On the Profitability of Bundling Sale Strategy for Online Service Markets With Network Effects

LI YE, The Chinese University of Hong Kong
WEIJIE WU, Huawei Technologies Co., Ltd
RICHARD T. B. MA, National University of Singapore
JOHN C. S. LUI, The Chinese University of Hong Kong

In recent years, we have witnessed a growing trend for online service companies to offer "bundling sales" to increase revenue. Bundling sale means that a company groups a set of products/services and charges this bundle at a fixed price, which is usually less than the total price of individual items in the bundle. In this work, our aim is to understand the underlying dynamics of bundling, particularly what is the optimal bundling sale strategy and under what situations it will be more attractive than the separate sales. We focus on online service markets that exhibit *network effects*. We formulate mathematical models to capture the interactions between buyers and sellers, analyze the market equilibrium and its stability, and provide an optimization framework to determine the optimal sale strategy for a service provider. We analyze the impact of various factors on the profitability of bundling, including the network effects, operating costs, and variance and correlation of customers' valuations toward these services. We show that bundling is more profitable when the variance of customers' valuations and the operational cost of the services are small. In addition, a positive network effect and a negative correlation among customers' valuation on services increase the profitability of bundling, whereas the heterogeneity of services and the asymmetry of operating costs reduce its advantage.

CCS Concepts: • **Networks** → **Network economics**; • **Applied computing** → **Decision analysis**; **Marketing**; • **Mathematics of computing** → *Distribution functions*;

Additional Key Words and Phrases: Bundling sale, Online service market, Company's profit

ACM Reference format:
Li Ye, Weijie Wu, Richard T. B. Ma, and John C. S. Lui. 2019. On the Profitability of Bundling Sale Strategy for Online Service Markets With Network Effects. *ACM Trans. Internet Technol.* 19, 3, Article 31 (May 2019), 32 pages.
https://doi.org/10.1145/3277667

1 INTRODUCTION

As the economy becomes more global and competitive, it is becoming more important for online service companies to find new ways to increase their revenue. One way to achieve this is via *bundling service*. Bundling service (or bundling sale) means that companies group a set of their

Authors' addresses: L. Ye and John C. S. Lui, H.S.H Engineering Building, The Chinese University of Hong Kong, Hong Kong; emails: {yli, cslui}@cse.cuhk.edu.hk; Richard T. B. Ma, 15 Computing Drive, COM2 Building, #04-27, National University of Singapore, Singapore; email: tbma@comp.nus.edu.sg; W. Wu, email: weijie@0x.org.

1533-5399/2019/05-ART31 $15.00
https://doi.org/10.1145/3277667

Table 1. Example of Bundling Sales

	Service A	Service B	Bundle
Customer 1	\$5	\$10	\$15
Customer 2	\$10	\$5	\$15

products/services and use a single price to sell these products/services. Usually, the price of this bundling service is less than the total price of individual items. For online service markets, services are provided over the Internet infrastructure. Typical services include instant messaging, online social networks, online games, online recommendation systems, and so forth. Companies usually want to expand the service scale so as to increase their market share. Although most online service providers do not charge ordinary services, they do charge users for premium services. For example, the largest movie recommendation network, IMDb, has an "IMDbPro" session where premium services ("Get informed," "Get connected," and "Get discovered") are provided to paid users only. These premium services are often provided in a bundling manner. For example, customers are not allowed to buy these three "Get" services separately, but they have to pay one single price to obtain the premium services as a whole. In addition, for online services in telecommunications, ISPs often bundle broadband access, VoIP, and TV together with a single price (also called the *triple play*). Another example is that Microsoft bundles all software in Office as a suite (including Word, PowerPoint, Excel, Access, etc.) in the Office365 online service, which has been a great commercial success. Although the bundling sale strategy is common for online service markets, researchers have limited understanding of the underlying rationales and dynamics.

There are many reasons service providers offer bundling sales. An appealing reason is to reduce cost. Online services usually share the same network or storage infrastructure, and therefore the cost to provide an extra service on the same infrastructure is often marginal. Another reason is that by bundling, service providers can reduce the variance of customers' reservation prices on the services, thereby increasing the revenue of the product. Here, a customer's reservation price refers to a value such that she will purchase the service if and only if its price is no higher than this value. For example, if customer 1's reservation price on service A is \$5, and if the sale price of service A is less than or equal to \$5, then customer 1 will *subscribe* to this type of service. Note that different customers have different reservation prices toward each service. In Table 1, we use a simple example to illustrate this concept. Suppose that a company sells two services (A and B) to two customers. The second and third columns depict both customers' reservation prices on the services. Assume a customer's reservation price of the bundle is the sum of reservation prices of individual services, and the two customers have the same reservation price on the bundle. If the services are sold separately, they can be priced at \$5 and attract both customers (hence, the revenue is \$20), or they can be priced at \$10 and attract only one customer for each service (hence, the revenue is also \$20). In contrast, if services are bundled and priced at \$15, then both customers will purchase the bundled service, and the total revenue is \$30. This shows that bundling can reduce the *variance* of customers' reservation prices on these services, and thus the company can increase its revenue.

One important feature of online services is the "network effects." This refers to the market effect at the customer's side where a particular customer's interest on a service is influenced by other customers' purchasing decisions. For example, in online social networks (e.g., Facebook, LinkedIn, Twitter, IMDb), when the number of membership increases, the benefit that each member receives also increases due to a higher degree of interaction and efficiency of information spreading, and this causes more users to subscribe to the service. This is a prime example for an online service market in which a large population size indicates a positive influence on each customer's valuation, and we call this the *positive network effects*.[1] As we will show, this effect has a major impact on the choice of pricing strategies for online service providers.

Several existing research works [1, 21, 22, 26] discussed bundling strategies, but most of them focused on *non-digital goods or services* and were mainly based on graphical explanations, case studies, or algorithmic approaches. Very limited work focused on formal mathematical models to provide deeper insights. Furthermore, most existing works did not consider the impact of network effects, so they can only provide limited insights for the online service market. In this article, we aim to answer the following questions:

- Is it more profitable for online service providers to bundle a number of services and sell them at a single price?
- What are the factors that impact the optimal pricing strategy with network effects?

Our contributions are the following:

- We provide a mathematical model that captures the online service market with network effects.
- We analyze the market equilibrium and formulate an optimization framework to determine the optimal sale strategy.
- We discuss the impact of different factors on the profitability of the bundling strategy. We show that bundling is more profitable when the variance of customers' valuations and the operational cost of the services are small. In addition, a positive network effect and a negative correlation among customers' valuation on services increase the profitability of bundling, whereas the heterogeneity of services and the asymmetry of operating costs reduce its advantage.

Our article is organized as follows. Section 2 presents a general model to capture customers' purchasing decision and the service provider's profit. Section 3 focuses on the online service market, analyzes the market equilibrium and its stability, and presents an optimization framework to capture the optimal sale strategies. In Sections 4 and 5, we analyze the role of network effects and operating cost on the profitability of bundling. In Sections 6 and 7, we discuss the role of customers' valuations toward the services on the profitability of bundling. In particular, Sections 6 and 7 discuss the variance and the correlation of the valuations, respectively. Section 8 states related work. Section 9 concludes and highlights several limitations on the assumptions of our modeling approach.

2 GENERAL MODEL

We present a general model to characterize the Internet service market, as well as how customers and the service provider make their purchasing/pricing decisions. Let us first provide some formal definitions on sale strategies.

Definition 2.1. *Separate sale* is a strategy by which a service provider sells each individual service S_i at price p_i. Customers can choose to purchase such service or not.

Definition 2.2. *Bundling sale* is a strategy by which a service provider offers to sell a set of services as a single unit. The bundling service is priced at p_b. Customers can only choose to purchase the whole bundling service or not.

[1]There is also a *negative network effect* if a large number of users causes congestion. However, congestion is a physical-level infrastructural problem but is not the focus of our article on the application-level pricing problem.

