# OPTIMAL BUNDLING STRATEGY FOR DIGITAL INFORMATION GOODS: NETWORK DELIVERY OF ARTICLES AND SUBSCRIPTIONS 

John Chung-I CHUANG* and Marvin A. SIRBU<br>Department of Engineering and Public Policy, Carnegie Mellon University


#### Abstract

The digitization of information goods necessitates a rethinking of their production and distribution economics. An N -good bundling model with multi-dimensional consumer preferences is developed to study the key factors that determine the optimal bundling strategy. Using analytical and empirical methods, mixed bundling is established as the dominant (i.e. profit maximizing) strategy. Pure unbundling is also shown to outperform pure bundling, even in the presence of some degree of economies of scale, if consumers positively value only a subset of the bundle components, which is the predominant case in the academic journal context. These results provide strong incentives for academic journal publishers to engage in mixed bundling, i.e. offer both individual articles and journal subscriptions, when selling and delivering over the Internet.


Keywords: Bundling, information goods, journals.

## 1. Introduction

Academic journals have traditionally been sold in the form of subscriptions. Individual articles are bundled into journal issues; issues are bundled into subscriptions. This aggregation approach has worked well in the paper-based environment, because there exist strong economies of scale in the production, distribution and transaction of journals.

Yet, the demand for scholarly information is diverse, unique, and sometimes whimsical. Scholars are often willing to expend a great deal of effort to secure a copy of a specific article unavailable from their personal subscription staple. With the proliferation of journal titles, it is impossible for every scholar to subscribe to all journals relevant to his/her work. Libraries, through their institutional subscriptions to the journals, serve to satisfy the scholars' demand for individual articles. Ordover and Willig (1978) treat journals as "sometimes shared goods" in the study of their optimal provision. Under the fair-use provision of the Copyright Act, scholars are permitted to reproduce single copies of individual articles from the library subscription copy for non-commercial purposes. There are frequent occasions, however, when the scholar's information needs go beyond the scope of the library's journal collection. In such circumstances the library is permitted to duplicate and share articles with other member libraries of an inter-library loan (ILL) consortium, as long as such "borrowing" does not lead to copying "in such aggregate quantities as to substitute for a subscription". ${ }^{2}$ Empirical studies have found that libraries are incurring costs of up to $\$ 20$ per ILL item obtained. This suggests that a potential market does exist for unbundled articles at both the individual and institutional levels.

The publishers, unable to directly appropriate charges for these forms of shared use, recompense for their loss of potential revenue by charging libraries an institutional subscription rate higher than that for individuals. This form of indirect appropriation

[^0]constitutes price discrimination of the third degree. ${ }^{3}$ While the legality of such practices are seldom questioned ${ }^{4}$, effective third degree price discrimination requires clear demarcation of market segments and minimal leakage across the segments. Both the segmentation of market and the preclusion of effective resale channels are fairly easy to enforce in the academic journal market, since institutions cannot easily disguise themselves as individual subscribers. Along with the apparent inelasticity of demand exhibited by the subscribing institutions, these have been blamed for the escalation of journal prices in recent years. ${ }^{5}$

With the global expansion and rapid commercialization of the Internet, the economics of journal publishing is quickly changing. Many publishers are experimenting with various forms of on-line access to their journals. It is now technically feasible for the publisher to electronically deliver, and charge for, individual journal articles requested by a scholar sitting at his/her desktop. The establishment of a ubiquitous electronic payment infrastructure, and the deployment of micropayment services in particular, will dramatically lower the cost of purchasing digital information goods over the Internet. From the scholars' perspective, this form of access is instantaneous, on-demand, and avoids the costs associated with traditional library access, such as travel to the library, physical duplication of the article, and congestion due to shared use of journals. But in order for network-delivery of journal articles to become a reality, economic incentives must exist for publishers to unbundle their journals.

By developing an N -good bundling model with multi-dimensional consumer preferences, this work seeks to demonstrate the existence of such incentives, and to quantify how these incentives to unbundle are affected by (i) readers' journal-reading behavior and (ii) cost trends of the underlying information network technologies. Using analytical and empirical methods, mixed bundling is established as the dominant strategy. This suggests that publishers should expand their product-line and sell individual articles in addition to journal subscriptions. By extension, a publisher with multiple journal titles should also engage in mixed bundling, offering site-licenses that are effectively 'superbundles' in addition to single-title subscriptions and individual articles.

Section 2 provides a short survey of the product bundling literature. The N-good bundling model is developed in Section 3, first the demand side in Section 3.1, followed by the supply side in Section 3.2. In Section 4, the model is applied to the academic journal industry for empirical results and analysis. Specifically, we look at how technology trends in distribution and transaction may change the supply side of the model but not change the fundamental result. We conclude with Section 5.

## 2. Economics of bundling

A multi-product monopolist may choose to bundle its goods for a variety of reasons. On the supply side, commodity bundling can result in cost savings due to the presence of economies of scale. On the demand side, bundling can be used as an effective tool for extracting consumer surplus. Both factors must be taken into account in the design of optimal bundle prices. Additionally, producers in imperfectly competitive markets may

[^1]choose to bundle their products for strategic reasons. However, bundling for strategic leverage has no direct implications on pricing design and is outside the scope of this work. ${ }^{6}$

Burnstein (1960) and Stigler (1963) are generally credited with the first references to the bundling phenomenon in the economics literature. Adams and Yellen (1976) operationalize the model for a bundle consisting of two goods, and identify three modes of bundling strategies, namely pure bundling, mixed bundling, and component selling (or pure unbundling). In pure bundling, consumers are restricted to purchasing either the entire bundle or nothing at all. In pure unbundling, no bundle is offered but consumers can put together their own bundle by buying both the component goods. Finally, a monopolist who chooses to engage in mixed bundling will allow the consumers to purchase the bundle or either one of the individual components. Consumers who choose to purchase the bundle will usually pay less than they otherwise would if they had purchased both component goods separately.

Figure 1 illustrates the consumer choice regions under each of the three bundling strategies. The axes in each plot represent the consumers' valuation for each of the two component goods, $\mathrm{G}_{1}$ and $\mathrm{G}_{2}$. An individual consumer who is willing to pay $\mathrm{w}_{1}$ for $\mathrm{G}_{1}$ and $w_{2}$ for $G_{2}$ can thus be represented as a point $\left(w_{1}, w_{2}\right)$ in this consumer space. Depending on the type of bundling strategy employed by the producer, the consumer will make the appropriate purchasing decision based upon his/her position in this twodimensional $\left\{\mathrm{W}_{1}, \mathrm{~W}_{2}\right\}$ space. For example, consumer Alice at $\left(\mathrm{a}_{1}, \mathrm{a}_{2}\right)$ will purchase only


Figure 1
Consumer choice regions for two-good bundling model

[^2]$G_{1}$ in Figure 1(a) because her willingness-to-pay (WTP) for $G_{1}, a_{1}$, is greater than its offer price $P_{1}$, but her WTP for $G_{2}, a_{2}$, is less than the offer price $P_{2}$. Bob, on the other hand, purchases nothing under this pure unbundling scenario since both $b_{1}$ and $b_{2}$ are less than $P_{1}$ and $\mathrm{P}_{2}$ respectively. Interestingly, the situation almost reverses itself if the producer switches to pure bundling instead, as in Figure 1(b). Alice rationally chooses to purchase nothing since her aggregate WTP $\left(a_{1}+a_{2}\right)$ is less than the price of the bundle, $\mathrm{P}_{\mathrm{B}}$. Bob now purchases the bundle since the sum of $b_{1}$ and $b_{2}$ is greater than $P_{B}$. Using similar logic, Alice consumes $\mathrm{G}_{1}$ and Bob consumes the bundle in the mixed bundling case, as illustrated in Figure 1(c). This simple, yet powerful illustration shows that the choice of the optimal bundling strategy and the selection of the optimal prices are strongly dependent on the distribution of the consumer population in this $\left\{\mathrm{W}_{1}, \mathrm{~W}_{2}\right\}$ space.

Schmalensee (1982) and McAfee, McMillan and Whinston (1989) build upon the Adams/Yellen framework, with careful treatment of the consumers' correlation of value between the two components. Among other results, they show that both pure bundling and pure unbundling are boundary cases of mixed bundling and are weakly dominated by the latter strategy in general. Chae (1992) applies the commodity bundling model to information goods in his study of the subscription TV market. He concludes that the bundling of CATV channels is practiced not to extract consumer surplus, but simply because there are economies of scope in the distribution technology.

