion reveals that in the general case all possible patterns of market evolution can happen depending on the distribution of usage volumes (F(Q)) and service design trade offs (pricing, database sizes, update intervals, search costs) inherent in the technologies.

6. NUMERICAL RESULTS

In this section, we provide numerical results for the equilibrium model where we assume that firms play a single period game and choose their service design parameters (database sizes and prices) simultaneously. We use the Component Information Services market to exemplify cost and benefit functions described earlier.

6.1. Cost and Benefit Functions

Assume that the query volume of the subscriber population has a triangular distribution on $[0,Q_{max}]$ with a given mode m_f. The value function is taken to be exponential. Database building and update costs are assumed to be linear in the database size. The search cost function is also assumed to be linear. Thus, $C_i(N) = c_iN$, $K_i(N) = K_i(0) + k_iN$, $V(N) = d(1\text{-}\exp(\text{-}bN))$, $i = s,p,o$, denoting self service, packaged and online options. With these assumptions, we find the following:

Proposition 6. *For the above cost and benefit functions and a triangular* F(Q), *the subscriber's self service net benefit is given by*

$$W_s^*(Q) = \begin{cases} 0 & Q \le k_s/(db - c_s) \\ QU_s(N_s(Q)) - K_s(N_s(Q)) & Q > k_s/(db - c_s) \end{cases}$$

where

$$N_s(Q) = -\ln\big((c + (k/Q))/db\big)/b, \quad Q > k_s/(db - c_s)$$

The upper bound on the fixed price for a service provider is obtained by solving the monopolist's problem for that firm. For service provider i, i = p,o this turns out be as below:

Proposition 7. *For any* N_i, *the maximum price that a monopolist provider would charge is* $U_i(N_i)\max\{Q_{max}/3, \sqrt{m_f Q_{max}/3}\}$, *where* m_f *is the mode of the triangular distribution. The upper bound for the fixed price is given by* $P_i^{max} = U_i^* \max\{Q_{max}/3, \sqrt{m_f Q_{max}/3}\}$.

The lower bound on all price parameters is, of course, zero. Even with these simplifying assumptions, it is not possible to establish the concavity of $PR_i(N_p,P_p,N_o,P_o,p_o)$, i = o,p, necessary and sufficient conditions for the existence of an equilibrium (Friedman (1977)). Thus, we solve the duopoly game numerically.

6.2. Component Information Services Example

Consider the following example from the electronic component information services industry. A typical electronic design for a product (say a notebook computer) priced in the range \$1000-\$2000 might have approximately 30 discrete components. Suppose potential subscribers to the information service look for exactly one particular component for their designs, say an 8-bit comparator or a floppy disk controller. N (database size) is interpreted as the number of alternatives that are contained in the database for that component. Judicious component selection for a product priced in the given range could reduce the cost of the design or increase its selling price in the market by, perhaps, \$150, or roughly \$5 (=150/30) per component in the design. For example, by reducing the power consumption of a notebook computer, a firm could increase its battery life, which could translate into a higher selling price for the product. Thus, for each search, subscribers stand to gain a maximum of \$5, if they searched exhaustively for the ideal component. We assume that the value function for search is given as V(N) = 5(1-exp(-N/5)), that is, d = \$5, and b = 0.2. This implies that the first 5 alternatives generated by a search deliver 63% of the maximum value, which is realistic.

The maximum usage volume, Q_{max}, is related to such factors as the size of the subscriber firm, the number of design engineers, the number of concurrent product development teams, and its rate of introduction of new products. The largest of the subscriber firms is assumed to contain 100 concurrent design teams, each developing one product per year, each of which requires one search for the particular component. Thus, we set $Q_{max} = 100$. As before, we assume that $K_p(N) > K_o(N)$, for a given database size N, since the packaged service incurs higher marginal costs for digitizing data from data books, whereas the online service has semiconductor manufacturers perform this activity. The remaining cost parameters are displayed in Table 2, for which we solve each of the market models described previously.

