

Bundling Information Goods of Decreasing Value

Xianjun Geng

Department of Management Science, University of Washington Business School, Seattle, Washington 98195,
gengxj@u.washington.edu

Maxwell B. Stinchcombe

Department of Economics, The University of Texas at Austin, Austin, Texas 78712, maxwell@eco.utexas.edu,

Andrew B. Whinston

Department of Management Science and Information Systems, The University of Texas at Austin,
Austin, Texas 78712, abw@uts.cc.utexas.edu

Consumers' average value for information goods, websites, weather forecasts, music, and news declines with the number consumed. This paper provides simple guidelines to optimal bundling marketing strategies in this case. If consumers' values do not decrease too quickly, we show that bundling is approximately optimal. If consumers' values to subsequent goods decrease quickly, we show by example that one should expect bundling to be suboptimal.

Key words: bundling; electronic commerce; price discrimination; digital products

History: Accepted by Rajiv D. Banker, information systems; received November 10, 2003. This paper was with the authors 1 week for 1 revision.

1. Introduction

It is the norm rather than the exception that consumers' average value for consuming a stream of information goods—goods with zero or very low marginal cost—declines with the number consumed. Travelers value successive future weather forecasts less because the accuracy declines quickly with forecast length. Internet surfers eventually get bored with streaming music or other online entertainment. More generally, consistent with a positive risk-free interest rate, people discount the value of future consumption relative to present consumption.

Pure bundling is the simplest pricing strategy—sell all of the goods in a single bundle at a single price. In this paper we provide a simple framework for examining the optimality of pure bundling of information goods of decreasing value. We assume that consumers' value for a bundle of goods is additively separable in their values of the individual goods, but we do *not* necessarily assume that the values of information goods are independent. Our framework allows for correlation of values across goods and a wide class of declining average values.

The most extreme case of positive and declining values for more than one good has positive values for two goods and no value for subsequent goods. It is clear that pure bundling is usually suboptimal in this case (Schmalensee 1984, McAfee et al. 1989). We give a much less extreme example in §5, where we

explain that all the goods have positive values and pure bundling causes a loss of half of the possible profits. We note in passing that the example satisfies all of the conditions that Bakos and Brynjolfsson (1999) claim lead to the optimality of bundling.

The simple framework allows us to show how a slow decline in value can ensure the optimality of pure bundling. Moreover, under slow decline in value, pure bundling is optimal even when there are strong positive or negative correlations of values across goods. We also give correct conditions for the results in Bakos and Brynjolfsson (1999) to be valid.

In what follows, we review the bundling literature in §2; we give our framework in §3; in §4 we discuss our simple sufficient condition for the optimality of pure bundling; in §5 we illustrate how violations of our simple condition can lead to the nonoptimality of pure bundling; in §6 we show that our simple condition implies that pure bundling is optimal in a broad range of cases of declining values; and §7 concludes.

2. Literature Review

Economists have long known that bundling can be an effective way for a multiproduct monopolist to increase profits when it has limited information about individual consumer preferences (Adams and Yellen 1976; Schmalensee 1984; McAfee et al. 1989; Armstrong 1996, 1999; Norman and Fang 2003). McAfee

et al. (1989) consider the case of two products with unit demand from each consumer and show that mixed bundling—the bundling strategy where both the bundle and subbundles of products are offered—is more profitable than selling the two products separately in a broad range of instances, including in the case of independent valuations. However, the complexity involved in analyzing bundling has impeded researchers from obtaining general results on the profitability of monopolistic bundling when there are more than two products.

More recently, Bakos and Brynjolfsson (1999) initiated research on the bundling of information goods. They consider the case of a monopolist selling a large number of information goods with zero or very low marginal cost, unit demand from each consumer, and random utilities that are additive in the goods consumed. They state their result as, “bundling very large numbers of unrelated information goods can be surprisingly profitable” (p. 1613). Their arguments are incorrect, and when corrected, the pricing recommendations that flow from their result are misleading.