## 3. $N$-good bundling model

All of the above-mentioned works are limited to bundles consisting of only two components. A typical academic journal, on the other hand, has between 80 to 100 articles per subscription period. An appropriate N -good bundling model is needed for this context. Unfortunately, a complete N -good model with $2^{\mathrm{N}}$ bundle combinations and N -dimensional consumer preferences quickly becomes computationally unwieldy as N gets large. Hanson and Martin (1990), by formulating the model as a mixed integer linear programming problem, manage to attack a bundle pricing problem with $\mathrm{N}=21 .^{7}$

Recognizing the need to balance profit-maximization and consumer rejection of a complex pricing schedule, we opt for a simpler model where no sub-bundles are available. The consumer either purchases the journal subscription for a price $P_{J}$ or individual articles at a price of $P_{A}$ apiece. This simplifies the model from that of setting $2^{\mathrm{N}}$ optimal prices to setting only two prices, $\mathrm{P}_{\mathrm{A}}$ and $\mathrm{P}_{\mathrm{J}}$. This is reminiscent of setting a menu of optional twopart tariffs in the nonlinear pricing literature (Willig, 1978 and Wilson, 1993). ${ }^{8}$ Lowdemand readers purchase articles individually, while high-demand readers pay the flat fee $P_{\mathrm{J}}$ and enjoy unlimited access to all articles (Figure 2).

[^3]

Figure 2
Total outlay vs. number of articles consumed

Optional two-part tariffs can be either ex ante or ex post in nature (Mitchell and Vogelsang, 1991, page 95). In an ex ante arrangement, readers elect to join either the subscriber group or the "article-on-demand" group prior to consumption. Knowing one's expected consumption behavior is critical in making the "right" decision. An "article-ondemand" reader who expects to read only a few articles but ends up reading more than $\mathrm{N}_{\mathrm{c}}$ ( $=\mathrm{P}_{\mathrm{J}} / \mathrm{P}_{\mathrm{A}}$ ) articles would have to pay more than if he/she had become a subscriber in the first place. Many consumers (especially those with fixed budgeting and fund allocation considerations) are reluctant to sign up for these pay-per-use arrangements precisely because of this uncertainty factor. An ex post approach eliminates this problem by allowing the consumer to choose the cheaper of the two pricing schemes at the end of the billing period, thereby placing a predictable upper bound on the final bill. However, the need for a final settlement incurs an administrative and metering overhead over true pay-per-use models.

### 3.1. Modeling heterogeneity in consumer preferences

The N-good bundling model departs from the traditional nonlinear pricing model in that consumers are not choosing to purchase n units of non-distinguishable articles, as if purchasing x kilowatt-hours of electricity or y minutes of cellular-phone air-time. Instead, each of the N articles is unique and distinct from one another. Consumers may value one article dramatically differently from the next. Unfortunately, a complete description of consumer heterogeneity using an N -dimensional vector $\left\{\mathrm{w}_{1}, \mathrm{w}_{2}, \ldots, \mathrm{w}_{\mathrm{N}}\right\}$ again leads to intractability. We seek a concise way to capture the essence of consumer's willingness-topay across the different articles.

Zahray and Sirbu (1990) attempt to capture the heterogeneity in consumer preferences for academic journals, albeit in one variable, the reservation price for the journal. A similar approach is taken by Bakos and Brynjolfsson (1997), where consumers are characterized by a single type variable $w$, and consumer valuations of goods are i.i.d. (independent and identically distributed). By employing a single variable, both models can only capture consumer valuations for the bundle in its aggregate. This is adequate in the pure bundling context, where journals are sold only in the form of subscriptions. In the mixed bundling context, however, it is important to account for the correlation of values across the components as well.

Consider, for example, a publisher selling a two-article journal in a market with only two consumers, our friends Alice and Bob. Alice is willing to pay $\$ 10$ for the first article and $\$ 0$ for the second, while Bob values the articles at $\$ 7$ and $\$ 5$ respectively. A publisher engaging exclusively in pure bundling (i.e. subscription only) is only interested in the aggregate willingness-to-pay of the two consumers. He/she will price the subscription at $\$ 10$ for a total revenue of $\$ 20$. A mixed bundling strategist, on the other hand, will desire additional information on the correlation of values for the component articles. In this example, the publisher will price individual articles at $\$ 10$ and raise the subscription price to $\$ 12$, thereby realizing a revenue of $\$ 22$ and completely extracting the consumer surplus in the process. In effect, the publisher has managed to separate the market into two -- the segment with high correlation of value across articles (Bob) is sold the subscription; the segment with low correlation (Alice) is offered individual articles.

The present work employs two variables, $\mathrm{w}_{\mathrm{o}}$ and k , to describe the N -dimensional consumer preference. We allow each journal reader to rank the N articles in the journal in decreasing order of preference, such that his/her favorite article is ranked first, the least favorite is ranked last, and weak monotonicity is observed. The reader may place zero value on any number of the N articles. By assuming a linear demand function for all positive-valued articles, we can plot an individual reader's valuation of all the articles in the journal in Figure 3. Each of the articles are positionally ranked between 0 and N along the horizontal axis. The individual's most highly valued article has $\mathrm{n}=0$, and so the y intercept, $\mathrm{w}_{\mathrm{o}}$, represents the WTP for his/her most favored article. The valuation for the subsequently ranked articles is assumed to fall off at a constant rate until it reaches zero at n $=\mathrm{k} \cdot \mathrm{N}$. No articles have negative value with the assumption of free disposal -- readers are free to discard unwanted articles at zero cost. The variable k dictates the slope of the demand curve, and it also indicates the fraction of articles in the journal that has non-zero value to the individual. For example, a reader with $\mathrm{k}=0.01$ is willing to pay a non-zero amount for only one article in a journal with a hundred articles, while another reader who positively values half of the articles in the journal will have a k of 0.5 . If an individual's k is greater than unity, that means he/she places positive value on all N articles in the journal and the demand curve will never cross the horizontal axis. Figure 4 shows a diverse range of consumer preferences that can be described using this two-dimensional $\left\{\mathrm{w}_{\mathrm{o}}, \mathrm{k}\right\}$ index.


Figure 3
Article valuation by an individual reader indexed by $\left\{\mathrm{w}_{\mathrm{o}}, \mathrm{k}\right\}$


Figure 4
Diversity of consumers indexed by $\left\{\mathrm{w}_{\mathrm{o}}, \mathrm{k}\right\}$
Empirical studies performed by King and Griffiths (1995) indicate that the correlation of article valuations is not very high for academic journals (Table 1). Out of the 80 to 100 articles (per subscription period) in an average journal, over $40 \%$ of readers surveyed read no more than five articles. Only $0.9 \%$ of readers read more than 50 , or about half of all articles in the subscription period. This suggest that a majority of readers have small k's, and very few readers have their k-value close to or exceeding unity. This result is incorporated into our analysis below as a fitted probability distribution for $\mathrm{k}, \mathrm{f}_{\mathrm{k}}(\mathrm{k})$.

TABLE 1
Distribution of number of articles read in a journal (King and Griffiths [1995])

| Number of Articles <br> Read in a Journal | Proportion of <br> Readers (\%) | Cumulative Proportion <br> of Readers (\%) |
| :---: | :---: | :---: |
| 1 to 5 | 43.60 | 43.60 |
| 6 to 10 | 34.40 | 78.00 |
| 11 to 15 | 8.21 | 86.21 |
| 16 to 20 | 5.50 | 91.71 |
| 21 to 25 | 3.37 | 95.08 |
| 26 to 30 | 1.97 | 97.05 |
| 31 to 40 | 1.23 | 98.28 |
| 41 to 50 | 0.82 | 99.10 |
| more than 50 | 0.90 | 100.00 |

Formally, an individual's valuation for the n-th article can be expressed as:

$$
\begin{equation*}
w(n)=\max \left\{0, w_{o} \cdot\left[1-\frac{1}{k}\left(\frac{n}{N}\right)\right]\right\}, \quad 0 \leq \mathrm{n} \leq \mathrm{N}-1 \tag{1}
\end{equation*}
$$

with $\mathrm{w}_{\mathrm{o}} \geq 0$ and $\mathrm{k} \geq 1 / \mathrm{N}$. Using this formulation, we can proceed to determine the individual's reservation price of the journal, his/her consumption decision in face of the
prices $\mathrm{P}_{\mathrm{A}}$ and $\mathrm{P}_{\mathrm{J}}$, and the optimal number of articles consumed in each of the three bundling scenarios.