6.3. Results

The pricing, profit and market penetration under different service designs and market structures are shown in Table 3. We see that the lower search cost package service is able to command a higher fixed price in a duopoly, and serves the high end of the market. Conversely, the online provider covers the lower end subscribers. The packaged service's choice of database size is very close to $N_p^* = 24.46$ because of its high fixed costs of building the database. For that reason, its database size is same when the service is a monopolist. As the packaged design reduces only its fixed price when moving from a monopoly to a duopoly, its profits certainly decline, but not to the same extent as the online service provider. The online provider's profits drop from \$54,179 to \$6,421 when the market structure changes from monopoly to duopoly, because of a corresponding increase in database size and a decrease in marginal price.

Parameter	Value
Q_{max}	100
Q_{min}	0
Mode	10
m	700
b	0.2
d	5
K(0): $(K_s(0),K_p(0),K_o(0))$	(0,30000,3000)
k: (c_s,c_p,c_o)	(0,3,1)
a: (a_s,a_p,a_o)	(0.9,x,y)
OBSRATE	0

Table 2. Test Matrix for Market Model

Market coverage in the duopoly is 95%, whereas it is 100% when the online provider is a monopolist and when the packaged service provider offers both technologies. The online provider is unable to serve the lower 5% in the former case because the average usage level in the lower 48% (= 100%-52%) of the market is too low to support a zero fixed price for online usage. When the online provider is a monopolist, the high end 52% of the market effectively subsidize the low end users, making a zero fixed price viable. Likewise, when the package provider offers both technologies, the upper 27% of the market subsidizes the remaining 73%, as seen by the \$60 (=\$218-\$158, see Table 2) premium they pay compared to the case when the provider has only the packaged technology.

Low end subscribers covered by the online provider derive greater surplus (measured by the gap between $W_s^*(Q)$ and $U_o(N_o)Q - P_o$)) when the online service competes with the

Market	Duopoly		Monopoly		Monopolist with Both Service Designs	
Parameter	Packaged	Online	Packaged	Online	Packaged	Online
N	24.28	8.05	24.28	5.65		10.57
P	122.34	8.61	158.66	0.00	218.88	0.00
p_o	-	1.19	-	2.23	-	2.26
PR	14,267.91	6,421.57	24,695.70	54,179.78		40,369.26
MS	0.52	0.43	0.49	1.00	0.27	0.73

Table 3. Database size, pricing, market share and profits for information service providers for Table 1 with $c_p = 0.0075$, $c_o = 0.2$.

packaged provider. In fact, in the duopoly, $N_o = N_o^* = 8.05$. Choosing a lower size would limit the fixed price that provider could charge, making it impossible to recover their fixed costs. In a monopolistic market, the online provider barely exceeds the benefit for the self service case, though his search cost is less than one-fourth that of self service. Note that the size of the database, 5.65, is also well below $N_o^* = 8.05$. As a monopolist, the provider covers the entire market with a zero fixed price, but revenues from marginal usage of the database more than recover their fixed costs.

<u>Effect of Online Search Costs</u>

In the component information services industry, subscriber search costs played a crucial role in the evolution of the industry. Apparently, these charges were very significant for subscribers to EnGenius, which took 45 seconds to download a datasheet image through a 9600 baud modem. As search costs for online access improve, profits for the package service provider (Figure 7) drop steadily because while fixed costs remain the same, the price that the market is

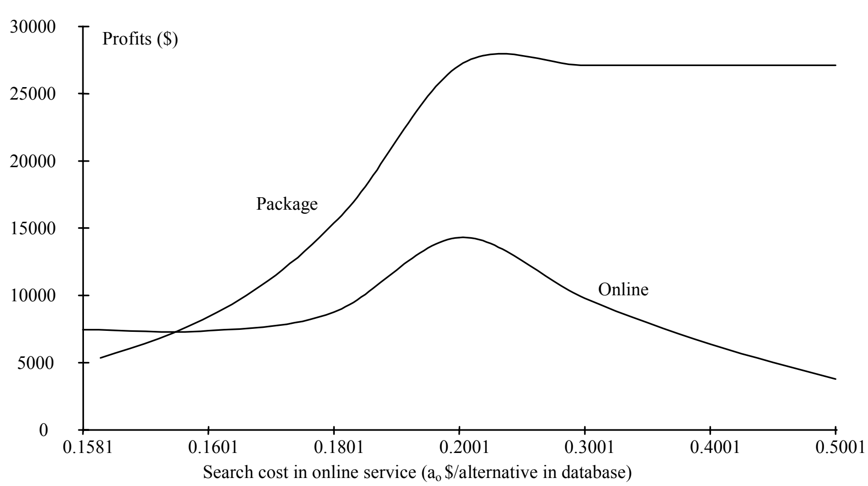

Figure 5. Effect of online search costs on profits for package and online providers in a duopoly. Search costs are linear: $C_i(N) = a_iN$, with $a_p = 0.0001$. All other parameters are given in Table 1.

willing to support drops. The profits for the online provider first increase because low volume firms switch from self service to the online service. Profits decrease because of price competition erodes profits. In the limit, when the search costs are equal, the packaged service provider stays out of the market, and the online provider becomes a monopolist.