2.1 Utility Functions of Separate Sales

Customers' utility. A customer decides whether to purchase the service(s) provided by the service provider. We consider a single service provider and a continuum of customers with different reservation prices on the service. The customers' heterogeneity in reservation prices is represented by their *types*—for instance, each infinitesimal customer is characterized by a one-dimensional type parameter $\theta \in \Theta$, which has a continuous distribution over Θ. The customer's utility function describes her purchasing behavior: a customer subscribes to a service if and only if she achieves a non-negative utility. This utility function depends on (1) the customer's reservation price on the service and (2) the sale price p_i of the service. We assume customer θ's reservation price on S_i is $v_i(\theta)\rho_i(\delta_i)$, where $v_i(\theta)$ is her intrinsic valuation on S_i, δ_i is the fraction of customers that subscribe to S_i, and $\rho_i(\delta_i)$ is a non-decreasing function in δ_i representing the network effects. We choose the multiplication form to represent the reservation price for the following two reasons. First, this form captures the fact that a customer with a higher intrinsic valuation is more sensitive to the network effects. Consider a user of the "triple-play" telephone service. If the user has a high "intrinsic value" for this service (or potentially a heavy user), then he would expect that many users are also using the telephone, as the valuation of using this service highly depends on how many other users can be reached by using the telephone service. In contrast, a light user may not be impacted much by the number of other users who use the telephone service. Second, this form represents that the reservation price is zero if no one uses it (or $\rho_i(\delta_i) = 0$). Consider the telephone user again. If he is the only user of telephone, then he cannot make phone calls to anyone, and thus the valuation for him is zero. We note that another commonly used model is in the additive form (e.g., $v_i(\theta) + \rho_i(\delta_i)$). One can easily verify that an additive form cannot represent heterogeneous sensitivity of network effects on users, nor can it represent a zero valuation of the service if no one else uses it. We define $u_i(\theta)$ as customer θ's utility on service S_i:

$$u_i(\theta) = v_i(\theta)\rho_i(\delta_i) - p_i. \tag{1}$$

Customers of different types have different intrinsic valuations $v_i(\theta)$ on S_i, and we assume $v_i(\theta)$ has a continuous distribution in θ over Θ. We further denote $f(\theta)$ as the density function of θ and define

$$H_i(x) \triangleq \int_{\theta\in\Theta} \mathbf{1}_{\{v_i(\theta)\le x\}} f(\theta) d\theta \tag{2}$$

as the cumulative distribution function of $v_i(\theta)$—that is, given any value x, $H_i(x)$ represents the fraction of customers whose intrinsic valuation on service S_i is less than or equal to x.

Service provider's utility. The service provider determines whether it should provide a separate or bundling sale and proposes the price(s) for the service(s). We model the service provider's utility as its total profit, and we will use "utility" and "profit" interchangeably in later analysis. We consider two factors that impact the service provider's utility: (1) the service fee received from customers, which we model as $p_i\delta_i$, (2) and the variable operating cost,[2] which we model as $m_i\delta_i$. Here, m_i represents the *per-unit variable cost*,[3] and we call it the *unit operating cost* or the *unit cost* for short. We define

$$U_i = (p_i - m_i)\delta_i \tag{3}$$

[2]Variable cost and fixed cost consist of the total cost. We consider the variable cost only because the fixed cost only represents a linear shift on the utility and does not affect our conclusion.

[3]Some existing literature uses the term *marginal cost* to represent this concept. In fact, if the marginal cost is a constant, its value is equal to the per-unit variable cost that we define here.

as the service provider's utility on service S_i. Suppose that we have services $S_1, S_2, \ldots, S_I$. Then the service provider's utility from all separate sale services is

$$U_s = \sum_{i=1}^{I} (p_i - m_i)\delta_i. \tag{4}$$

2.2 Utility Functions of the Bundling Sale

In the previous section, we expressed the utilities of customers and service provider when services are sold *individually*. We now consider the bundling strategy that combines all the services $S_1, S_2, \ldots, S_I$. Customers often view the bundled services as a whole. We use the notation b to denote the bundling service S_b. For consistency, we still assume that the network effect function impacts the utilities of the bundle in a multiplication manner. By substituting b for i, we denote the corresponding notations for S_b. In particular, $u_b(\theta)$, δ_b, and p_b represent customer θ's utility on purchasing S_b, the fraction of users purchasing S_b, and the price charged for S_b, respectively. We have the customer's and the service provider's utility functions as

$$u_b(\theta) = \left(\sum_{i=1}^{I} v_i(\theta)\rho_i(\delta_b) + \Delta v\right) - p_b \text{ and } U_b = (p_b - m_b)\delta_b, \tag{5}$$

where $v_i(\theta)\rho_i(\delta_b)$ is the valuation of the service i. The term Δv represents the extra valuation from the bundle composition compared to total valuation of separate services. A positive Δv means that the services are *complementary* to each other, or that a service will have a higher value if it is bought together with other services. Similarly, a negative Δv indicates that the services are *substitutable.* Now, customer θ's valuation of the bundle is $\sum_{i=1}^{I} v_i(\theta)\rho_i(\delta_b) + \Delta v$. Here we assume that the network effect function ρ_i is the same for service i whether it is bundled or not. Moreover, m_b denotes the unit cost for S_b. In particular, we assume $m_b = \sum_{i=1}^{I} \beta_i m_i$, where $\beta_i \in [0, 1]$ denotes the *scaling factors* of the operating cost. In fact, $\beta_i \leq 1$ implies that bundling can reduce the unit costs. For example, if we bundle a number of bandwidth-related functionalities in online game services, then the services can rely on the same infrastructure and save cost. In addition, we define $F_b(\cdot; \delta_b)$ as the cumulative distribution function of customers' valuation for the bundle when the fraction of users who purchase the bundle is δ_b, where

$$F_b(x; \delta_b) \triangleq \int_{\theta \in \Theta} \mathbf{1}_{\{\sum_{i=1}^{I} v_i(\theta)\rho_i(\delta_b) + \Delta v \leq x\}} f(\theta) d\theta. \tag{6}$$

2.3 Market Equilibrium

Due to the network effects, and that customers subscribe to services at different times, the preceding model is in fact a dynamic system. We use the following definition to describe the steady state of the system.

Definition 2.3. Given price p_i, $\delta_i > 0$ is a *market equilibrium* if

$$\int_{\theta \in \Theta} \mathbf{1}_{\{u_i(\theta) \geq 0\}} f(\theta) d\delta = \delta_i, \tag{7}$$

where $f(\theta)$ is the density function of θ, for $i \in \{1, \ldots, I, b\}$.

This definition states that for any given customer's utility $u_i(\theta)$, if exactly δ_i fraction of customers have a non-negative utility to purchase S_i, then δ_i is a market equilibrium. This represents the fraction of customers who purchase the service S_i when the system reaches a steady state. For instance, given this fraction, no customer has an incentive to change her decision. In the following, our analysis is based on this equilibrium. We will discuss pricing strategies under such a scenario. Unless we state otherwise, we will use δ_i to denote the equilibrium in the remainder of this article.

Note that when $\delta_i = 0$, it may also be a steady state with no user. But for this case, the service is closed and there is no real market. Thus, we exclude $\delta_i = 0$ from the definition of the equilibrium. Now let us characterize the value of δ_i.

Lemma 2.4. *Assume $\rho_i(\delta_i) > 0$ for any $\delta_i > 0$ ($i \in \{1, \dots, I\}$). For separate sale, the value δ_i is an equilibrium if and only if it satisfies $\delta_i = 1 - H_i(\frac{p_i}{\rho_i(\delta_i)})$, where $H_i(\cdot)$ is the cumulative distribution function of $v_i(\theta)$. For bundling sale, the value δ_b is an equilibrium if and only if it satisfies $\delta_b = 1 - F_b(p_b; \delta_b)$.*

Proof. Please refer to the appendix. □

The preceding lemma gives an implicit form to characterize and compute the equilibrium. In later analysis, it is more convenient to use the following corollary.