As we show below, when the result in Bakos and Brynjolfsson (1999) is true, it arises from Tchebyshev’s inequality. Further, the use of Tchebyshev’s inequality shows that their result is true in cases where their assumptions are ruled out. Their framework and exposition uses the logic of the weak law of large numbers, a statement about the *average* of large numbers of random variables. Using the recommendations that flow from using the average price of bundled good leads either to charging \$0 and making \$0, or charging more than the consumers’ total wealth. As we argue here, the correct unit of analysis is the price of the bundle.

3. The Framework

Consider a monopolistic seller offering information goods $1, 2, \dots, m, \dots$ to a population of consumers, Ω . The size of the whole consumer population is normalized to 1. Marginal production cost for an information good is 0. We assume that for any consumer $\omega \in \Omega$, her value of a goods bundle is additively separable in her values of the individual goods,

$$Y(\omega) = \sum_{m=1}^{\infty} c_m X_m(\omega), \quad (1)$$

where $Y(\omega)$ is the value of the bundle for consumer ω , and $c_m X_m(\omega)$ is the value contributed to the bundle by good m , or simply the value of good m .

This value of good m for consumer ω contains two parts: a nonnegative number $X_m(\omega)$, which captures

the heterogeneity among consumers,¹ and is private information to consumer ω ; a nonnegative number c_m , which reflects the average value among these information goods. We will focus on the case where the sequence $(c_1, c_2, \dots, c_m, \dots)$ eventually decreases to 0.

The monopolist chooses a pricing strategy, that is, a set of bundles of goods and prices for these bundles. Pure bundling sells all of the goods together at a single price. One example of pure bundling is RealOne SuperPass by Real Networks, which bundles digital contents into a single online service.

Our basic result provides a bound on the ratio of the maximal possible profits under any bundling scheme, no matter how complicated, and the profits earned under simple bundling. This bound leads to a number of easy, sufficient conditions on c_m and the joint distribution of the $X_m(\omega)$ for our bound to be arbitrarily tight.

By setting $c_m = 0$ for all large m , the modeling setup in (1) accommodates finite numbers of goods. With $c_m > 0$ for infinite m s, it accommodates an infinite number of goods. We think of the study of “infinite goods” as a convenient method of studying very large finite numbers of goods.

All of the X_m s (and thus Y s) are random variables defined on Ω . We assume that each X_m has finite mean, μ_m , and finite variance, σ_m^2 . We also assume that $Y = \sum_{m=1}^{\infty} c_m X_m$ has finite, positive mean, μ , and finite variance,

$$\sigma^2 < \infty. \quad (2)$$

The maximal per capita profit the seller can receive from any bundling strategy, no matter how complicated, is denoted π^* . We use p to denote the price of the pure bundle, $\pi(p)$, the associated per capita profit, and p^* , the profit-maximizing price for the most profitable pure bundle. $Y, \mu, \sigma^2, p^*, \pi^*$ all depend on the c_m s and the joint distribution of the X_m s.

Our crucial assumptions are additive separability (1), and finite variance (2).

4. A Sufficient Condition for the Optimality of Pure Bundling

The largest possible profit for the monopolist occurs when they “perfectly discriminate,” that is, they receive $Y(\omega)$ from each consumer ω . If the monopolist could perfectly identify each consumer’s maximal willingness to pay for the bundle in this way, then

¹ Note that in practice consumers often face information goods with negative values, which voids the assumption that $X_m(\omega)$ is nonnegative. For example, many Web surfers find most Web pages boring or even irritating in the case of pop-ups. In these cases, the consumers can often ignore or suppress the goods they do not wish to view. We call this user-defined filtering. When user-defined filters exist, we can replace the possibly negative $X_m(\omega)$ with a nonnegative $X'_m(\omega) = \max(0, X_m(\omega))$ for our analysis.

they would have per capita profits of $\mu = E(Y)$. Therefore $\pi^*/\mu \leq 1$.

When σ/μ is very small, consumers' valuations are tightly concentrated around μ . Intuitively, this means that there can be very little loss to the monopolist from adopting the pure bundling strategy and selling at a price $(1 - \varepsilon)\mu$ for some small positive ε . This intuition is borne out in the following immediate implication of Tchebyshev's inequality.