### 3.1.1. Consumer choice in pure bundling

In pure bundling, potential readers can only choose to subscribe to the journal or buy nothing at all. Purchasing individual articles is not an available option. Therefore, an individual's decision is based solely on the price of the subscription, $\mathrm{P}_{\mathrm{J}}$, and his/her reservation price of the journal bundle in the aggregate. This reservation price, $\mathrm{W}_{\mathrm{J}}$, is simply the summation ${ }^{9}$ (or integration if we approximate n as a continuous variable) of his/her reservation prices for all the individual articles:

$$
\begin{equation*}
W_{J}=\int_{0}^{N} w(n) \cdot d n \tag{2}
\end{equation*}
$$

The net benefit $U_{J}$ derived from subscribing is the difference of the reservation price $W_{J}$ and the actual subscription price $\mathrm{P}_{\mathrm{J}}$ :

$$
\begin{equation*}
\mathrm{U}_{\mathrm{J}}=\mathrm{W}_{\mathrm{J}}-\mathrm{P}_{\mathrm{J}} . \tag{3}
\end{equation*}
$$

A potential reader will only choose to subscribe if the subscription results in a positive net benefit $U_{J}>0$. The $U_{J}=0$ curve, plotted in $\left\{w_{o}, k\right\}$ space in Figure 5, separates the readership population into two regions. Those that fall in the region $R_{J}$ will choose to subscribe, while those in region $\mathrm{R}_{0}$ will opt out. Please refer to the Appendix for derivation of this and subsequent results.


Figure 5
Consumer choice in pure bundling scenario

[^4]
### 3.1.2. Consumer choice in pure unbundling

In the pure unbundling scenario, all articles are available individually at a unit price of $\mathrm{P}_{\mathrm{A}}$. Consumers are free to purchase as many or as few articles as they desire, up to and including all N articles in the journal. A rational-choice utility-maximizing consumer will consume only those articles with $w(n) \geq P_{A}$, realizing a net benefit of $w(n)-P_{A}$ for each of those articles. The marginal article consumed by the consumer, $\mathrm{n}^{*}$, has a benefit $\mathrm{w}\left(\mathrm{n}^{*}\right)=$ $P_{A}$. Therefore, for $w_{o} \geq P_{A}$, the optimal number of articles read by an individual indexed by $\left\{\mathrm{w}_{\mathrm{o}}, \mathrm{k}\right\}$ can be expressed as

$$
\begin{equation*}
n^{*}=\min \left\{N, \quad \frac{k \cdot N \cdot\left(w_{o}-P_{A}\right)}{w_{o}}\right\}, \tag{4}
\end{equation*}
$$

with the maximum capped at N , the total number of articles available in the journal. On the other hand, for an individual with $\mathrm{w}_{0}<\mathrm{P}_{*}$, even the most favored article is deemed unworthy of the price tag $\mathrm{P}_{\mathrm{A}}$. In this case, $\mathrm{n}^{*}$ would be equal to zero and no articles will be purchased. Figure 6 presents the optimal article consumption level in $\left\{w_{0}, k\right\}$ space. In addition to the optimal consumption level, the net benefit derived from consuming $\mathrm{n}^{*}$ articles, $\mathrm{U}_{\mathrm{A}}$, can also be expressed as

$$
\begin{equation*}
\mathrm{U}_{\mathrm{A}}=\mathrm{W}_{\mathrm{A}}-\mathrm{n}^{*} \cdot \mathrm{P}_{\mathrm{A}}, \tag{5}
\end{equation*}
$$

where the gross benefit, $W_{A}$, is itself a function of $\mathrm{n}^{*}$ :

$$
\begin{equation*}
W_{A}=\int_{0}^{n^{*}} w(n) \cdot d n . \tag{6}
\end{equation*}
$$



Figure 6
Optimal article consumption level in pure unbundling scenario

### 3.1.3. Consumer choice in mixed bundling

In mixed bundling, consumers seek to maximize their utility by choosing one of three options: subscribe to journal, purchase individual articles, or neither. Depending on each individual's $U_{\mathrm{I}}$ and $U_{A}$ measures, as defined above, he/she can fall into one of five regions in Table 2. This is illustrated in the consumer choice diagram, Figure 7. For example, individuals who value their most favored article at less than the article price (i.e. $\mathrm{w}_{\mathrm{o}}<\mathrm{P}_{\mathrm{A}}$ ) have a negative $\mathrm{U}_{\mathrm{A}}$ and will not purchase any articles in unbundled form. If their valuation of all the articles in the aggregate is less than the subscription price $P_{J}$, they will not subscribe to the journal either. These individuals fall in the $\mathrm{R}_{0}$ region. On the other hand, if their aggregate valuation is greater than $P_{\mathrm{I}}$, they will fall in the $\mathrm{R}_{\mathrm{J1}}$ region and will choose to subscribe to the journal. Individuals with high $w_{o}$ and low $k$ tend to value only a few articles highly, and will be best off purchasing individual articles. These consumers are found in region $\mathrm{R}_{\mathrm{A} 1}$. Finally, consumers in $\mathrm{R}_{\mathrm{A} 2}$ and $\mathrm{R}_{\mathrm{J} 2}$ receive positive benefits from either journal subscription or article purchase, and make their respective purchasing decisions based on the relative magnitudes of their $U_{J}$ and $U_{A}$.

TABLE 2
Consumer choice in mixed bundling scenario

| Region | $\mathrm{U}_{\mathrm{J}}$ | $\mathrm{U}_{\mathrm{A}}$ | $\mathrm{U}_{\mathrm{J}}>\mathrm{U}_{\mathrm{A}}$ ? | Purchase |
| :---: | :---: | :---: | :---: | :---: |
| $\mathrm{R}_{0}$ | $<0$ | $<0$ | -- | Nothing |
| $\mathrm{R}_{\mathrm{A} 1}$ | $<0$ | $>0$ | No | Article(s) |
| $\mathrm{R}_{\mathrm{J} 1}$ | $>0$ | $<0$ | Yes | Journal |
| $\mathrm{R}_{\mathrm{A} 2}$ | $>0$ | $>0$ | No | Article(s) |
| $\mathrm{R}_{\mathrm{J} 2}$ | $>0$ | $>0$ | Yes | Journal |



Figure 7
Consumer choice in mixed bundling scenario

### 3.2. Production costs and economies of scale

Thus far we have focused on the demand side of the problem. We now turn to the supply side, specifically to the underlying technology and production functions of academic journals. As previously noted, the information industry in general and the journal publishing industry in particular are characterized with high fixed costs (FC) and low marginal costs (MC). A producer will only stay in the market if gross margin (gross revenue minus variable cost) is enough to cover fixed cost. As long as the total revenue is greater than total cost, the optimal pricing decision is then independent of FC. (Alternatively, we can think of the fixed cost as either zero or sunk). This assumption allows the treatment of FC as an exogenous variable in the present model.

We incorporate the presence (or absence) of economies of scale (EoS) in the production function by establishing the following relationship between the marginal costs $\mathrm{MC}_{\mathrm{J}}$ and $\mathrm{MC}_{\mathrm{A}}$ :

$$
\begin{equation*}
\mathrm{MC}_{\mathrm{J}}=\mathrm{N}^{\gamma} \cdot \mathrm{MC}_{\mathrm{A}} . \tag{7}
\end{equation*}
$$

N is the number of articles in the journal and $\gamma$ is the economies of scale index. When $\gamma<$ 1 , economies of scale are present and a subscription bundle of N articles is cheaper to produce and sell than N individual articles. Therefore the publisher can realize cost savings via bundling. When $\gamma=1$, there are no economies of scale in journal production or distribution. No cost savings can be realized by bundling. Finally, if there are diseconomies of scale in the production function, it can be described with $\gamma>1$. Prior work in bundling almost invariably assumes no cost savings from bundling, i.e., $\gamma=1$. Chae's assumption of extreme economies of scope in the CATV delivery technology translates to a special case of $\gamma=0$. By treating the extent of economies of scale as an endogenous variable, this model allows a parametric analysis of its influence on the producer's optimal bundling strategy.

Based upon the distribution of consumers in the $\left\{\mathrm{w}_{\mathrm{o}}, \mathrm{k}\right\}$ space and the underlying cost structure of journal production, the publisher proceeds to optimize $P_{A}$ and $P_{J}$ to maximize gross margin $\Pi$ :

$$
\begin{equation*}
\prod\left(P_{J}, P_{A}\right)=\iint_{R J}\left[P_{J}-M C_{J}\right] f\left(w_{o}, k\right) \cdot d w_{o} \cdot d k+\iint_{R A} n^{*}\left[P_{A}-M C_{A}\right] f\left(w_{o}, k\right) \cdot d w_{o} \cdot d k, \tag{8}
\end{equation*}
$$

where the term $\mathrm{f}\left(\mathrm{w}_{\mathrm{o}}, \mathrm{k}\right)$ is the joint probability density function (p.d.f.) of the readership described in $\left\{\mathrm{w}_{\mathrm{o}}, \mathrm{k}\right\}$ space. It is worthwhile to note that, in the case where the optimal strategy turns out to be pure bundling (pure unbundling), the second (first) integral component will be zero.