Corollary 2.5. *Assume $H_i(\cdot)$ is a strictly increasing function in $[0, V_i]$, and $\rho_i(\delta_i) > 0$ for any $\delta_i > 0$. For separate sale, given any price p_i, if there exists an equilibrium δ_i, then it is a solution to the following equation:*

$$p_i = P_i(\delta_i) \triangleq \rho_i(\delta_i) H_i^{-1}(1 - \delta_i), \tag{8}$$

where $H_i^{-1}(\cdot)$ is the inverse function of $H_i(\cdot)$ defined in $[0, 1]$. For bundling sale, given any price p_b, if there exists an equilibrium δ_b, then it is a solution to the following equation:

$$p_b = P_b(\delta_b) \triangleq F_b^{-1}(1 - \delta_b; \delta_b), \tag{9}$$

where $F_b^{-1}(x; \delta_b)$ is the inverse function of $F_b(x; \delta_b)$ with respect to x defined in $x \in [0, 1]$.

Until now, we have set up a general model to capture the customers' and the service provider's utilities. Based on this model, we proceed to analyze the properties of the market.

3 ONLINE SERVICE MARKET: EQUILIBRIUM AND OPTIMAL SALE STRATEGY

In this section, we study an online service market. We first model the network effects and the users' valuation distribution, and then we analyze the market equilibrium. Last, we establish a framework to determine the optimal sale strategies.

3.1 Network Effects and Utility Functions

We model the network effects in the form of $\rho_i(\delta_i) = \delta_i^{\alpha_i}$, where $\alpha_i \in [0, +\infty)$ represents the shape of the network effect function and the *degree of network effects*. A larger α_i means that the network effect function is more convex, which indicates stronger positive network effects. As discussed in Section 2.2, we assume that bundling does not change the network effects for each included service, which means that the parameters of network effects are still α_i when the service S_i is bundled.

We use the preceding form for a number of reasons. First, $u_i(\theta) = 0$ if $\delta_i = 0$—for instance, no customer has an incentive to enter an empty market. This is a common fact in many interactive applications, such as online social network or recommendation systems, and this also shows that it is important for a service provider to promote the service and have some initial users when launching the service. Second, $\delta_i^{\alpha_i}$ is increasing in δ_i, so it represents a positive network effect. Last, but not least, this is an *iso-elasticity* function that allows us to use a single parameter α_i to model the elasticity, or the shape of the network effects. Large α_i (or a *convex* function) means that given a small δ_i, $\delta_i^{\alpha_i}$ is small and many users will lose their interest, so a large start-up population is necessary. However, if α_i is small (or a *concave* function), then $\delta_i^{\alpha_i}$ is large given a moderate or small δ_i. This means that a small number of initial users is sufficient to induce a large network effect, and later on, the service can potentially attract many more customers. Note that our model generalizes the linear network effect models in many existing works [9, 14, 24, 25]; in fact, when $\alpha_i = 1$, our model exactly represents the linear network effects.

Now, service S_i is uniquely defined by (1) the users' maximal intrinsic valuation V_i, (2) the unit operating cost m_i, and (3) the network effect parameter α_i. In later analysis, we use a tuple $S_i = \langle V_i, m_i, \alpha_i \rangle$, $i = 1, 2$ to denote a separate sale service. Based on the preceding discussion, customer θ's utility and the service provider's utility on service S_i are

$$u_i(\theta) = v_i(\theta)\delta_i^{\alpha_i} - p_i, \quad U_i = (p_i - m_i)\delta_i \tag{10}$$

and the service provider's utility on all separate sales is

$$U_s = \sum_{i=1}^{2} U_i = \sum_{i=1}^{2} (p_i - m_i)\delta_i. \tag{11}$$

For the bundling sale, the utility functions are

$$u_b(\theta) = v_1(\theta)\delta_b^{\alpha_1} + v_2(\theta)\delta_b^{\alpha_2} + \Delta v - p_b, \quad U_b = \left(p_b - \sum_{i=1}^{2} \beta_i m_i\right)\delta_b, \tag{12}$$

respectively, where β_i is the scaling factor of the operating cost, defined in the previous section.

3.2 Distributions of Users' Valuations

Recall that we characterize the valuation of an customer θ on the bundle as $\sum_{i=1}^{I} v_i(\theta)\rho_i(\delta_b) + \Delta v$, given the customer intrinsic valuation on each individual service S_i as $v_i(\theta)$. Let us first assume the customers' intrinsic valuation on different services are independent. Note that in practice, a user's valuation for different services cannot be entirely independent. The independence case serves as a baseline that allows us to focus on other factors (e.g., network effects). With the independence assumption, if we let $H_i(x)$ and $F_B(x;\delta_b)$ be the cumulative distribution functions of $v_i(\theta)$ and $\sum_{i=1}^{I} v_i(\theta)\rho_i(\delta_b)$, respectively, we have $F_B(x;\delta_b) = (\rho_1(\delta_b)H_1(x)) \otimes (\rho_2(\delta_b)H_2(x)) \otimes \cdots \otimes (\rho_I(\delta_b)H_I(x))$, where the convolution operation is defined by $H_i(x) \otimes H_j(x) = \int H_i(x - t)dH_j(t)$.[4] In Section 7, we will relax this assumption and consider the correlation among services.

Let $F_b(x;\delta_b)$ be the cumulative distribution of $\sum_{i=1}^{I} v_i(\theta)\rho_i(\delta_b) + \Delta v$. We have $F_b(x;\delta_b) = F_B(x - \Delta v;\delta_b)$. Namely, given Δv, $F_B(x - \Delta v;\delta_b)$ is the fraction of users whose valuation for the bundle is less than x. In what follows, we set $\Delta v = 0$ and use $F_b(x;\delta_b) = F_B(x - \Delta v;\delta_b)$ as the baseline analysis. On the one hand, when the services are complementary (i.e., $\Delta v \geq 0$), the cumulative distribution $F_b(x;\delta_b) = F_B(x - \Delta v;\delta_b)$ is *upper bounded* by $F_B(x;\delta_b)$. The "result 3" on page 15 in Venkatesh and Kamakura [27][5] shows that if in this baseline analysis bundling achieves a profit gain over the separate sale under a certain circumstance, then bundling can achieve at least as good or even higher profit gain under the distribution $F_b(x;\delta_b)$. On the other hand, when the services are substitutable (i.e., $\Delta v \leq 0$), Venkatesh and Kamakura [27] show that bundling is less likely to be profitable compared to our baseline distribution $F_B(x;\delta_b)$. Correspondingly, we will also use $v_b(\theta) = \sum_{i=1}^{I} v_i(\theta)$ as the baseline analysis in the remainder of this article.

We focus on bundling two services: S_1 and S_2. This model represents a wide range of bundling strategy decisions, because any bundling of multiple services can be constructed by bundling two services. For analytical tractability, we do not consider a general distribution of customers'

[4]This is a standard result in probability theory, and we omit its proof.
[5]Venkatesh and Kamakura [27] consider a similar model to ours and discuss the bundle of two items.