Improvements in online technology have serious consequences for the package service provider, given their larger costs for database creation and updates. This is likely to happen in the near future, as more subscriber firms connect to high bandwidth online access. Providing online access may be the only strategic options for the packaged service provider in the future. By going online, such a firm would capture high end users through the packaged design and low end users through an online service.

7. CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS

We have examined how technology choice (online versus package) and service design parameters (database sizes, pricing) impact market segmentation and structure in information services markets by relating tradeoffs in the different options (search costs, update intervals, pricing mechanisms) to benefits for subscribers. We showed that package providers will serve high volume subscribers if they are able to ensure a low search cost that overcomes their longer update intervals compared to online services. Where the update intervals in the package design are longer and are compounded by declining online search costs, we showed that the package provider will be increasingly vulnerable to being driven out of the market by the online provider. Relative to the questions we set out to answer in our Introduction, our results for the two stage game (technology choice followed by selection of database sizes and pricing) show that both package and online can be feasible options for providers as long as they are positioned in the appropriate market segments. As for the failure of EnGenius, our results would argue that they could have been a successful business but were brought down by their service design (high price compounded by a small database), incorrect positioning (they could perhaps have targeted low volume users who could tolerate the high search costs) and premature market entry (specifically, before they had a significant-size database).

Several extensions of the present model are possible. Some hint at survival strategies for package providers faced with declining online search costs. First, markets where subscribers value different information—that is, they are indexed by both their usage volume Q, and the type of information they desire. In this case, it should be possible to show that providers will create smaller databases tailored for each type of customer segment. For a package provider, in particular, this strategy will result in higher revenues than a "one database size fits all approach". Using Figure 4, it is easy to see that a package provider may miss out low volume subscribers with a "one database size fits all approach"—this may be remedied by choosing a set of smaller database sizes (a set of lower slopes for the benefit line, with lower prices to match), effectively increasing market coverage. This is a revenue maximizing tactic Aspect Development has used—they sell several databases tailored for vertical markets such as medical instruments companies, military users and telecommunications—subscribers in these segments have specific component categories that they are interested in searching for. Second, the self-service search cost may itself be a function of subscriber characteristics. For example, $C_s \sim C_s(Q,N)$, where in the self service search costs depend on the usage volume. This may reflect the fact that large volume users may be able to create a fully functional database in-house, such as the example of Motorola in the Component Information Services example. It is possible to show that the self service benefit function will remain convex as long as $C_s(Q,N)$, is decreasing and concave in Q, which is a reasonable assumption. Of course, in this case, the package and online provider's benefit functions are also convex (not linear as before) but we expect segmentation results to hold largely unchanged. Third, database queuing delays impact a subscriber's response time (i.e. search costs). Queuing delays can be expected to be proportional to the market share for the provider and impact online subscribers more. As such, this extension would be intractable, but for simple assumptions on usage volume distribution (F(Q)) interesting results may be derived. For example, with a Binomial distribution (x% of the subscriber base with usage volume q1, 1-x% with volume q2), it would be possible to show cases where the online provider may avoid high volume users because of high search costs and increase in their fixed costs. Finally, in Bashyam and Karmarkar (2004) we have analyzed a similar setting to that of this paper, but where consumers can be indexed based on two segmentation variables: the volume (number of searches) and the value associated with each search.

REFERENCES

Bashyam, T.C.A and U.S. Karmarkar (2000), "Aspect Development Inc (A)", *Managing the Global Corporation: Case Studies in Strategy and Management*, J. De La Torre, Y.Doz and T. Devinney (editors), 2nd Edition, pp 335-348, McGraw Hill, New York.