## 4. Analysis and empirical results

The N -good bundling model is used to quantify how the choice of the optimal bundling strategy and optimal pricing are affected by MC and $\gamma$ on the supply side, and $\mathrm{f}\left(\mathrm{w}_{\mathrm{o}}, \mathrm{k}\right)$ on the demand side. Recalling that $\mathrm{w}_{\mathrm{o}}$ is an individual's valuation of his/her most favored article in the journal, and k is the fraction of articles in the journal that have nonzero value to the individual, we assume independent distributions for $\mathrm{w}_{\mathrm{o}}$ and k . We normalize $w_{0}$ to be uniformly distributed between 0 and 1. Using the King/Griffiths data in Table $1, \mathrm{k}$ is fitted to an exponential distribution with $\lambda=13.8758$ or $\mu=1 / \lambda=0.072\left(\mathrm{R}^{2}\right.$
$=0.97117$ ). This means that the average reader reads only $7.2 \%$ of all articles in a typical journal. Figure 8 shows, for a journal with $\mathrm{N}=100$ articles, the producer surplus (as measured by gross margin) attainable via each bundling alternative as a function of MC and
$\gamma$. The marginal cost of a single article, MC, is restricted to be no greater than the highest individual valuation, $\max \left[\mathrm{w}_{\mathrm{o}}\right]$ (which we normalize to unity without loss of generality). There would be no market participation if it were more costly to produce an article than anyone is willing to pay.


Figure 8
Profit-maximizing bundling strategy: it is clear that mixed bundling is the dominant strategy across all marginal cost and economies of scale conditions.

In Figure 8(a), there are no economies of scale and the EoS factor $\gamma=1$. The marginal cost of the journal is N times that of a single article, and so no cost saving is realizable from bundling. The pure bundling strategy is clearly dominated by the other two strategies. The mixed bundling and the pure unbundling alternatives are essentially identical to each other. This suggests that even if the publisher opts for a mixed bundling strategy, virtually the entire revenue will come from article sales. As the production function begins to exhibit some economies of scale, cost-related bundling incentives begin to appear. Yet, in Figure 8(b), where $\gamma=0.75$, the situation remains unchanged. Mixed
bundling and pure unbundling continue to be the optimal strategies. As $\gamma$ continues to fall in the face of stronger economies of scale, mixed bundling becomes the strictly dominant strategy. In Figure $8(\mathrm{c})$, where $\gamma=0.5$, pure bundling and pure unbundling trade dominance depending on the magnitude of MC , but both are dominated by mixed bundling.
Finally, in the case of extreme economies of scale, where $\gamma=0$ in Figure 8(d), it costs as much (or as little) to produce and sell an entire journal as it does a single article. Pure unbundling becomes inferior to bundling of either form as MC becomes comparable in magnitude to the consumer article valuations. As MC approaches max [ $\mathrm{w}_{\mathrm{o}}$ ] or unity, pure bundling performs as well as mixed bundling. In this case most of the publisher's revenue will be derived from journal subscriptions.

The first conclusion is that mixed bundling is superior to pure bundling and pure unbundling across all values of MC and $\gamma$. This extends earlier results for two-good models to the present N -good model. This result makes intuitive sense since both pure bundling and pure unbundling are boundary cases of mixed bundling, and therefore can do no better than the mixed bundling strategy. The price discrimination mechanism is at work here, as the mixed bundling strategy creates an incentive-compatible condition, inducing the high and low-demanders to reveal their preferences by self-selecting into the appropriate consumption groups.

Secondly, we observe that pure bundling does not necessarily dominate over pure unbundling in the N -good scenario. Specifically, the model identifies plausible conditions under which unbundling is actually superior to bundling (in pure forms). When marginal cost is non-zero, pure bundling is undesirable not only in the absence of economies of scale $(\gamma=1)$, but also if the degree of EoS is too weak (as illustrated by $\gamma=0.75$ ) for the costsaving bundling incentive to become a dominating factor. Even in the presence of strong economies of scale $(\gamma=0.5,0)$, the relative merits of pure bundling and unbundling are still dependent on the magnitude of the marginal cost relative to consumer valuations of the articles. Inefficiency in resource allocation (and loss of surplus) will result if individuals are forced to purchase the bundle and consume some articles which they value below marginal cost. Adams and Yellen label this condition where consumption occurs at subMC levels as a violation of the 'Exclusion' assumption. This is of real concern to journal publishers since the distribution of k (as fitted to empirical data from King and Griffiths) is such that most readers actually place zero value on most of the articles in an average journal that they read. Except for the case of $\mathrm{MC}=0$, or the case of $\gamma=0$, where the marginal costs for all but the first article are effectively zero, exclusion is always violated for those readers with $\mathrm{k}<1$. In our numerical analysis, where k is exponentially distributed with a mean $\mu=0.072$, the probability of $\mathrm{k} \geq 1$, i.e. a reader having positive valuations for all articles in the journal, is on the order of $10^{-6}$, or one out of one million readers. (To place this number in context, Science, one of the most widely read academic journals, has a circulation of 165,000; IEEE Spectrum and American Economic Review, two mainstream periodicals in the electrical engineering and economics disciplines, have circulation of 30,000 and 27,000 respectively. ${ }^{10}$ ) Therefore, the choice of optimal bundling strategy lies in the balance between cost-savings from bundling and loss of surplus due to exclusion violation. The proposition by Adams and Yellen (p. 488) that pure unbundling "is a more desirable strategy the greater the cost of violating Exclusion" holds true here.

[^5]
### 4.1. Optimal pricing and revenue mix

The mixed bundling publisher is interested in the optimal pricing of its articles and subscriptions. Figures 9 and 10 show the optimal pricing ratio $\left(\mathrm{P}_{\mathrm{J}} / \mathrm{P}_{\mathrm{A}}\right)$ and the corresponding revenue mix for various marginal cost and EoS conditions, respectively. We observe that when marginal cost is negligible ( $\mathrm{MC}=0$ ), the subscription (to a bundle of 100 articles) should be priced approximately ten times that of an individual article, and this optimal pricing ratio would result in a revenue stream that is well balanced between the sale of articles and subscriptions. When the marginal cost is non-negligible, however, the optimal ratio becomes sensitive to the economies of scale condition. If there are extreme economies of scale $(\gamma=0)$, the cost-saving incentive induces the publisher to rely more heavily on the sale of bundled subscriptions as MC increases. With strong economies of scale ( $\gamma=0.5$ ), the optimal pricing ratio stays constant but the revenue mix shifts decisively towards subscription sales with increasing cost. On the other hand, when the economies of scale are absent or weak $(\gamma=1,0.75)$, the publisher is best served by increasing the price ratio, thereby realizing essentially all of its revenue through individual article sales.


Figure 9
Optimal price ratio for mixed bundling strategy across various EoS conditions


Figure 10
Optimal revenue mix for mixed bundling strategy

### 4.2. Internet-based document delivery technology

We can characterize the extent to which economies of scale are present in the current set of network-based document delivery technologies. Specifically, we ask what is a reasonable value of $\gamma$, and how might it change with technology? We identify two major components to the marginal cost of delivering a journal or an article. These are the cost to transmit raw data bits and transaction costs. Production and data storage are fixed costs to the publisher and should be excluded from consideration in this context.

We consider the scenario where the publisher outsources both data transmission and fee-collection functions to specialized services. Web hosting services are offered by a multitude of Internet presence providers. Entire digitized archives of journal articles can be hosted on a web server and made accessible for downloading by scholars. Several micropayment systems are also available to facilitate electronic payment for articles or other information goods sold via the Internet. ${ }^{11}$

We choose to characterize the marginal costs $\mathrm{MC}_{\mathrm{J}}$ and $\mathrm{MC}_{\mathrm{A}}$ using three cost coefficients:

$$
\left\{\begin{array}{l}
M C_{J}=\kappa_{f}+\kappa_{v} \cdot P_{J}+\mu_{s} \cdot N \cdot \kappa_{d}  \tag{9}\\
M C_{A}=\quad \kappa_{f}+\kappa_{v} \cdot P_{A}+\kappa_{d}
\end{array}\right.
$$

where $\kappa_{\mathrm{f}}, \kappa_{\mathrm{v}}, \kappa_{\mathrm{d}}$ are cost coefficients and $\mu_{\mathrm{s}}$ is the expected fraction of articles downloaded by a subscriber. We discuss each variable in turn. Transactional costs are modeled after the two-part fee structure of credit-card transactions. $\kappa_{\mathrm{f}}$ is a fixed fee levied for each transaction, while $\kappa_{\mathrm{v}}$ is the variable component charged in proportion to the value of the

[^6]transaction ( $\mathrm{P}_{\mathrm{J}}$ and $\mathrm{P}_{\mathrm{A}}$ respectively). ${ }^{12}$ This implies, significantly, that the marginal costs are no longer constants as we have assumed thus far, but have become functions of $P_{J}$ and $P_{A}$, respectively.