LEMMA 1. *If we set $\varepsilon = (\sigma/\mu)^{2/3}$, then pure bundling at a price $p = (1 - \varepsilon)\mu$ gives profits of at least $\pi = (1 - 2\varepsilon)\mu$. In particular,*

$$1 \geq \frac{\pi^*}{\mu} \geq 1 - 2\varepsilon. \quad (3)$$

For example, if $\sigma/\mu = 0.001$, then pure bundling achieves 98% of the profits of a perfectly discriminating monopolist.

Following Lemma 1, conditions on c_m s and X_m s guaranteeing that σ/μ is vanishingly small will ensure that the profit difference between a perfectly discriminating monopolist and a pure bundler is vanishingly small.² In this case we say that the pure bundler achieves *approximate optimality*. Although we do not know the optimal selling strategy for the seller, which could involve complex mixed bundling, we do know that the profit under pure bundling is very close to the optimal one. We record this observation as Proposition 1.

PROPOSITION 1. *Pure bundling is approximately optimal if σ/μ is almost zero.*

Approximate optimality is the optimality concept used in this paper. Except when specifically noted, in this paper, by "optimality," we always mean "approximate optimality." Similar use of the term "optimality" also presents in both Bakos and Brynjolfsson (1999) and Armstrong (1999). Proposition 1 provides a simple sufficient condition for determining the optimality of pure bundling. Lemma 1 gives guidelines on the approximately optimal bundle price.

In the extreme case that only a few goods are valued positively, σ/μ is determined by those few goods. In this case, and in similar cases of rapidly declining values satisfying all of the conditions given in Bakos and Brynjolfsson (1999), bundling can be a money loser compared to selling the goods individually, that is, compared to unbundling them (see §5).

When the values of subsequent goods decline, but not too rapidly, and the X_m s are of approximately the same size and not too dependent, then σ/μ is small. Section 6 gives a number of plausible examples of slowly decreasing values, each of which ensures the optimality of pure bundling in a large range of cases.

5. When Pure Bundling Is Not Optimal

An extreme case of positive and declining values has positive values for two goods and no value for subsequent goods. Pure bundling can be a money loser in this case (Schmalensee 1984), and mixed bundling (selling both goods separately and selling the bundle for less than the sum of the prices) will typically dominate (McAfee et al. 1989). In the following less extreme example, satisfying Bakos and Brynjolfsson's (1999) conditions for the optimality of bundling, selling the goods individually doubles the profit achievable by pure bundling.

EXAMPLE 1. Let X_1, X_2, \dots be *i.i.d.* For any m , the continuous density function of X_m is given by $f_m(x) = (1/2\varepsilon) \max(1 - (1/\varepsilon)d(x, \{\varepsilon, a\}), 0)$, where a is a constant positive number, ε is a small positive number and $\varepsilon \ll a$, and $d(x, \{\varepsilon, a\}) = \min\{|x - \varepsilon|, |x - a|\}$ is the shortest distance between x and the two point set $\{\varepsilon, a\}$. The densities are concentrated in the neighborhoods of ε and a . (They represent the consumers having a 50 : 50 chance of loving the product or not caring about it.) Let $c_m = 1/2^{m-1}$ so that the consumers' values of goods decrease at an exponential rate with a discount factor of $1/2$. (Take $v_{nk} = X_k/2^{k-1}$ and Bakos and Brynjolfsson's 1999, Assumptions A1, A2, and A3 are satisfied.)

We study the case where $\varepsilon \approx 0$. If the goods are sold separately, the optimal price for the m th good is $1/2^{m-1}$ times a number between $a - \varepsilon$ and a . Thus, the total profit the seller can earn by complete unbundling converges to $(1 + 1/2 + 1/4 + \dots) \cdot a/2 = a$ as ε shrinks to 0. If the seller bundles all goods in a single bundle, the value is $Y = X_1 + (1/2)X_2 + \dots + (1/2^{m-1})X_m + \dots$. To see the density function of Y , first consider the finite case, $Y_n = X_1 + (1/2)X_2 + \dots + (1/2^{n-1})X_n$, and observe that when $\varepsilon \rightarrow 0$, the density of Y_n concentrates equally at the neighborhood of the following 2^n points: $0, a/2^{n-1}, 2(a/2^{n-1}), 3(a/2^{n-1}), \dots, (2^n - 1)(a/2^{n-1})$. As $n \rightarrow \infty$ and $\varepsilon \rightarrow 0$, this converges to the uniform distribution on the interval $[0, 2a]$, and it is optimal to set the price equal to a , yielding a profit of $a/2$. Pure bundling causes a loss of half of the potential profits.