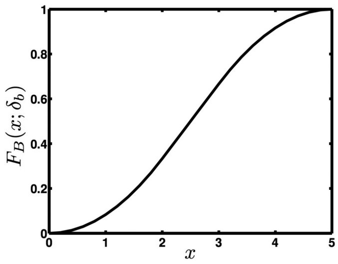


Fig. 1. Cumulative distribution $F_B(x;\delta_b), V_1 = 2, V_2 = 3, \rho_1(\delta_b) = \rho_2(\delta_b) = 1$.

valuations represented by the abstract $H_i(x)$ function[6] but focus on some specific distributions. We first consider the uniform distribution of customer's intrinsic valuation, which is widely adopted in economic literature. For example, throughout the work of Bakos and Brynjolfsson [2], uniform distribution is used to illustrate the profitability of bundling. In addition, in Section 2 of Matutes and Regibeau [21] and Section 3 of Guérin et al. [14], the authors explicitly use an uniform distribution to analyze bundling. Note that the uniform distribution cannot be "accurate" to capture a practical valuation distribution. The choice of uniform distribution enables us to decouple the impact of distribution from other important factors (e.g., operating cost, network effects). Furthermore, the simple form of the uniform distribution helps us present our framework and clarify the key results. Later in Section 6, we will extend our model to consider the Gaussian distribution of valuations and discuss the impact of different valuation distributions. In the following Sections 3, 4, and 5, we consider the uniform distribution and define the cumulative distribution of $v_i(\theta), i = 1{,}2$ as

$$H_i(x) = \begin{cases} 0 & \text{if } x < 0, \\ x/V_i & \text{if } 0 \le x \le V_i, \\ 1 & \text{if } x > V_i, \end{cases} \tag{13}$$

where $V_i \ (i = 1, 2)$ is the maximal intrinsic valuation of S_i. Without loss of generality, we let $V_1\rho_1(\delta_b) \le V_2\rho_2(\delta_b)$, and we have the following lemma.

Lemma 3.1. *The baseline distribution function $F_B(x;\delta_b)$ is*

$$F_B(x;\delta_b) = \begin{cases} 0 & \textit{if } x < 0, \\ x^2/(2V_1\rho_1(\delta_b)V_2\rho_2(\delta_b)) & \textit{if } 0 \le x \le V_1\rho_1(\delta_b), \\ (2x - V_1\rho_1(\delta_b))/(2V_2\rho_2(\delta_b)) & \textit{if } V_1\rho_1(\delta_b) < x \le V_2\rho_2(\delta_b), \\ 1 - (V_1\rho_1(\delta_b) + V_2\rho_2(\delta_b))^2/(2V_1\rho_1(\delta_b)V_2\rho_2(\delta_b)) & \textit{if } V_2\rho_2(\delta_b) < x \le V_1\rho_1(\delta_b) + V_2\rho_2(\delta_b), \\ 1 & \textit{if } x > V_1\rho_1(\delta_b) + V_2\rho_2(\delta_b). \end{cases}$$

Proof. The convolution operations on $H_1(x)\rho_1(\delta_b)$ and $H_2(x)\rho_2(\delta_b)$ directly lead to the result. □

In Figure 1, we illustrate the shape of this distribution when $\rho_1(\delta_b) = \rho_2(\delta_b) = 1$. It shows that $F_B(x;\delta_b)$ increases more rapidly in the middle range of the interval; in other words, customers are more *concentrated* to have a moderate valuation of the bundle, compared to the uniform distribution of separate services. This shows that bundling can *reduce the variance* of customers' valuations, and it is an important underlying reason to make bundling profitable: if the service provider sets a relatively low bundling price, then it will be easier for him to attract more customers because there are a lot of customers with moderate valuations, and hence the service provider can make

[6]Although our model is general for any c.d.f. $H_i(x)$, analysis on general $H_i(x)$ is intractable. First, the c.d.f. of the valuation of the bundle $F_B(x;\delta_b)$ involves integration. Second, the optimal sale strategy in (15) is hard to analyze for a general $H_i(x)$.

more profit. However, if the service provider only targets a small amount of customers with high valuations, then bundling may not have an advantage. This is because $F_B(x;\delta_b)$ indicates fewer customers with high valuations compared to the uniform distribution.

3.3 Analysis of Market Equilibrium

In this section, we derive the conditions for the existence of the market equilibrium (or equilibria).

Theorem 3.2. *Consider any service S_i for separate sale, $i \in \{1,2\}$. There exists a threshold $\bar{p}_i = \frac{V_i}{\alpha_i+1}(\frac{\alpha_i}{\alpha_i+1})^{\alpha_i}$ such that for any given service price p_i, we have*

$$\textit{no. of equilibrium (or equilibria)} = \begin{cases} 0 & \textit{if } p_i > \bar{p}_i, \\ 1 & \textit{if } p_i = 0 \textit{ or } \bar{p}_i, \\ 2 & \textit{if } p_i \in (0,\bar{p}_i). \end{cases}$$

For the bundling sale S_b, there exists a threshold $\bar{p}_b$ such that for any given service price p_b, we have

$$\textit{no. of equilibrium (or equilibria)} \begin{cases} = 0 & \textit{if } p_b > \bar{p}_b, \\ \geq 1 & \textit{if } p_b = 0 \textit{ or } \bar{p}_b, \\ \geq 2 & \textit{if } p_b \in (0,\bar{p}_b). \end{cases}$$

Proof. Please refer to the appendix. □

This theorem states that the condition for the existence of equilibrium is that the service price cannot be too high; otherwise, no customer will purchase the service. We also note that if the existence is guaranteed, then almost surely there are at least two equilibria for both the separate sale and bundling sale. Next we discuss the stability property and explain why we are interested in the largest equilibrium.

Discussion on stability. We say that an equilibrium δ is *stable* if there exists an $\epsilon > 0$ such that if at any time a non-equilibrium fraction $\delta' \in (\delta - \epsilon, \delta + \epsilon)$ of customers subscribe to the service, then the dynamic market will eventually reach the equilibrium δ. In fact, for the separate sale case, if we consider the two equilibria $\delta_i^1 < \delta_i^2$ ($i \in \{1,2\}$) for the preceding theorem, the only *stable* one is δ_i^2. If the market is with $\delta_i^1 - \epsilon$ fraction of customers, then eventually all customers will leave the market, and the service will be closed; if the market is with $\delta_i^1 + \epsilon$ fraction of customers, it will not reach δ_i^1 but will reach δ_i^2. Hence, δ_i^1 is an *unstable* equilibrium. In contrast, if we consider the market with any fraction $\delta_i^{2'} \in (\delta_i^2 - \epsilon, \delta_i^2 + \epsilon)$ of users, the dynamic market will eventually reach the equilibrium δ_i^2; hence, δ_i^2 is a *stable* equilibrium. For the bundling sale, we can show that the largest equilibrium is stable, and in many cases it is the only stable equilibrium, as stated in the appendix. Interested readers may also refer to the work of Easley and Kleinberg [9] for a detailed discussion. In addition, the features of our equilibria are quite similar to discussions in the work of Buragohain et al. [6]. In an online service market, a service provider will try his best to attract potential users and operate the service with this stable equilibrium. Due to this stability property, in later analysis we can safely focus our analysis for the largest equilibrium. We use $\max\{\delta_i(p_i)\}$ to denote the maximal equilibrium for a given p_i and define $\max \emptyset = 0$ to capture an empty market when the price is too high and the equilibrium does not exist.

3.4 An Optimization Framework for the Sale Strategies

In this section, we establish an optimization framework to determine the optimal sale strategies. In natural way to model the optimal sales, the service provider aims to find a price (or prices) for the separate or the bundling sale that maximizes its profit. For instance, the optimal separate or bundling sale $S_i, i \in \{1, 2, b\}$, can be modeled as

$$\begin{aligned} \max_{p_i} \quad & U_i(p_i) = (p_i - m_i)\max\{\delta_i(p_i)\}, \\ \text{subject to} \quad & p_i \geq 0. \end{aligned} \tag{14}$$

However, this form is not easy to analyze, and we opt to change the decision variable from p_i to δ_i. We can transform (14) into the following problem[7]:

$$\begin{aligned} \max_{\delta_i} \quad & U_i(\delta_i) = (P_i(\delta_i) - m_i)\delta_i, \\ \text{subject to} \quad & 0 \leq \delta_i \leq 1. \end{aligned} \tag{15}$$

For separate sales, according to (8) in Corollary 2.5, we have $P_i(\delta_i) = \rho_i(\delta_i)H_i^{-1}(1 - \delta_i)$ for $i \in \{1, 2\}$. For bundling sale, we have $P_b(\delta_b) = F_b^{-1}(1 - \delta_b; \delta_b)$ from Easley and Kleinberg (9).