The variable $\kappa_{d}$ is the cost of transmitting or downloading one journal article. Web hosting services currently charge between $\$ 0.05$ and $\$ 0.50$ per MB (megabyte) of data accessed by a client from the server. ${ }^{13}$ A journal page, scanned at 600 dpi and compressed in Group 4 Fax/TIFF format, takes up about 100kB (kilobytes). Assuming a typical journal article has ten pages, downloading a journal article requires the transmission of 1 MB of data. This translates to a $\kappa_{\mathrm{d}}$ of between $\$ 0.05$ and $\$ 0.50$. With continued improvements in data transmission and compression technologies, it is reasonable to expect further declines in $\kappa_{d}$.

Most providers sell downloads at a fixed cost per bit, so the publisher enjoys no economies of scale in data transmission per se. However, selling a journal subscription on-line does not necessarily require the transmission of all N articles to the subscriber. The subscribers are free to download all N articles, but most will choose to download only a fraction of all articles. This "just-in-time" (as opposed to "just-in-case") delivery paradigm results in an expected transmission cost of $\mu_{s} \cdot N \cdot \kappa_{d}$ instead of $N \cdot \kappa_{d}$ for each journal subscription. We can quantify $\mu_{\mathrm{s}}$ as the conditional expectation of the fraction of articles read by the subscribing sub-population (the region $\mathrm{R}_{\mathrm{J}}$ in Figure 7),

$$
\begin{equation*}
\mu_{s}=\frac{\iint_{R J} k \cdot f\left(w_{o}, k\right) \cdot d w_{o} \cdot d k}{\iint_{R J} f\left(w_{o}, k\right) \cdot d w_{o} \cdot d k} . \tag{10}
\end{equation*}
$$

We have shown that the area of integration $R_{J}$ is a function of the prices set by the publisher. Therefore $\mu_{\mathrm{s}}$ is dependent on the prices as well. Substituting equations (9) and (10) into equation (8) with the appropriate values for the $\kappa$ coefficients and re-optimizing, we can gain insight into how $\mu_{\mathrm{s}}$ and $\gamma$ are affected by a decline in transmission cost $\kappa_{\mathrm{d}}$, which in turn determine the optimal pricing and revenue mix decisions. Figure 11 shows that the optimal subscription price $\mathrm{P}_{\mathrm{J}}$ (right hand axis) varies significantly in the current range of $\kappa_{d}$. The expected fraction of articles read from a subscription copy $\mu_{s}$ (left hand axis) follows a similar trend, which is not surprising given its dependency on $\mathrm{P}_{\mathrm{J}}$. The higher the price of a subscription, the more articles one will have to read in order to justify becoming a subscriber. It is interesting to note that, even when transmission costs become negligible ( $\kappa_{d}=0$ ), $\mu_{\mathrm{s}}$ is still significantly greater than $\mu$ of 0.072 for the overall journal readership.

[^7]

Figure 11
Effect of transmission cost on journal subscription pricing
Figure 12 further illustrates how the economies of scale $(\gamma)$ and optimal revenue mix are likely to be impacted by a declining $\kappa_{\mathrm{d}}$. For $\kappa_{\mathrm{d}}$ greater than $\$ 0.20 / \mathrm{MB}$, there are essentially no economies of scale and most of the revenue is derived from article sales. At $\kappa_{d}=\$ 0.05 / \mathrm{MB}$, the current low-end estimate, $\gamma$ falls to 0.6 and we begin to see a well balanced revenue mix between article and journal sales. But even when $\kappa_{d}=0$, we see that $\gamma$ will not fall below 0.3 , and $30 \%$ of the revenue is still derived from selling individual articles. Under no circumstance should we expect the entire publishing revenue to come from subscription sales alone.


Figure 12
Effect of declining $\kappa_{d}$ on economies of scale and revenue mix

While we have held $\kappa_{\mathrm{f}}$ and $\kappa_{\mathrm{v}}$ constant in our analysis, it is reasonable to expect a decline in these coefficients as well. The Millicent protocol, for example, proposes a lightweight micropayment mechanism with cryptographic operations that cost one-tenth to onehundredth of a cent (Manasse, 1995). Yet one should not expect $\kappa_{\mathrm{f}}$ and $\kappa_{\mathrm{v}}$ to fall at a similar rate as $\kappa_{d}$. This is because transaction costs are not solely dictated by progress in hardware technology or the state of the art in cryptography. Other sources of payment system costs such as customer service, fraud protection, chargebacks and back-office accounting may decline only slowly over time, if at all.

## 5. Conclusions

Several recent independent works suggest that bundling is desirable for information goods (Bakos and Brynjolfsson, 1997; Fishburn et al., 1997; Varian, 1995). The current work demonstrates, however, that a different conclusion may be drawn when the important distinction between mixed and pure bundling is made. While mixed bundling is always the dominant strategy, our results also show that pure bundling may, under certain conditions, be inferior to pure unbundling. We therefore caution against any wholesale adoption of pure bundling without a thorough analysis of the supply and demand of the information product in question. Specifically, for information goods that presently exist in bundled form (e.g. academic journals), unbundling (i.e. switching from pure bundling to mixed bundling) can actually increase producer surplus. This result suggests that an academic journal publisher should expand its on-line product offering to include unbundled articles in addition to traditional subscriptions. By offering a menu of choice that includes both the original bundle and the components, the publisher can extract consumer surplus more completely via consumer self-selection. By extension, the publisher can do even better by simultaneously bundling and unbundling the journal, adding "super-bundles" of multiple journal subscriptions or site-licenses to the product mix. Mackie-Mason and Riveros (1997) offer another bundle option in addition to unbundled articles and the traditional subscription, namely the generalized subscription. In this arrangement, the user purchases unlimited access to N units of articles, and is free to select any N articles from the entire archive of M articles (with $\mathrm{M} \gg \mathrm{N}$ ).

Our model assumes that a journal is made up of N individual articles. In reality there are other separable components to a journal subscription, such as the table of content, indices, abstracts and other announcements. Readers can assign different valuations for each of these components just as they do for the individual articles. Therefore these components can be candidates for unbundling as well. RevealAlert, a recent product offered by CARL, delivers via email the tables of contents of up to fifty user-selected journal titles.

A casual survey will reveal that all the major players in the academic journal publishing industry are actively pursuing the possibility of network access to their journal products. Many have made impressive strides in a very short period of time. Some publishers provide on-line access to article abstracts, tables of content and indices to their journal titles; others offer fully searchable text, complete with images and mark-up tags, of the journal articles. Most publishers have installed (or plan to install) some form of access control and billing mechanism so that charges can be appropriated for the usage of these materials. However, lessons learnt from various research/demonstration projects indicate that significant economic, behavioral and institutional barriers need to be crossed before on-
demand network delivery of academic journals can become ubiquitous. ${ }^{14}$ Intelligent pricing designs must take into consideration the information needs and usage behavior patterns of the journal reading population, as well as the economies-of-scale characteristics of the underlying technologies.

## Acknowledgment

This work is supported by the National Science Foundation under Cooperative Agreement No. IRI-9411299 and by the Council of Library and Information Resources A. R. Zipf Fellowship. The authors wish to thank Lester Lave, Fallow Sowell, seminar participants at Harvard and CMU, and two anonymous referees for comments on an earlier version of this paper.

[^8]
## Appendix A. Derivation of producer surplus for alternative bundling strategies

A profit-maximizing journal publisher will seek to optimize the prices $\mathrm{P}_{\mathrm{J}}$ and $\mathrm{P}_{\mathrm{A}}$ by maximizing the objective function $\Pi$, restated here from equation (8):

$$
\begin{equation*}
\prod\left(P_{J}, P_{A}\right)=\iint_{R J}\left[P_{J}-M C_{J}\right] f\left(w_{o}, k\right) \cdot d w_{o} \cdot d k+\iint_{R A} n^{*}\left[P_{A}-M C_{A}\right] f\left(w_{o}, k\right) \cdot d w_{o} \cdot d k \tag{A.1}
\end{equation*}
$$

To derive the gross margin or producer surplus attainable from each of the three alternative bundling strategies, we need to identify the regions $R_{J}$ and/or $R_{A}$ in each scenario. This allow the limits of integration for the definite integrals to be quantitatively specified.