Our example uses a particular bimodal density. Beyond the clustering of consumers' values, the specific details of the density are not particularly important. It is discounting that drives the conclusion.

6. When Pure Bundling Is Optimal

In this section we give sufficient conditions for σ/μ to be small. These show that if consumers' values of goods in a bundle decrease slowly, then pure bundling is optimal in a wide range of cases.

² All proofs are in the appendix. Armstrong (1999) has a similar result to Lemma 1, although he does not discuss the decreasing value case.

A. Uncorrelated Values

We say that the X_m s are *regular* if

- (i) all, but at most finitely many, of the μ_m s belong to a bounded interval $[\mu, \bar{\mu}]$ with $\mu > 0$, and
 - (ii) the σ_m^2 s are bounded above by $\bar{\sigma}^2$ with $\bar{\sigma} > 0$.
- If the X_m s are uncorrelated, then

$$\frac{\sigma}{\mu} < \sqrt{\frac{\sum_{m=1}^{\infty} c_m^2}{(\sum_{m=1}^{\infty} c_m)^2}} \cdot \frac{\bar{\sigma}}{\underline{\mu}}.$$

Therefore, if the X_m s have the same mean and variance, then

$$\frac{\sum_{m=1}^{\infty} c_m^2}{(\sum_{m=1}^{\infty} c_m)^2}$$

being small is sufficient for the optimality of pure bundling. It turns out that the same mean and variance are not needed for this intuition to be correct, as shown below.

PROPOSITION 2. *Pure bundling is optimal for uncorrelated, regular X_m s if*

$$\frac{\sum_{m=1}^{\infty} c_m^2}{(\sum_{m=1}^{\infty} c_m)^2} \tag{4}$$

is sufficiently small.

Exponential decline of the consumers' valuations corresponds to $c_m = \delta^m$ for some discount factor δ satisfying $0 < \delta < 1$. A discount factor close to 1 implies that the values of the goods decrease slowly. This is especially plausible when the time between consumption of the goods is very short. A direct proof of the following uses Proposition 2 and L'Hôpital's rule to show that

$$\frac{\sum_{m=1}^{\infty} (\delta^m)^2}{(\sum_{m=1}^{\infty} \delta^m)^2}$$

goes to 0 as δ goes to 1.

COROLLARY 1. *For uncorrelated, regular X_m s and exponential decline of the c_m s, σ/μ is small when δ is close to 1.*

The proportional weight of good m in the bundle Y is $c_m / \sum_{j=1}^{\infty} c_j$. In Corollary 1, for each m , the proportional weight of m is $\delta^m / \sum_{j=1}^{\infty} \delta^j$, and this is close to 0 when δ is close to 1. This is a special case of Corollary 2.

COROLLARY 2. *For uncorrelated, regular X_m s, if the maximal ratio of $c_m / \sum_{j=1}^{\infty} c_j$ is small, then σ/μ is small.*

The special case of this that Bakos and Brynjolfsson (1999) study has $c_m = 1$ for $m = 1, \dots, N$, $c_m = 0$ for $m > N$, where N is the number of goods in the bundle. Here the maximal ratio $c_m / \sum_{j=1}^{\infty} c_j = 1/N$, and as N becomes large, this ratio goes to 0. When pure bundling is optimal in the cases they consider, the profit-maximizing price for a bundle is $(1 - \varepsilon)\mu$ for

some ε close to 0, but $\mu \geq N\mu$. This means that the price charged for the bundle goes to infinity. If one adds yet another assumption to their model, that $(1/N) \sum_{m=1}^N \mu_m$ converges—say to r —then one can conclude that the average price per good in the bundle should be $(1 - \varepsilon)r$ for some small ε . This has the advantage of being of a sensible size, though its implications for bundle pricing are still implausible.