Since the preceding optimizations for $i \in \{1, 2, b\}$ have continuous objective functions over a compact set, they are guaranteed to have optimal solutions. By solving the preceding optimizations, we choose the largest equilibrium fraction for any given price, which is the stable one as we desire.[8] Until now, we have established an optimization framework to determine the optimal sale strategies. In what follows, we use U_s^* and U_b^* to denote the maximal profit of the service provider under the separate and bundling sales, respectively. If we could calculate U_s^* and U_b^*, then we can determine whether bundling is more profitable than separate sale by comparing their values. Formally, we have the following definition to capture the profit gain of bundling over the separate sale.

Definition 3.3. The *profit gain ratio* γ is the difference between the maximal profit of the bundling sale and that of the separate sale, divided by maximal profit of the separate sale—that is,

$$\gamma = (U_b^* - U_s^*)/U_s^*. \tag{16}$$

If $\gamma > 0$, it means the optimal bundling sale is more profitable than the optimal separate sale and vice versa. A larger value of γ indicates a larger profit gain by the bundling sale.

Note that this framework is critical for online service providers to evaluate their best sale strategies, but we would also like to point out that it may not be easy to have general results at this stage, particularly when a general α_i induces difficulty in solving the optimizations. In the following sections, we explore the impact of various factors—that is, the network effect parameter α_i and the unit cost m_i on the profitability of bundling. Before we proceed, let us present the following lemma that reflects the scaling properties of the sales.

Lemma 3.4 (Scaling property). *Let c be a positive number:*

(1) If the equilibrium and profit of the optimal sale $S_i = \langle V_i, m_i, \alpha_i \rangle$ are δ_i^* and U_i^*, then the equilibrium and profit of the optimal sale $S_i' = \langle cV_i, cm_i, \alpha_i \rangle$ are δ_i^* and cU_i^*.
(2) If the profit gain ratio for bundling $S_1 = \langle V_1, m_1, \alpha_1 \rangle$ and $S_2 = \langle V_2, m_2, \alpha_2 \rangle$ is γ, then the profit gain ratio for bundling $S_1' = \langle cV_1, cm_1, \alpha_1 \rangle$ and $S_2' = \langle cV_2, cm_2, \alpha_2 \rangle$ is also γ.

Applying the optimization framework, we can easily prove the preceding lemma. This lemma points out that if V_i and m_i increases (or decreases) by the same factor, then it does not impact the equilibrium or the profitability of bundling. Hence, in later analysis, we can normalize V_1 to be 1

[7]Although $\delta_i = 0$ is excluded from Corollary 2.5, we can verify that if the solution to (14) is p_i^* and $\max\{\delta_i(p_i^*)\} = 0$, then the solution to (15) is 0.

[8]To see this, note if there exist $\delta_i^1 < \delta_i^2$ such that $P_i(\delta_i^1) = P_i(\delta_i^2) \geq m_i$, then obviously δ_i^1 cannot be the solution to (15).

and vary V_2 so as to explore the whole design space. This simplifies our analysis and does not lose any generality.

4 IMPACT OF NETWORK EFFECTS

Up to now, we have formulated a framework to capture the pricing strategies and the market equilibrium. In this and the following sections, we will discuss the impact of various factors on the profitability of bundling. We first focus on the impact of network effects.

Many online service providers incur a much larger fixed cost compared to their variable cost. For example, the telecommunication services in "triple-play" (broadband, telephone, TV) need a large amount of investment to initially set up the hardware and infrastructure, but the cost is minimal to support an additional user. In this section, we set the per-unit variable cost $m_i = 0$ and consider $S_i = \langle V_i, 0, \alpha_i \rangle$. The service provider's utility can be expressed as

$$U_s = p_1\delta_1 + p_2\delta_2. \tag{17}$$

This simplification captures many features of a wide range of digital online services, and it allows us to isolate different factors to better understand the impact of network effects.

4.1 Homogeneous Network Effects ($\alpha_1 = \alpha_2 = \alpha$)

We start our discussion with the network effect functions $\rho_1(\delta) = \rho_2(\delta) = \delta^\alpha$. Such setting represents bundling two services with similar network effects. The following theorem shows that bundling is more profitable than separate sales under this setting.

Theorem 4.1. *Consider $S_1 = \langle V_1, 0, \alpha \rangle$ and $S_2 = \langle V_2, 0, \alpha \rangle$. The profit gain ratio of the bundling sale is $\gamma > 0$. In particular, when $S_1 = S_2 = \langle V, 0, \alpha \rangle$, we have*

(a) *The optimal separate sale is*

$$\delta_i^* = \frac{\alpha+1}{\alpha+2},\ p_i^* = \frac{V}{\alpha+2}\left(\frac{\alpha+1}{\alpha+2}\right)^\alpha,\ U_s^* = 2\delta_i^* p_i^*.$$

(b) *The optimal bundling sale is*

$$\delta_b^* = \frac{2\alpha+2}{2\alpha+3},\ p_b^* = \left(\frac{2\alpha+2}{2\alpha+3}\right)^\alpha \sqrt{\frac{2}{2\alpha+3}}V,\ U_b^* = \delta_b^* p_b^*.$$

(c) *The profit gain ratio of the bundling sale is*

$$\gamma(\alpha) = \frac{\sqrt{2}}{4}\frac{(2\alpha+4)^{\alpha+2}}{(2\alpha+3)^{\alpha+3/2}} - 1,$$

and it is an increasing function in α.

Proof. Please refer to the appendix. □

This theorem states that when $\alpha_1 = \alpha_2$, bundling is always more profitable, and *large* α (i.e., a *convex* network effect function) indicates a high profit gain. Let us use examples to show how bundling achieves a higher profit. In Figure 2, we consider $S_1 = \langle 1, 0, \alpha \rangle$ and $S_2 = \langle V_2, 0, \alpha \rangle$, where we vary $V_2 \in \{1, 2, 5\}$ and $\alpha \in [0.1, 3.0]$. We can see γ is always positive, and this validates our results in Theorem 4.1. We can also observe that when there is a large gap between V_1 and V_2, the profit gain ratio γ reduces. This is because the joint distribution $H_B(x)$ becomes less concentrated in the middle range, so bundling sale can attract fewer customers. Let us interpret these findings in applications. For the bundle of "triple-play," there exist positive network externalities for those telecommunications services, which make bundling more profitable. We can see that the network effects play an important role in the success of the "triple-play." In addition, the positive network

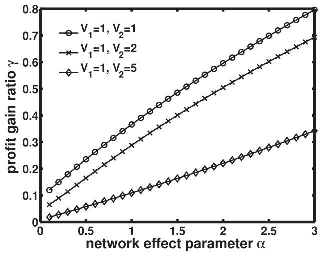


Fig. 2. Impact of network effects, $\alpha_1 = \alpha_2$.

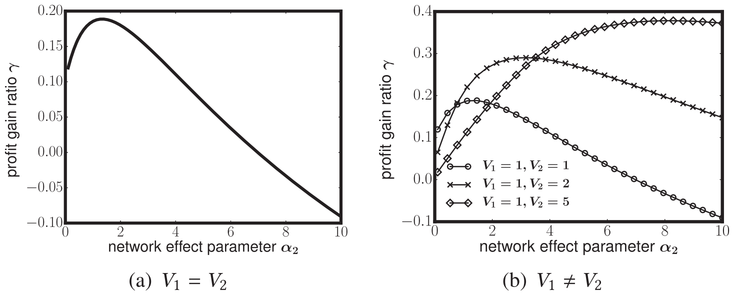


Fig. 3. Impact of network effects, $\alpha_1 \neq \alpha_2$.

effects of Microsoft Office software increase the profitability of the Office bundle. To summarize, we have the following observation.

Observation 1: The advantage of bundling becomes more apparent when the network effect function is more convex (i.e., larger α); however, the heterogeneity in intrinsic valuation distributions reduces the profit gain ratio of bundling.