Bashyam, T.C.A and U.S. Karmarkar (2004), "Usage Volume and Value Segmentation in Business Information Services", in *Managing Business Interfaces: Marketing, Engineering, and Manufacturing Perspectives*, Chakravarty A. and J. Eliashberg (Eds.), Kluwer Academic Publishers.

Boeri, R. and M. Hensel (1995), "What Good is SGML?", *CD-ROM Professional* v8, n4, 108-110.

Electronic Engineering Times, "Two launch component information services", n770, November 1, 1993, pp. 16-18.

Friedman, J. (1977), *Oligopoly and the Theory of Games*, North-Holland, Amsterdam.

Moorthy, K.S. (1988), "Product and Price Competition in a Duopoly", *Marketing Science* 7(2), 141-168.

Rockafellar, R.T. (1970), *Convex Analysis*, Princeton University Press, Princeton, New Jersey.

Weidemer, J.D. and D.B. Boelio (1995), “CD-ROM versus Online: An Economic Analysis for Publishers”, *CD-ROM Professional* v8, n4, 36-42.

Westland, J.C. (1989), "Topic Specific Market Concentration in the Information Services Industry: Evidence from the DIALOG Group of Databases", *The Information Society* 6, 127-138.

APPENDIX A: PROOFS OF PROPOSITIONS

PROOF OF PROPOSITION 1. (A). Observe that $W_S^*(Q)$ is the maximum over a set of functions linear in Q, and is therefore convex (Rockafellar 1970). (B). $W_S^*(Q) = U_S(N_S)Q - K_S(N_S) \geq (U_S(N_S^*)Q - K_S(N_S^*)$, from the definition of $N_S(Q)$. (C). It is sufficient to show that $N_S(Q) \rightarrow N_S^*$ for large Q. The first order condition for $W_S^*(Q)$ may be written as $Q\left(\frac{\partial U_S(N)}{\partial N} - \frac{K_S'(N)}{Q}\right) = 0$. $K_s'(N)$ is bounded above since $N_S(Q)$ is. Thus, the solution to the above first order condition approaches the solution to $Q\frac{\partial U_s(N)}{\partial N} = 0$.

PROOF OF PROPOSITION 2. Obvious from Figure 4.

PROOF OF PROPOSITION 3. Assume that if both suppliers use identical product design and pricing strategies, they split the market equally. Suppose both suppliers survive in the market, but $N_{a1} = N_{a2} = N$, $P_{a1} = P_{a2} = P$ in equilibrium. Keeping a database size of N, Firm 2 can then price at a level infinitesimally smaller than Firm 1, and capture the entire market. Firm 1 reacts likewise, and the two firms undercut each other till one firm charges the smallest price $P_0(N)$ such that $mP_0(N)MS(N,P_0(N)) - K_{ai}(N) = 0$. The other firm then stays out of the market, which reverts to a monopoly. A similar argument applies when both attempt to change their product designs while charging identical prices in equilibrium. Thus, if both firms survive, they utilize different pricing and design strategies. Clearly, the higher priced firm will then use a larger database, otherwise, it will be dominated. In addition, they must make identical profits, otherwise, the lower profit firm can mimic the size of the other firm, and undercut its price and do better. Next we need to establish that the market segments are non-adjacent/disjoint in equilibrium. Suppose an equilibrium exists where both providers make identical profits but use different database sizes and prices as described above and have adjacent market segments. Then, one provider can mimic the other's strategy (size and price), undercut price slightly and capture the entire market. Firms engage in this type of competition until both firms make zero profits. If they both make zero profits, it must be that $P_{a1} > P_{a2} > P_0(N_{a2})$, since $P_0(N_{a2})$ is the smallest price at which either firm makes a zero profit as the sole provider with a size N_{a2}. Then, Firm 1 can mimic Firm 2's size, undercut their price, and do better. When this happens, firms engage in intense price and size competition until one firm charges $P_0(N_a^*)$, while the other firm stays out, reverting the market to monopoly. Therefore, it must be that if both survive, they make identical profits, and occupy different and disjoint segments. In such a case, each provider, given the equilibrium strategy of the other firm, has no incentive to mimic the other provider because doing so will only reduce its profits.

PROOF OF COROLLARY 3.1. Take any providers two at a time and use Proposition 3.