B. Correlated Values

We now relax the assumption that values of all goods are uncorrelated, although we keep regularity. Let Σ be the variance-covariance matrix of the X_m 's, \vec{c} , the vector of c_m s, and $\vec{\mu}$, the vector of μ_m s. Let $\sigma_{i,j}$ denote the i, j th entry of Σ , that is, covariance of X_i and X_j . Suppose that the $\sigma_{i,j}$ s are bounded above by $\bar{\sigma}$. Then $\sigma/\mu < (\sqrt{\vec{c}'\Sigma\vec{c}}/(\vec{c}'\vec{\mu})) \cdot \bar{\sigma}/\underline{\mu}$. When the ratio $\sqrt{\vec{c}'\Sigma\vec{c}}/(\vec{c}'\vec{\mu})$ is small, we know that pure bundling is optimal. The original analyses of bundling (Adams and Yellen 1976, Schmalensee 1984) had negative off-diagonal entries, that is, negative correlation between the values of the goods. Making the correlation negative can have the effect of lowering σ while leaving μ unchanged.

We say that Σ is *finitely correlated* if there exists an integer K such that for all i , $\sigma_{i,j} \neq 0$ for at most K different j s. We say that Σ is *mixing* if a constant C and a β , $0 < \beta < 1$ exists, such that $\sigma_{i,j} \leq C\beta^{|i-j|}$. Both of these capture the idea that there is a "distance" between goods, and the more distant two goods are, the less related their values are.

PROPOSITION 3. *If X_m s are regular and have a fixed covariance structure that is either finitely correlated or mixing, then if the maximal ratio of $c_m / \sum_{j=1}^{\infty} c_j$ is small, then σ/μ is small.*

The analysis here points out that even positive correlation between the values of goods is consistent with pure bundling, provided the correlation does not dominate in Σ . This is the case, for example, in weather forecasts where serial correlation exists but diminishes quickly when the dates are far away from each other. Another example is online stock quotes—in the short term, quotes are very highly correlated, the long-term quotes are not.

C. On the Assumptions in Bakos and Brynjolfsson (1999)

As shown above, the results in Bakos and Brynjolfsson (1999) on the optimality of bundling are not true without some additional assumptions. Their special case

(1) has independent,³ not merely uncorrelated, X_m s,

(2) requires densities of the X_m s, and

(3) requires that the densities have uniformly bounded support.

With all of these in place, they study

(4) the case of $c_m = 1$ for $m = 1, \dots, N$, $c_m = 0$ for $m > N$.

In this setting, they look for conditions under which bundling becomes optimal as N grows, and for pricing guidance on the average price of a good in the bundle.

As seen above, the optimality of bundling has nothing to do with the existence of, or conditions on, the densities of consumers' valuations, and very little to do with independence. However, reading Bakos and Brynjolfsson's analysis (1999, p. 1627) shows that they use the condition that $(1/N) \sum_{m=1}^N \mu_m$ converges to a strictly positive number throughout their arguments. Note that this is a strictly stronger condition than regularity.

Regularity and small $c_m / \sum_{j=1}^{\infty} c_j$ imply that bundling is optimal, thus, their bundling optimality result follows, given the additional assumption of regularity. However, their focus on the average price of a good in a bundle whose value is increasing without bound gives no useful guidance to pricing. If $(1/N) \sum_{m=1}^N \mu_m$ converges to a strictly positive number—say r —the optimal price of the bundle is $N \cdot (1 - \varepsilon)r$, which goes to infinity and surpasses any consumer's total wealth when N goes to infinity. The goods are not to be sold separately, so the average price is not helpful. In some cases, it is worse than unhelpful: In the case of declining values, $(1/N) \sum_{m=1}^N \mu_m$ will often converge to 0. For example, this occurs in all of the exponentially declining values cases. Believing that the average price of a good in the bundle should be 0 leads to a "give it all away" business model and guarantees profits of 0.

7. Summary and Conclusions

This paper analyzes a simple condition for the approximate optimality of bundling when consumer valuations of goods are additively separable. The condition is that when sampling from the population of consumers, the ratio of the standard deviation and mean of the value of the bundle is small. By focusing on the total price of the bundle rather than the average price of the goods included in the bundle, the requisite techniques become quite simple and many pitfalls are avoided.