4.2 Heterogeneous Network Effects ($\alpha_1 \neq \alpha_2$)

Now let us consider bundling two services with different network effects. We first consider $V_1 = V_2$.

Let us show the impact of network effects on γ. In Figure 3(a), we consider $S_1 = \langle 1, 0, 0.1 \rangle$ and $S_2 = \langle 1, 0, \alpha_2 \rangle$ where we vary $\alpha_2 \in [0.1, 10.0]$. When α_2 increases from 0.1, γ also increases; this is because a large network effect parameter has a positive impact on bundling. But when α_2 is large, γ begins to decrease and eventually becomes negative; this is because when α_1 and α_2 differ a lot, the optimal equilibria, δ_1^* and δ_2^*, are also different. In such cases, it is not rational to bundle S_1 and S_2, as the bundling sale needs to find a unique equilibrium δ_b^*, which is either far away from δ_1^* or far away from δ_2^*, so bundling is not as profitable as the separate sale.

Similar to the previous discussions, we also consider the services with different V_i. In Figure 3(b), we consider $S_1 = \langle 1, 0, 0.1 \rangle$ and $S_2 = \langle V_2, 0, \alpha_2 \rangle$ where we vary $\alpha_2 \in [0.1, 10]$ and $V_2 \in \{1, 2, 5\}$. We have similar observations: when α_2 increases, γ first increases and then decreases. We also observe that the inflection point increases when V_2 increases. This is because when V_2 is large, service S_2 has a major impact on the bundle, so the positive impact of α_2 on bundling can be effective in a larger range. To summarize, we have the following observation.

Observation 2: The heterogeneity of network effect functions reduces the profitability of bundling.

5 IMPACT OF OPERATING COST

In the previous section, we discussed the impact of network effects when the variable operating cost equals zero. Although this approximation applies to many existing services, there might be exceptions. For example, in online storage systems (e.g., Dropbox), the unit cost of storing the data might not be negligible. In this section, we discuss how the operating cost impacts the pricing strategies. Our discussions generalize the results we obtained in the previous section.

5.1 Impact of Operating Cost When $\alpha_1 = \alpha_2 = \alpha$

We start our discussion when both services have the same network effect function (i.e., $\rho_i(\delta_i) = \delta_i^\alpha, i = 1, 2$). We explore how our results in Theorem 4.1 can be generalized with non-zero unit operating costs. According to Lemma 3.4, we can normalize V_i so that the effectiveness of the unit cost is represented by $\frac{m_i}{V_i}$. We will discuss when $m_1 : m_2 = V_1 : V_2$ and $m_1 : m_2 \neq V_1 : V_2$—that is, symmetric and asymmetric unit costs, respectively (where the notation $m_1 : m_2$ is for m_1/m_2). We start from the symmetric case when $m_1 : m_2 = V_1 : V_2$.

Theorem 5.1. *If $\alpha \geq 1$ and $m_1 : m_2 = V_1 : V_2$, then $U_b^* \geq U_s^*$.*

Proof. Please refer to the appendix. □

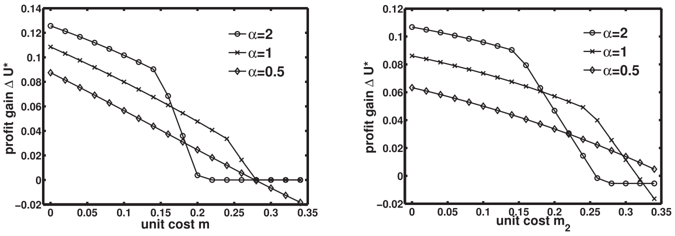
This theorem states that if S_1 and S_2 have the same convex network effect function and symmetric unit costs, then the profit of the optimal bundling is no less than that of the optimal separate sale. The key reason is that under this setting, the optimal separate sale always obtains an equilibrium larger than 1/2, so bundling can attract more customers. However, if $\alpha_i < 1$ and m is large, then δ_i^* may be less than 1/2 and bundling may not be always profitable. Let us use some examples to show this phenomenon. Since U_s^* may be zero, which leads $\gamma = \infty$, we opt to use the *profit gain*, defined by $\Delta U^* = U_b^* - U_s^*$, as the performance measure. In Figure 4(a), we consider $S_1 = S_2 = \langle 1, m, \alpha \rangle$ and vary $m \in [0, 0.34]$, $\alpha \in \{0.5, 1, 2\}$. We observe for any given α, ΔU^* reduces with respect to m. This indicates that unit cost *reduces* the profit gain of bundling. When $\alpha = 1$ or 2, bundling is always no worse than separate sale. This validates our result in Theorem 5.1. When $\alpha = 0.5$, ΔU^* can be negative when m is large. This indicates when the network effect functions are concave, a large unit cost can make bundling less profitable than separate sale.

Now let us consider the impact of asymmetric unit costs (i.e., $m_1 : m_2 \neq V_1 : V_2$). The asymmetry induces different equilibria, and bundling is not always more profitable. It is not easy to quantify the dominant domain of the bundling or the separate sales. We have the following theorem as a sufficient condition to guarantee the profitability of bundling.

Theorem 5.2. *Let $\delta_i^*, i = 1, 2$ be equilibria of optimal separates sales S_1 and S_2. If $\frac{\beta_2}{2-\beta_2}\delta_1^* \leq \delta_2^* \leq \delta_1^*$ (where β_2 is the scaling factor such that $m_b = \beta_1 m_1 + \beta_2 m_2$), then $U_b^* > U_s^*$.*

Proof. Please refer to the appendix. □

This theorem states that if the optimal equilibria of separate sales are close, then bundling is more profitable. The underlying reason is similar to the previous analysis: if two services are highly asymmetric and have very different equilibria, then it is not feasible to find a suitable service price for the bundle, because the corresponding equilibrium δ_b of the bundle is either too far from δ_1^* or too far from δ_2^*; only when δ_1^* and δ_2^* are close can bundling be more profitable. Let us use examples to show how the asymmetry impacts the profit gain. In Figure 4(b), we consider $S_1 = \langle 1, 0.14, \alpha \rangle$, $S_2 = \langle 1, m_2, \alpha \rangle$ and vary $m_2 \in [0, 0.34]$ and $\alpha \in \{0.5, 1, 2\}$. We can observe that when m_2 increases, the profit gain ΔU^* decreases; when m_2 is greatly larger than m_1, then ΔU^* can be negative for $\alpha = 1$ or 2. Comparing with Figure 4(a) where ΔU^* is always non-negative for $\alpha = 1$ or 2, we can see the asymmetry in unit costs further reduces the profitability of bundling.



(a) $m_1 = m_2$ (b) $m_1 \neq m_2$

Fig. 4. Impact of unit cost, $\alpha_1 = \alpha_2$.

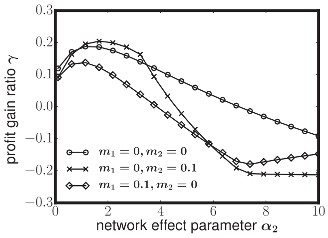


Fig. 5. Impact of unit cost, $\alpha_1 \neq \alpha_2$.

To summarize, we have the following observation.

Observation 3: Under the symmetric operating costs and homogenous network effects, the operating costs reduce the profitability of bundling; in particular, when the network effect function is concave, bundling may be less profitable than separate sales. When the operating costs are asymmetric, the profitability of bundling is further reduced.