Additionally, the p.d.f. of the journal reading population in $\left\{\mathrm{w}_{\mathrm{o}}, \mathrm{k}\right\}$ space has to be specified. The assumption of independence between random variables $\mathrm{w}_{\mathrm{o}}$ and k , and the choice of the probability distributions gives $f\left(w_{0}, k\right)=f_{w o}\left(w_{o}\right) \cdot f_{k}(k)$, where

$$
\begin{align*}
& f_{w o}\left(w_{o}\right)= \begin{cases}1 & 0 \leq w_{o} \leq 1 ; \\
0 & \text { elsewhere } ;\end{cases}  \tag{A.2}\\
& f_{k}(k)=\lambda \cdot e^{-\lambda\left(k-\frac{1}{N}\right)} \quad k \geq \frac{1}{N} . \tag{A.3}
\end{align*}
$$

From equation (4), $\mathrm{n}^{*}$, the optimal number of articles consumed, is

$$
\begin{equation*}
n^{*}=\min \left\{N, \frac{k \cdot N \cdot\left(w_{o}-P_{A}\right)}{w_{o}}\right\} . \tag{A.4}
\end{equation*}
$$

Finally, from equation (7), we have the marginal cost per journal ( $\mathrm{MC}_{\mathrm{J}}$ ) expressed in terms of the marginal cost per article (MC): $\mathrm{MC}_{\mathrm{J}}=\mathrm{N}^{\gamma}$. MC. With these substitutions, producer surplus under each bundling alternative $\left(\Pi_{\mathrm{PB}}, \Pi_{\mathrm{PU}}\right.$ and $\left.\Pi_{\mathrm{MB}}\right)$ can be expressed as functions of $\mathrm{P}_{\mathrm{J}}, \mathrm{P}_{\mathrm{A}}, \mathrm{MC}, \gamma$ and the model parameters N and $\lambda$.

## A. Pure Bundling

The limits of integration for the pure bundling strategy are based on the boundaries of the region $R_{J}$, as defined by the $U_{J}=0$ curve in Figure 5. Solving for $U_{J}=0$ requires the quantification of $\mathrm{W}_{\mathrm{J}}$. Integrating $\mathrm{w}(\mathrm{n})$ over all articles $(0 \leq \mathrm{n} \leq \mathrm{N}-1)$ results in

$$
\begin{equation*}
W_{J}=\int_{0}^{N} w(n) \cdot d n+\Delta_{c} \tag{A.5}
\end{equation*}
$$

or

$$
W_{J}=\left\{\begin{array}{ccc}
\frac{k N w_{o}}{2}+\Delta_{c} & \text { if } & k \leq 1  \tag{A.6}\\
\frac{(2 k-1) N w_{o}}{2 k}+\Delta_{c} & \text { if } & k>1
\end{array}\right.
$$

The compensating term $\Delta_{\mathrm{c}}$ is the sum of all triangular areas not integrated under the demand curve $\mathrm{w}(\mathrm{n})$ in Figure 3. There are kN (or N if $\mathrm{k}>1$ ) triangles, each with an area of $\mathrm{w}_{\mathrm{o}} / 2 \mathrm{kN}$. Therefore $\Delta_{\mathrm{c}}$ is independent of N (and k if $\mathrm{k} \leq 1$ ):

$$
\Delta_{c}=\left\{\begin{array}{lll}
\left(\frac{w_{o}}{2}\right) & \text { if } & k \leq 1  \tag{A.7}\\
\left(\frac{1}{k}\right)\left(\frac{w_{o}}{2}\right) & \text { if } & k>1
\end{array}\right.
$$

Substituting these results into $\mathrm{U}_{\mathrm{J}}=\mathrm{W}_{\mathrm{J}}-\mathrm{P}_{\mathrm{J}}=0$ yields

$$
w_{o}= \begin{cases}w_{o 1}=\frac{2 P_{J}}{N k+1} & k \leq 1 ;  \tag{A.8}\\ w_{o 2}=\frac{2 k P_{J}}{2 k N-N+1} & k>1 .\end{cases}
$$

Using $\mathrm{w}_{\mathrm{o} 1}$ and $\mathrm{w}_{\mathrm{o} 2}$ as the limits of integration for Eqn. A.1, we can express producer surplus under the pure bundling scenario as

$$
\prod_{P B}=\int_{k=\frac{1}{N}}^{1} \int_{w_{o}=w_{01} 1}^{1}\left[P_{J}-N^{\gamma} \cdot M C\right] f\left(w_{o}, k\right) \cdot d w_{o} \cdot d k+\int_{k=1}^{\infty} \int_{w_{b}=w_{o 2}}^{1}\left[P_{J}-N^{\gamma} \cdot M C\right] f\left(w_{o}, k\right) \cdot d w_{o} \cdot d k
$$

or

$$
\begin{equation*}
\prod_{P B}=\left[P_{J}-N^{\gamma} \cdot M C\right] \cdot\left\{\int_{k=\frac{1}{N}}^{1} \int_{w_{o}=\text { nol }^{\prime}}^{1} \lambda e^{-\lambda\left(k-\frac{1}{N}\right)} \cdot d w_{o} \cdot d k+\int_{k=1}^{\infty} \int_{w_{0}=w_{o 2}}^{1} \lambda e^{-\lambda\left(k-\frac{1}{N}\right)} \cdot d w_{o} \cdot d k\right\} . \tag{A.9}
\end{equation*}
$$

Note that the absence of a $\mathrm{R}_{\mathrm{A}}$ region in pure bundling means that the second term of Eqn. A. 1 can be dropped. Differentiating $\Pi_{P B}$ with respect to $P_{J}$ and setting it to zero, the optimal bundle subscription price $\mathrm{P}_{\mathrm{J}}$ and the corresponding $\Pi_{\mathrm{PB}}$ can be computed numerically.

## B. Pure Unbundling

Under pure unbundling, there is no region $\mathrm{R}_{\mathrm{J}}$ since no subscription bundles are sold. Instead, the region of integration is $R_{A}$, the area to the right of the line $w_{o}=P_{A}$ :

$$
\begin{equation*}
\prod_{P U}=\int_{k=\frac{1}{N}}^{\infty} \int_{w_{o E}}^{1}\left[P_{A}-M C\right] \cdot n^{*} \cdot f\left(w_{o}, k\right) \cdot d w_{o} \cdot d k \tag{A.11}
\end{equation*}
$$

However, $\mathrm{n}^{*}(\mathrm{wo}, \mathrm{k})$ has a discontinuity at $\mathrm{n}^{*}=\mathrm{N}$. This mandates the integration to be carried out in two parts. We can locate the region where $\mathrm{n}^{*}=\mathrm{N}$, which is northeast of the $\mathrm{n}^{*}=100$ curve in Figure 6, by solving $\mathrm{w}(\mathrm{N}-1)=\mathrm{P}_{\mathrm{A}}$. This yields $\mathrm{k}=\mathrm{w}_{\mathrm{o}} /\left(\mathrm{w}_{\mathrm{o}}-\mathrm{P}_{\mathrm{A}}\right)$. Therefore, producer surplus under pure unbundling can be expressed as:

$$
\begin{equation*}
\prod_{P U}=\left[P_{A}-M C\right]\left\{\int_{w_{o}=P_{A}}^{1} \int_{k=\frac{1}{N}}^{\frac{w_{o}}{w_{-}-P_{A}}}\left(\frac{k N\left(w_{o}-P_{A}\right)}{w_{o}}\right) \lambda e^{-\lambda\left(k-\frac{1}{N}\right)} d w_{o} d k+\int_{w_{o}=P_{A}}^{1} \int_{k=\frac{w_{o}}{w_{\sigma} P_{A}}}^{\infty} N \lambda e^{-\lambda\left(k-\frac{1}{N}\right)} d w_{o} d k\right\} \tag{A.12}
\end{equation*}
$$

Again, $\Pi_{P U}$ can be differentiated with respect to $\mathrm{P}_{\mathrm{A}}$ and set to zero, and the optimal article price $P_{A}$ and the maximum $\Pi_{P U}$ can be solved numerically.

The actual utility gained from the purchase of $n *$ articles, $W_{A}$, can be determined by summing (or integrating) all the $\mathrm{w}(\mathrm{n})$ 's for $0 \leq \mathrm{n} \leq \mathrm{n}$ *:

$$
\begin{equation*}
W_{A}=\int_{0}^{n^{*}} w(n) \cdot d n+\Delta_{c}^{*} \tag{A.13}
\end{equation*}
$$

or

$$
\begin{equation*}
W_{A}=(k N)\left(w_{o}-P_{A}\right)-\left(\frac{k N}{2 w o}\right)\left(w_{o}-P_{A}\right)^{2}+\Delta_{c}^{*} . \tag{A.14}
\end{equation*}
$$

In this case, $\Delta_{c}^{*}$ is not the entire area $\Delta_{c}$, only a fraction proportional to $\mathrm{n}^{*}$ :

$$
\Delta_{c}=\left\{\begin{array}{lll}
\left(\frac{n^{*}}{N}\right)\left(\frac{w_{o}}{2}\right) & \text { if } & k \leq 1  \tag{A.15}\\
\left(\frac{n^{*}}{N}\right)\left(\frac{1}{k}\right)\left(\frac{w_{o}}{2}\right) & \text { if } & k>1
\end{array}\right.
$$

which can be reduced to:

$$
\Delta_{c}=\left\{\begin{array}{lll}
k\left(\frac{w_{o}-P_{A}}{2}\right) & \text { if } & k \leq 1  \tag{A.16}\\
\left(\frac{w_{o}-P_{A}}{2}\right) & \text { if } & k>1
\end{array}\right.
$$

From here, the net benefit derived from purchasing $\mathrm{n}^{*}$ individual articles under pure unbundling can be calculated as $\mathrm{U}_{\mathrm{A}}=\mathrm{W}_{\mathrm{A}}-\mathrm{n}^{*} \cdot \mathrm{P}_{\mathrm{A}}$.