³ Bakos and Brynjolfsson's (1999) Proposition 1A makes claims about the optimality of bundling with correlation. There is a class of counterexamples to their claim that does not work through discounting (as in §5).

Note that our results also apply to cases where goods may have, on average, negative values as long as consumers can easily block parts of the bundle they dislike. This means that the widespread attitudes that "most of the Web is boring," "most news is worthless," or "most TV is boring" are entirely consistent with being an avid Web surfer, newspaper reader, or TV junkie. In this context, our analysis provides an explanation of having an ISP, a newspaper subscription, and a subscription to cable TV. In each of these cases, the supplier bundles their services.

Acknowledgments

The authors thank participants at the 14th annual Workshop on Information Systems and Economics (WISE) at Barcelona, Spain, in December 2002 and seminar attendants at the University of Washington in December 2003 for valuable comments. They also thank the department editor, associate editor, and referees for their helpful comments.

Appendix

PROOF OF LEMMA 1. The upper bound of π^* is straightforward from the definition of μ . To obtain the lower bound, we pick a small positive number ε and choose $p = (1 - \varepsilon)\mu$.

$$\begin{aligned} \pi^* &\geq (1 - \varepsilon)\mu \cdot \text{Prob}(Y > (1 - \varepsilon)\mu) \\ &\geq (1 - \varepsilon)\mu \cdot \text{Prob}(|Y - \mu| < \varepsilon\mu) \\ &\geq (1 - \varepsilon)\mu \cdot [1 - (\sigma/\varepsilon\mu)^2] \\ &\geq \mu \left[1 - \varepsilon - \frac{1}{\varepsilon^2}(\sigma/\mu)^2 \right]. \end{aligned}$$

(By the Tchebyshev Inequality.) Set $\varepsilon = (\sigma/\mu)^{2/3}$, and we have $\pi^* \geq \mu[1 - 2(\sigma/\mu)^{2/3}]$. □

PROOF OF PROPOSITION 2. When $n = \infty$, the mean of the bundle is $\mu = \sum_{m=1}^{\infty} c_m \mu_m$, and variance is $\sigma^2 = \sum_{m=1}^{\infty} c_m^2 \sigma_m^2$. We have

$$\begin{aligned} \frac{\sigma^2}{\mu^2} &= \frac{\sum_{m=1}^{\infty} \delta^{2(m-1)} \sigma_m^2}{\sum_{m=1}^{\infty} \delta^{m-1} \mu_m} < \frac{\sum_{m=1}^{\infty} c_m^2 \bar{\sigma}^2}{\left(\sum_{m=1}^{\infty} c_m \underline{\mu}\right)^2} \\ &= \left(\frac{\bar{\sigma}}{\underline{\mu}}\right)^2 \cdot \frac{\sum_{m=1}^{\infty} c_m^2}{\left(\sum_{m=1}^{\infty} c_m\right)^2}. \quad \square \end{aligned}$$

PROOF OF COROLLARY 1. Now we have

$$\begin{aligned} \frac{\sigma^2}{\mu^2} &< \left(\frac{\bar{\sigma}}{\underline{\mu}}\right)^2 \cdot \frac{\sum_{m=1}^{\infty} c_m^2}{\left(\sum_{m=1}^{\infty} c_m\right)^2} \\ &= \left(\frac{\bar{\sigma}}{\underline{\mu}}\right)^2 \cdot \frac{1 - \delta}{1 + \delta} \rightarrow 0 \quad \text{when } \delta \rightarrow 1. \quad \square \end{aligned}$$

PROOF OF COROLLARY 2. Define $\alpha_m = c_m / \sum_{j=1}^{\infty} c_j > 0$ so that $\sum_{m=1}^{\infty} \alpha_m = 1$. Then

$$\frac{\sum_{m=1}^{\infty} c_m^2}{\left(\sum_{j=1}^{\infty} c_j\right)^2} = \sum_{m=1}^{\infty} \left(\frac{c_m}{\sum_{j=1}^{\infty} c_j}\right)^2 = \sum_{m=1}^{\infty} \alpha_m^2.$$