PROOF OF PROPOSITION 4. Segment the strategy space of the online provider into **B'** = $\{(N_o, P_o, p_o)| N_{omin} \le N_o \le U_o^{-1}(U_p^*), 0 \le P_o \le P_o^{max}, 0 \le p_o \le U_o^*\}$ and **B-B'** = $\{(N_o, P_o, p_o)| N_o > U_o^{-1}(N_p^*), 0 \le P_o \le P_o^{max}, 0 \le p_o \le U_o^*\}$. For $b \in$ **B'**, the online provider will simply dominate the package provider because (1) the online provider has a better cost structure ($K_o(N)<K_p(N)$ for all N and (2) the online provider can mimic whatever strategy (database size and price) the package provider uses, undercut them and do better because the online provider gets usage revenues while the package provider does not. So, if the package provider enters at all, it has to be the case that online provider chooses a strategy in **B-B'**. In this case the slope and intercept of the online provider's benefit line satisfies (U_o-$p_o > U_p$) and $P_o > P_p$. Choosing that high end strategy should result in greater profits for the online provider compared to a strategy where they mimic the package provider, and undercut them. If not, the online provider will revert to dominating the package provider and returning the market to a monopoly. If the condition of online provider profits is satisfied, it has to be the case that online provider makes greater profits than the package provider when they serve the high end market because the online provider has lower costs than the package provider and gets additional usage revenues.

PROOF OF PROPOSITION 5. Segment the strategy space of the package provider into **A'** = $\{(N_p,P_p)|\ N_{pmin} \le N_p \le U_p^{-1}(U_o^*),\ 0 \le P_p \le P_p^{max}\}$ and **A-A'** $=\{(N_p,P_p)|\ N_p > U_p^{-1}(U_o^*),\ 0 \le P_p \le P_p^{max}\}$. For a $\in$ **A'**, the package provider will attempt to dominate the online provider by mimicking that provider's strategy (slope of the benefit line and fixed price). However, in the competition that ensues (both providers increase the slope of benefit line and drop fixed price), the package provider will be the first to drop out because (1) the online provider has a better cost structure ($K_o(N)<K_p(N)$ for all N and (2) the online provider will always have larger revenues than the package provider because of their usage revenues. Thus, if the package provider enters at all, they will pick a strategy in **A-A'**. This leaves the low end market completely open to the online provider who will enter as long at they can make a profit. The slope and intercept of the package provider's benefit line will satisfy ($U_p > U_o - p_o$) and $P_p > P_o$. Choosing this high end strategy should result in greater profits for the package provider compared to a strategy where they mimic the online provider's low end strategy, and undercut them. If not, the package provider will attempt to dominate the online provider, eventually lose and returning the market to a monopoly for the online provider. Thus, the online provider picks a strategy such that the package provider, if they were to mimic the online provider's strategy (benefit line slope and fixed price), will make less profits than if they stayed at the high end. Note that the online provider is guaranteed positive profits for this strategy since their costs are lower and revenue higher with any strategy that the package provider would imitate.

PROOF OF PROPOSITION 6. With V(N) concave, C(N) and K(N) linear in N, $QU_s(N)-K_s(N)$ is concave in N for each Q. Thus, the first order conditions are sufficient. When $Q \leq c_s / (db - a_s)$, $N_s(Q)$ = -ln((c + k/Q)/db)/b < 0, which being infeasible, we set $N_s(Q)$ = 0. Thus, $N_s(Q)$ = max{0,-ln((c + k/Q)/db)/b}, which upon substitution into the benefit function gives the proposition.

PROOF OF PROPOSITION 7. From Figures 4 and 5, it is enough to consider $U_i(N_i) \geq U_s^*$ to derive an upper bound for the fixed price. Then, for a given U_s^*, $MS_i(N_i,P_i) < MS_i(N_i,1-P_i/(U_i(N_i)Q_{max}))$. Substitute that market share bound in the monopolist packaged service provider's objective function in Section 5.1.1 along with the formula for the triangular distribution. Differentiate it with respect to P_i to obtain $P_i^{max}(N_i) = \max\{U_i(N_i)Q_{max}/3, U_i(N_i)\sqrt{(m_f Q_{max}/3)}\}$, where m_f is the mode of the triangular distribution. Substituting $U_i(N_i) = U_i^*$ gives the bound in the proposition.