5.2 Impact of Operating Cost When $\alpha_1 \neq \alpha_2$

In this section, we study the impact of the unit operating cost when the two services have different network effect functions. In particular, we show how our result in Figure 3(a) changes with the consideration of the unit cost. To be consistent with Figure 3(a), we evaluate the profit gain ratio γ. In Figure 5, we consider bundling $S_1 = \langle 1, m_1, 0.1\rangle$ and $S_2 = \langle 1, m_2, \alpha_2\rangle$. We fix $\beta_1 = \beta_2 = 1$ and vary $\alpha_2 \in [0.1, 10.0]$. We consider three cases of γ: $\gamma = \gamma_1$ if $m_1 = m_2 = 0$, $\gamma = \gamma_2$ if $m_1 = 0, m_2 = 0.1$, and $\gamma = \gamma_3$ if $m_1 = 0.1, m_2 = 0$. We first note $\gamma_3 < \gamma_1$. This means that a unit cost on S_1 discourages bundling, which is the same as our finding in the previous section. Next we focus on the curve of γ_2 and have some interesting observations. We note $\gamma_2 > \gamma_1$ when α_2 is moderately large. This shows that the unit cost of S_2 can sometimes increase the profitability of bundling. The reason is this unit cost reduces δ_2^*, so the gap between δ_1^* and δ_2^* reduces. Therefore, the unit cost reduces the asymmetry of S_1 and S_2, so it increases the profitability of bundling. However, when α_2 is large, the unit cost further reduces δ_2^* and its negative impact on the profit becomes dominant. To summarize, we have the following observation.

Observation 4: The operating costs play a significant role when network effect functions are different. In particular, a moderate operating cost on the service with larger α_i may increase the profitability of bundling.

6 IMPACT OF VARIANCE OF CUSTOMERS' VALUATIONS

In the previous two sections, we discussed the impact of the network effects and the operation cost on the profitability of bundling. Recall that one important reason for bundling to be more profitable is that it can reduce the variance of customers' valuations. One natural and interesting question is the impact of the *variance* itself with the existence of network effects, which we will focus on in this section.

6.1 Gaussian Distribution of Customers' Intrinsic Valuations

Recall that we have been using the uniform distribution to capture the distribution of customers' intrinsic valuations. This simple form has also been repeatedly used in the literature, capturing users' heterogeneity with mathematical tractability. However, this simple form does not allow us to capture how users' intrinsic valuations are distributed. Thus, in this section, we use a Gaussian distribution to capture users' intrinsic valuations. We will show that although it brings some analytical difficulty, it will reveal important insights on the impact of variance of customers' valuations. In addition, it will enable us to capture correlations among services, which we will demonstrate later.

Formally, let the cumulative distribution of $v_i(\theta)$ be

$$H_i(x) = \Phi\left(\frac{x-\mu_i}{\sigma_i}\right), i = 1, 2, \tag{18}$$

where $\Phi(\cdot)$ is the cumulative distribution function of standard Gaussian distribution, and μ_i and $\sigma_i^2 (\sigma_i \geq 0)$ are the mean and variance, respectively. We emphasize that Gaussian distribution has also been used in the literature to capture users' valuation distribution. For example, Schmalensee [26] argues that " the frequency with which normal distributions arise in the social sciences makes the Gaussian family a plausible choice to describe the distribution of tastes in a population of buyers."

In this section, we study the services with the same network effects (i.e., $\alpha_1 = \alpha_2 = \alpha$). In addition, we assume that customers' intrinsic valuations for these two services are independent. Hence, the cumulative distribution function of the sum of intrinsic valuations $v_1(\theta) + v_2(\theta)$ is

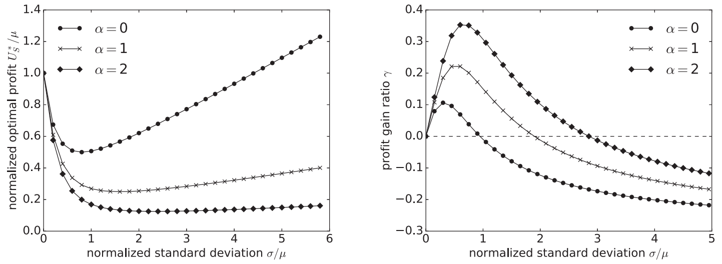
$$H_B(x) \triangleq \Phi\left(\frac{x-(\mu_1+\mu_2)}{\sqrt{\sigma_1^2+\sigma_2^2}}\right), \tag{19}$$

which is also a Gaussian distribution with mean $\mu_b \triangleq \mu_1 + \mu_2$ and variance $\sigma_b^2 \triangleq \sigma_1^2 + \sigma_2^2$.

We want to emphasize that our analytical framework applies for a wide range of realistic distributions, such as Gaussian/gamma/chi-square/Poisson distributions. This is because the sums of independent Gaussian/gamma/chi-square/Poisson random variables still follow Gaussian/gamma/chi-square/Poisson distributions. If we want to answer the question "whether bundling is more profitable than separate sale," what we need to compare are the valuation distributions for separate services versus the valuation distribution for the bundle. Under a family of distributions (e.g., Gaussian), the valuation distribution of the bundle has the same form as the valuation distributions of separate services, so we only need to compare the parameters to answer the aforementioned question.

Now we can derive the cumulative distribution function of customers' valuations as $F_B(x;\delta_b) = H_B(x)\delta_b^\alpha$. Our previous optimization framework also applies here. Similarly, we use δ_i as the decision variable. Then the problem for choosing $i \in \{1, 2, b\}$ can be formulated based on (15), simply by instantiating $P_i(\delta_i)$ as $\delta_i^\alpha\left(\Phi^{-1}(1-\delta_i)\sigma_i + \mu_i\right)$, or

$$\begin{aligned} \max_{\delta_i} \quad & U_i(\delta_i) = (\delta_i^\alpha\left(\Phi^{-1}(1-\delta_i)\sigma_i + \mu_i\right) - m_i)\delta_i, \\ \text{subject to} \quad & 0 \leq \delta_i \leq 1. \end{aligned} \tag{20}$$



(a) Optimal separate sale (b) Profit gain of bundling

Fig. 6. Impact of variance of customers' valuations, homogeneous case.

Similar to the previous sections, we use a tuple $S_i = \langle(\mu_i, \sigma_i^2), m_i, \alpha\rangle$ to denote an individual service and $S_b = \langle(\mu_b, \sigma_b^2), m_b, \alpha\rangle$ to denote the bundle of services with the same network effects. Moreover, the scaling property in Lemma 3.4 still holds.

Lemma 6.1 (Scaling property under the Gaussian model). *Let $c > 0$. If the profit gain ratio for bundling $S_1 = \langle(\mu_1, \sigma_1^2), m_1, \alpha\rangle$ and $S_2 = \langle(\mu_2, \sigma_2^2), m_2, \alpha\rangle$ is γ, then the profit gain ratio for bundling $S_1' = \langle(c\mu_1, (c\sigma_1)^2), cm_1, \alpha\rangle$ and $S_2' = \langle(c\mu_2, (c\sigma_2)^2, cm_2, \alpha\rangle$ is also γ.*

It is easy to prove this lemma via a similar approach we used previously.

6.2 Homogeneous Variance ($\sigma_1^2 = \sigma_2^2 = \sigma^2$)

Let us first discuss when the variances are homogeneous (i.e., $\sigma^2 \triangleq \sigma_1^2 = \sigma_2^2$). Recall that the network effects for the two products are characterized by a single parameter α (i.e., $\alpha_1 = \alpha_2 = \alpha$). For the other factors, we set $c_1 = c_2 = 0$ and $\mu_1 = \mu_2 = \mu$. This setting corresponds to the scenario of bundling two similar services. We make the mild assumption that the mean valuation for the product is non-negative (i.e., $\mu_i \geq 0, \forall i$). Then, based on Lemma 6.1, without loss of generality, we can normalize the customers' intrinsic valuation distribution to be $\mathcal{N}(1, (\sigma/\mu)^2)$, with the coefficient of variation σ/μ as the only parameter.

Let us first investigate in the impact of variance of valuations on the optimal separate sale. In Figure 6(a), we vary the normalized standard deviation σ/μ and observe the maximal profit under separate sales. We can see that the maximal profit first decreases and then increases when σ/μ increases for different network effects. This impact of the variance of customers' valuations is discussed and explained in other works [17, 26]. These works found that bundling benefits for thin-tailed valuations distributions (small σ/μ) but not for heavy-tailed valuation distributions (large σ/μ).