If $\bar{\alpha} = \sup_m(\alpha_m) \rightarrow 0$ and $\sum_{m=1}^{\infty} \alpha_m = 1$, we know that $\sum_{m=1}^{\infty} \alpha_m^2 \leq \sum_{m=1}^{\infty} \bar{\alpha} \alpha_m = \bar{\alpha} \rightarrow 0$. □

PROOF OF PROPOSITION 3. First, consider the case where the variance-covariance matrix is finitely correlated. Note that

$$\frac{\sigma^2}{\mu^2} < \left(\frac{\bar{\sigma}}{\underline{\mu}}\right)^2 \cdot \frac{(K+1) \sum_{m=1}^{\infty} c_m^2}{\left(\sum_{m=1}^{\infty} c_m\right)^2},$$

where $\bar{\sigma}^2$ is the upper bound of all σ_m^2 s. From Corollary 2 we know that $\sum_{m=1}^{\infty} c_m^2 / (\sum_{j=1}^{\infty} c_j)^2$ is small if the maximal ratio of $c_m / \sum_{j=1}^{\infty} c_j$ is small. Then σ^2 / μ^2 is also small because K is a fixed number.

Second, consider the case where the variance-covariance matrix is mixing. Note that

$$\begin{aligned} \frac{\sigma^2}{\mu^2} &\leq \frac{\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} c_i c_j \beta^{|i-j|}}{\left(\sum_{m=1}^{\infty} c_m \mu\right)^2} \\ &= \frac{C}{\underline{\mu}^2} \cdot \frac{\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} c_i \beta^{|i-j|} c_j}{\left(\sum_{m=1}^{\infty} c_m\right)^2} \\ &= \frac{C}{\underline{\mu}^2} \cdot \frac{c_1 \cdot \sum_{j=1}^{\infty} \beta^{|1-j|} c_j + c_2 \cdot \sum_{j=1}^{\infty} \beta^{|2-j|} c_j + \dots}{c_1 \cdot \sum_{m=1}^{\infty} c_m + c_2 \cdot \sum_{m=1}^{\infty} c_m + \dots}. \end{aligned}$$

Therefore we only need to show that $\sum_{j=1}^{\infty} \beta^{|i-j|} c_j / \sum_{m=1}^{\infty} c_m$ is small for any i . Notice that if

$$\bar{\alpha} = \sup_j \left(\frac{c_j}{\sum_{m=1}^{\infty} c_m} \right) \rightarrow 0,$$

we have

$$\begin{aligned} \frac{\sum_{j=1}^{\infty} \beta^{|i-j|} c_j}{\sum_{m=1}^{\infty} c_m} &= \sum_{j=1}^{\infty} \left(\beta^{|i-j|} \frac{c_j}{\sum_{m=1}^{\infty} c_m} \right) \leq \sum_{j=1}^{\infty} (\beta^{|i-j|} \bar{\alpha}) \\ &= \bar{\alpha} \sum_{j=1}^{\infty} \beta^{|i-j|} \leq \bar{\alpha} \frac{2}{1-\beta} \rightarrow 0. \quad \square \end{aligned}$$

References

- Adams, W. J., J. L. Yellen. 1976. Commodity bundling and the burden of monopoly. *Quart. J. Econom.* **90**(3) 475–498.
- Armstrong, M. 1996. Multiproduct nonlinear pricing. *Econometrica* **64**(1) 51–75.
- Armstrong, M. 1999. Price discrimination by a many-product firm. *Rev. Econom. Stud.* **66** 151–168.
- Bakos, Y., E. Brynjolfsson. 1999. Bundling information goods: Pricing, profits and efficiency. *Management Sci.* **45**(12) 1613–1630.
- Bakos, Y., E. Brynjolfsson. 2000. Bundling and competition on the internet. *Marketing Sci.* **19**(1) 63–82.
- McAfee, R. P., J. McMillan, M. D. Whinston. 1989. Multiproduct monopoly, commodity bundling, and correlation of values. *Quart. J. Econom.* **104**(2) 371–383.
- Norman, P., H. Fang. 2003. To bundle or not to bundle. Working paper, University of British Columbia, Vancouver, BC, Canada.
- Schmalensee, R. 1984. Gaussian demand and commodity bundling. *J. Bus.* **57**(1) S211–S230.