## C. Mixed Bundling

The consumer choice regions under mixed bundling may take on one of two slightly different shapes and boundaries depending on the $\mathrm{P}_{\mathrm{J}} / \mathrm{P}_{\mathrm{A}}$ ratio. Figure A. 1 illustrates the two alternate scenarios. In each case, we need to solve $\left\{\mathrm{U}_{\mathrm{J}}=0, \mathrm{U}_{\mathrm{A}}=0, \mathrm{U}_{\mathrm{J}}\right.$ $=\mathrm{U}_{\mathrm{A}}$ \} to establish the boundaries between the different regions. The solution to $\mathrm{U}_{\mathrm{J}}=0$ is the same as that in the pure bundling scenario. The solution to $U_{A}=0$ is simply $w_{o}=P_{A}$. Solving for $\mathrm{U}_{\mathrm{J}}=\mathrm{U}_{\mathrm{A}}$ yields (for $\mathrm{k} \leq 1$ and $\mathrm{w}_{\mathrm{o}} \geq \mathrm{P}_{\mathrm{A}}$ ):

$$
\begin{equation*}
\tilde{w}_{o}=\frac{k N P_{A}^{2}}{(2 k N+1) P_{A}-2 P_{J}}, \tag{A.17}
\end{equation*}
$$

or equivalently,

$$
\begin{equation*}
\tilde{k}=\frac{w_{o}\left(P_{A}-2 P_{J}\right)}{N\left(P_{A}{ }^{2}-2 w_{o} P_{A}\right)} . \tag{A.18}
\end{equation*}
$$


(a) $1<\mathrm{PJ} / \mathrm{PA}_{\mathrm{A}}<\mathrm{N} / 2$

(b) $\mathrm{N} / 2<\mathrm{PJ}^{2} / \mathrm{P}_{\mathrm{A}}<\mathrm{N}$

Figure A. 1
Consumer choice in mixed bundling scenario

For $\mathrm{P}_{\mathrm{J}} / \mathrm{P}_{\mathrm{A}} \leq \mathrm{N} / 2$, as illustrated by Fig. A.1(a), we can express producer surplus as $\Pi_{\mathrm{MB}(\mathrm{S})}$ :

$$
\begin{equation*}
\Pi_{\mathrm{MB}(\mathrm{~S})}=\Pi_{\mathrm{MB}(\mathrm{~S}), \mathrm{A}}+\Pi_{\mathrm{MB}(\mathrm{~S}), \mathrm{J}} \tag{A.19}
\end{equation*}
$$

where

$$
\begin{equation*}
\prod_{M B(S), A}=\left[P_{A}-M C\right] \cdot\left\{\int_{w_{o}=P_{A}}^{1} \int_{k=\frac{1}{N}}^{k} n^{*} \cdot f\left(w_{o}, k\right) \cdot d w_{o} \cdot d k\right\} \tag{A.20}
\end{equation*}
$$

and

$$
\begin{equation*}
\prod_{M B(S), J}=\left[P_{J}-N^{\gamma} M C\right]\left\{\int_{w_{o}=P_{A}}^{1} \int_{k=\tilde{k}}^{\infty} f\left(w_{o}, k\right) d w_{o} d k+\int_{k=\frac{2}{N P_{J}}}^{1} \int_{w_{o}=w_{o l}}^{P_{A}} f\left(w_{o}, k\right) d w_{o} d k+\int_{k=1}^{\infty} \int_{w_{o}=w_{0} 2}^{P_{A}} f\left(w_{a} k\right) d w_{d} d k\right\} \tag{A.21}
\end{equation*}
$$

For $\mathrm{N} / 2 \leq \mathrm{P}_{\mathrm{J}} / \mathrm{P}_{\mathrm{A}} \leq \mathrm{N}$, we have, instead, $\Pi_{\mathrm{MB}(\mathrm{B})}$ :

$$
\begin{equation*}
\Pi_{\mathrm{MB}(\mathrm{~B})}=\Pi_{\mathrm{MB}(\mathrm{~B}), \mathrm{A}}+\Pi_{\mathrm{MB}(\mathrm{~B}), \mathrm{J}} \tag{A.22}
\end{equation*}
$$

where

$$
\prod_{M B(B), A}=\left[P_{A}-M C\right] \cdot\left\{\int_{\left\{w_{0}=P_{A}\right.}^{\left.\tilde{w}_{k}\right|_{k=1}} \int_{k=\frac{1}{N}}^{1} n^{*} f\left(w_{a}, k\right) d w_{o} d k+\int_{w_{0}=\tilde{w}_{0} l_{k=1}}^{1} \int_{k=\frac{1}{N}}^{\tilde{k}} n^{*} f\left(w_{o}, k\right) d w_{d} d k+\int_{w_{o}=P_{A}}^{w_{0} 2} \int_{k=1}^{\frac{N P_{A}}{2\left[N P_{A}-P_{1}\right]}} n^{\pi^{*}} f\left(w_{o}, k\right) d w_{o} d k\right\} \text { (A.23) }
$$

and

$$
\begin{equation*}
\prod_{M B(B), J}=\left[P_{J}-N^{\gamma} M C\right]\left\{\int_{\left.w_{o}=\left.\tilde{w}_{o}\right|_{k=1} ^{1} \int_{k=\tilde{k}}^{1} f\left(w_{o}, k\right) d w_{o} d k+\int_{k=1}^{\infty} \int_{w_{o}=w_{b} 2}^{1} f\left(w_{o}, k\right) d w_{o} d k\right\} . . . . . . . .}\right. \tag{A.24}
\end{equation*}
$$

Since we do not know, a priori, the ratio $\mathrm{P}_{\mathrm{J}} / \mathrm{P}_{\mathrm{A}}$, we need to compute both $\Pi_{\mathrm{MB}(\mathrm{S})}$ and $\Pi_{\mathrm{MB}(\mathrm{B})}$, obtain the two sets of optimal prices, and by inspection of the ratios determine which set of the results is valid.

## Appendix B. Sample list of web-hosting service providers

| Service Provider | URL | \$/MB download |
| :---: | :---: | :---: |
| AT\&T Easy World Wide Web | http://www.att.com/ | \$0.50 |
| Cowboy.Net | http://cowboy.net/commercial_prices.html | \$0.05 |
| Citizens Internet Service | http://www.swva.net/citizen/services/webprice.html | \$1.00 |
| DC-AdNet | http://www.dc-adnet.com/prices.htm | \$1.00 |
| Internet Industries Web Hosting | http://www.industries.net/webhosting.html | \$0.05 |
| Internet Video Services' netvideo | http://www.netvideo.com/netvideo/price.html | \$0.02-\$0.08 |
| Multiboard Communications | http://www.multiboard.com/services.html | \$0.07-\$0.10 |
| PreciseNet Web Site Hosting | http://www.precisenet.com/host.htm | \$0.20 |
| Pro-NetMedia Creations, Inc. | http://www.pcinc.com/pricing.htm | \$0.25 |
| Serview Premium Webhosting | http://serview.com/pricing.html | \$0.10 |
| Sustance | http://www.he.net/~sustance/prices.html | \$0.039-\$0.10 |

## References

Adams, W.J. and J.L. Yellen, 1976, Commodity bundling and the burden of monopoly, Quarterly Journal of Economics 90, 475-498.
Armstrong, M., 1997, Price discrimination by a many-product firm, mimeo.
Bakos, Y. and E. Brynjolfsson, 1997, Bundling information goods: pricing, profits and efficiency, in: Proceedings of the Conference on Internet Publishing and Beyond: Economics of Digital Information and Intellectual Property, Cambridge MA.
Besen, S.M. and S.N. Kirby, 1989, Private copying, appropriability, and optimal copying royalties, Journal of Law and Economics 32, no. 2, pt. 1, 255-280.
Burnstein, M.L., 1960, The economics of tie-in sales, Review of Economics and Statistics 42, 68-73.
Byrd, G.D., 1990, An economic "commons" tragedy for research libraries: scholarly journal publishing and pricing trends, College \& Research Libraries 51, 184-195.
Carbajo, J., D. de Meza and D.J. Seidmann, 1990, A strategic motivation for commodity bundling, Journal of Industrial Economics 38, 283-298.
Chae, S., 1992, Bundling subscription TV channels: a case of natural bundling, International Journal of Industrial Organization 10, 213-230.
Dyl, E.A., 1983, A note on price discrimination by academic journals, Library Quarterly 53, no. 2, 161-168.
Fishburn, P.C., A.M. Odlyzko and R.C. Siders, 1997, Fixed fee versus unit pricing for information goods: competition, equilibria, and price wars, in: Proceedings of the Conference on Internet Publishing and Beyond: Economics of Digital Information and Intellectual Property, Cambridge MA.
Hanson, W. and R.K. Martin, 1990, Optimal bundle pricing, Management Science 36, no.2, 155-174.
Joyce, P. and T.E. Merz, 1985, Price discrimination in academic journals, Library Quarterly 55, no. 3, 273-283.
King, D.W. and J.M. Griffiths, 1995, Economic issues concerning electronic publishing and distribution of scholarly articles, Library Trends 43, no. 4, 713-740.
King, D.W., D.D. McDonald, and N.K. Roderer, 1981, Scientific journals in the United States: their production, use, and economics. (Hutchinson Ross Publishing, Stroudsburg PA).
Laffont, J.-J., E. Maskin and J.-C. Rochet, 1985, Optimal nonlinear pricing with twodimensional characteristics, in: Groves, Radner and Reiter, eds. Information, Incentives and Economic Mechanisms.
Lewis, D.W., 1989, Economics of the scholarly journal, College \& Research Libraries 50, 674-688.
Liebowitz, S.J., 1985, Copying and indirect appropriability: photocopying of journals, Journal of Political Economy 93, no. 5, 945-957.
MacKie-Mason, J.K. and J. Riveros, 1997, Economics and electronic access to scholarly information, in: Internet Publishing and Beyond: Economics of Digital Information and Intellectual Property, forthcoming.
MacKie-Mason, J.K. and K. White, 1996, Evaluating and selecting digital payment mechanisms, in: Proceedings of the 1996 Telecommunications Policy Research Conference.
Manasse, M.S., 1995, The Millicent protocols for electronic commerce, in: Proceedings of the First USENIX Workshop on Electronic Commerce.
McAfee, R.P., J. McMillan, and M.D. Whinston, 1989, Multiproduct monopoly, commodity bundling, and correlation of values, Quarterly Journal of Economics 104, 371383.

Metz, P. and P.M. Gherman, 1991, Serial pricing and the role of the electronic journal, College \& Research Libraries 52, 315-327.

Mitchell, B.M., and I. Vogelsang, 1991, Telecommunications pricing: theory and practice (University Press, Cambridge).
Okerson A.S., and J.J. O'Donnell, eds. 1995, Scholarly journals at the crossroads: a subversive proposal for electronic publishing (Association of Research Libraries, Washington, DC).
Ordover, J.A. and R.D. Willig, 1978, On the optimal provision of journals qua sometimes shared goods, American Economic Review 68, 324-338.
Schmalensee, R., 1984, Gaussian demand and commodity bundling, Journal of Business 57, no. 1, pt. 2, S211-230.
Sirbu, M.A. and J.D. Tygar, 1995, NetBill: an Internet commerce system optimized for network delivered services, in: Proceedings of IEEE CompCon Conference.
Sirbu, M.A., 1997, Credits and debits on the Internet, IEEE Spectrum, Feb. 1997, 23-29. Spigai, F., 1991, Information pricing, in: Annual Review of Information Science and Technology 26, 39-73.
Stigler, G.J., 1963, United States v. Loew's Inc.: a note on block booking, Supreme Court Review, 152-157.
Stoller, M.A., R. Christopherson and M. Miranda, 1996, The economics of professional journal pricing, College \& Research Libraries 57, 9-21.
Varian, H.R., 1995, Pricing information goods, in: Proceedings of Scholarship in the New Information Environment Symposium, Harvard Law School.
Whinston, M.D., 1990, Tying, foreclosure, and exclusion, American Economic Review 80, 837-859.
Willig, R.D., 1978, Pareto-superior nonlinear outlay schedules, Bell Journal of Economics 9, 56-69.
Wilson R., 1993, Nonlinear pricing (University Press, Oxford).
Zahray, W.P. and M. Sirbu, 1990, The provision of scholarly journals by libraries via electronic technologies: an economic analysis, Information Economics and Policy 4, 127154.


[^0]:    * Corresponding author: John Chuang. 129 Baker Hall, Carnegie Mellon Univesity, Pittsburgh Pennsylvania 15213, USA. (tel) 1-412-268-5605; (fax) 1-412-268-3757; (email) chuang+@cmu.edu; (url) http://www.andrew.cmu.edu/user/chuang.
    ${ }^{1} 17$ U.S.C. § 107 (1988 \& Supp. V 1993).
    ${ }^{2} 17$ U.S.C. § 108(g)(2) (1988). Specifically, the CONTU Guidelines (reprinted in 1976 U.S.C.C.A.N. $5810,5813-14)$ set forth a copying limit of five copies per year of articles from the most recent five years of any journal article.

[^1]:    ${ }^{3}$ See Liebowitz (1985) and Besen and Kirby (1989) for detailed treatment of journal photocopying and indirect appropriability; Joyce and Merz (1985) study the extent of price discrimination by journals across various academic disciplines.
    ${ }^{4} \mathrm{Dyl}$ (1983) muses upon the applicability and antitrust implications of the Robinson-Patman Act to price discrimination by academic journals.
    ${ }^{5}$ Interested readers can consult Lewis (1989), Byrd (1990), Metz and Gherman (1991), Spigai (1991) and Stoller, Christopherson and Miranda (1996) for works on the economics of scholarly publishing and serials pricing from the library and information sciences communities' perspective.

[^2]:    ${ }^{6}$ Carbajo, de Meza and Seidmann (1990) and Whinston (1990) provide further treatment of this topic.

[^3]:    ${ }^{7}$ Armstrong (1997) shows that an approximate solution for the optimal tariff problem is a cost-based twopart tariff, i.e. a fixed up-front membership fee plus a per-article charge set equal to the marginal cost. However, this approximation reasonably converges only for N in the 'several thousands' range, and the absence of a price cap may make it unacceptable to consumers who are used to the traditional subscription model.
    ${ }^{8}$ Technically, an $(\mathrm{n}+1)$-part tariff can be made to be Pareto-superior to an n -part tariff. Indeed, Laffont, Maskin and Rochet (1985) derived the optimal nonlinear tariff for consumers with two dimensional characteristics, which has a gradually declining marginal price schedule. Using the same argument in the bundling context, any mixed bundling strategy with more than two prices (up to $2^{\mathrm{N}}$ ) will necessarily perform better than a mixed bundling strategy with two prices, and thus pure bundling and pure unbundling strategies as well. Again, the extent to which a publisher chooses to offer multiple prices is clearly dependent upon its multi-variate optimization capabilities, and more importantly, consumer acceptance/rejection of a complex pricing structure.

[^4]:    ${ }^{9}$ We assume here and in subsequent sections that there is no economies of scope in demand, i.e. the marginal benefits of individual articles are additive but not superadditive.

[^5]:    ${ }^{10}$ circulation data from Ulrich's International Periodicals Directory, 34 ed. R.R. Bowker Publishing, 1996.

[^6]:    ${ }^{11}$ See MacKie-Mason and White (1996) and Sirbu (1997) for surveys of digital payment mechanisms.

[^7]:    ${ }^{12}$ A typical credit card operation has $\kappa_{\mathrm{f}}$ and $\kappa_{\mathrm{v}}$ set at $\$ 0.30$ and $1.66 \%$ and is not suited for small value transactions because of this high $\kappa_{\mathrm{f}}$. NetBill (http://www.netbill.com), an experimental electronic micropayment system developed at Carnegie Mellon University, has $\kappa_{f}=\$ 0.02$ and $\kappa_{v}=5 \%$, enabling it to support transactions down to $5-10$ cents. This latter set of cost coefficients is used for this analysis. See Sirbu and Tygar (1995) for a description of the NetBill electronic micropayment system.
    ${ }^{13}$ Price schedules for incremental data downloads obtained from a website survey of web hosting service providers, January 1997. See Appendix B.

[^8]:    ${ }^{14}$ Okerson and O'Donnell (1995) edit an interesting forum discussion, which took place entirely on the Internet, on the future of scholarly journals; drawing experience from various electronic journal endeavors such as Psycoloquy, Chicago Journal of Theoretical Computer Science , and the electronic pre-print archive for high-energy physics at Los Alamos National Labs.