Recall that bundling reduces the customers' valuations. Thus, for any point on the curve in Figure 6(a), by bundling two independent and identical services, the maximal profit of the bundle should be some point to the left of the original point on the curve. Thus, bundling can be more (or less) profitable if the normalized standard deviation σ/μ of the separate service has (or has not yet) reached the turning point in Figure 6(a). Moreover, we can see that the network effect parameter α determines such turning point in the curve. If a service has a more convex network effect function (larger α), the turning point will be at a larger coefficient of variation, which means that bundling is profitable even when the variance of customers' valuations is high. Formally, we have the following theorem.

Theorem 6.2. *Consider two independent services $S_1 = S_2 = \langle(1, (\sigma/\mu)^2), 0, \alpha\rangle$ (the mean of valuations is 1 and the variance of valuations is $(\sigma/\mu)^2$, the marginal cost is 0 and the network effect parameter is α), where $\alpha \geq 0$:*

(1) *If* $\sigma/\mu < \sqrt{2/\pi} \cdot (\alpha+1)$, *then* $U_b^* > U_s^*$;
(2) *If* $\sigma/\mu > 2/\sqrt{\pi} \cdot (\alpha+1)$, *then* $U_b^* < U_s^*$.

Proof. Please refer to the appendix. □

This theorem explicitly states the relationship between the profitability of bundling and the network effects. When the network effect function is more convex, there is a larger range of variance for bundling to be more profitable. It has similar consequences to the case when the distribution of customers' valuation is uniform. However, this theorem also states that if the variance of customers' intrinsic valuations is large, bundling two identical services is not always more profitable even when $c = 0$. This is different from our previous results that bundling two services with zero operational cost can always be more profitable if their intrinsic valuation distributions are uniform. Furthermore, the profitability of bundling depends on both the variance of customers' valuations and the network effects. In particular, when the variance is small or the network effects are strongly positive, bundling is more profitable. However, when the variance is large or the network effects are not strongly positive, bundling is less profitable.

We use numerical examples to illustrate our finding. In Figure 6(b), we vary the coefficient of variation σ/μ and investigate the profit gain ratio of the bundling sale, We can see that there is a turning point of σ/μ and bundling is more profitable if and only if σ/μ is smaller than this turning point. Furthermore, a larger α indicates a higher profit gain of bundling and a larger range of variance for bundling to be more profitable. To summarize, we have the following observation.

Observation 5: When customers' valuations follow a Gaussian distribution with high variance, bundling can be less profitable than separate sales even if the marginal cost is zero. In particular, when the network effect function is more convex, there is a larger range of variance for bundling to be more profitable.

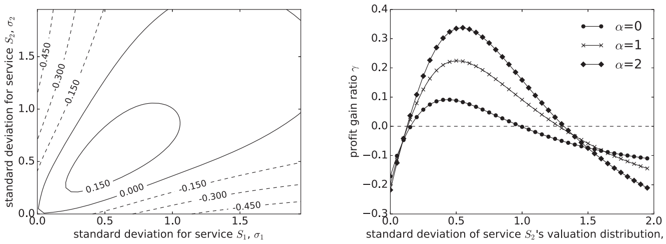
6.3 Heterogeneous Variance ($\sigma_1^2 \neq \sigma_2^2$)

So far, we have discussed the profitability of bundling when the variances of valuations are the same for different services. Now let us explore the case when services are with different variance of customers' valuations, which is not explored by previous works [17, 26] that study the impact of valuation distributions on bundling.

Theorem 6.3. *Consider two bundles: L: services* $\langle(\mu, (\sigma-\Delta_L)^2), 0, \alpha\rangle$ *and* $\langle(\mu, (\sigma+\Delta_L)^2), 0, \alpha\rangle$, *H: services* $\langle(\mu, (\sigma-\Delta_H)^2), 0, \alpha\rangle$ *and* $\langle(\mu, (\sigma+\Delta_H)^2), 0, \alpha\rangle$, *where* $\Delta_L < \Delta_H \le \sigma, \alpha \ge 0$. *Let* γ_L, γ_H *be the profit gain ratios of bundle L and H, respectively. If* $\gamma_H > 0$, *then* $\gamma_L > \gamma_H$.

Proof. Please refer to the appendix. □

The parameter Δ_L (or Δ_H) represents the heterogeneity of valuation variances for the services in the bundle L (or H). This theorem states that the profitability of bundling decreases when the heterogeneity of valuation variance increasesfor various degrees of positive network effects. In Figure 7, we have two services, $S_1 = \langle(1, \sigma_1^2), 0, \alpha\rangle$ and $S_2 = \langle(1, \sigma_2^2), 0, \alpha\rangle$, where the variances σ_1^2 and σ_2^2 could be different. Figure 7(a) is the contour plot of the profit gain ratio corresponding to a different standard deviation configuration (σ_1, σ_2) of the services. In the contour plot, points on the same curve indicates the same value of profit gain ratio. We can observe that only when σ_1 is close to σ_2 is the profit gain ratio positive. In Figure 7(b), we fix $\sigma_1 = 0.5$ and vary $\sigma_2 \in [0, 2]$. Figure 7(b) is in fact a projection of Figure 7(a) on a one-dimensional space where $\sigma_1 = 0.5$. We can see that for $\alpha \in \{0, 1, 2\}$, the profit gain ratio of bundling reaches the highest when σ_2 is close to σ_1. When σ_2 is far different from σ_1, the profit gain decreases, and bundling may be less profitable than separate sales. Furthermore, as we can see from Figure 7(b), if the network effect function is



(a) Contour plot of profit gain of bundling, with α=1. Dotted curve for negative value.

(b) Profit gain ratio, $\sigma_1 = 0.5$

Fig. 7. Impact of heterogeneous variance, $\sigma_1 \neq \sigma_2, \mu_1 = \mu_2 = 1$.

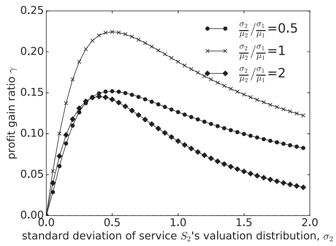


Fig. 8. Impact of heterogeneous variance, $\mu_1 \neq \mu_2, \frac{\sigma_2}{\mu_2} \propto \frac{\sigma_1}{\mu_1}, \alpha = 1, \sigma_1 = 0.5$.

more convex, then the value of the profit gain ratio will increase when the variance of customers' valuations is heterogeneous.

We also explore the impact of heterogeneous variance where the mean valuation of customers is no longer fixed. In Figure 8, we consider one fixed service, $S_1 = \langle(1, 0.5^2), 0, \alpha\rangle$, and vary the other $S_2 = \langle(\mu_2, \sigma_2^2), 0, \alpha\rangle$ while keeping $\frac{\sigma_2}{\mu_2} \propto \frac{\sigma_1}{\mu_1}$. This setting characterizes the scenario where the products have different scales of valuations. In Figure 8, we still see that profitability of bundling decreases as the heterogeneity of products increases. To summarize, we have the following observation.

Observation 6: The heterogeneity of variance in customers' valuations reduces the profitability of bundling. Still, the profit gain increases as the network effect function becomes more convex.

7 IMPACT OF CORRELATION OF CUSTOMERS' VALUATIONS

Recall that a key factor of bundling's advantage is that it may potentially reduce the variance of customers' valuations. This is true when the services are independent, but what if such valuations are *correlated*? Correlations among the bundled items are common. In the example of Microsoft Office, it is rare that one is interested in using all software (Word, PowerPoint, Excel, Access, etc.); instead, an individual usually uses only a small subset of them. In other words, people's valuation toward these products are often negatively correlated. There is evidence on the negative correlation of valuations between Word and Excel [12]. In this section, we study the impact of the correlation of valuations, especially with the existence of network effects. Note that the impact of correlation of customers' valuations has been studied in various works [22, 26]. Our work differs from the others with regard to network effects. In addition, the impact of correlation with network effects was also studied in other works [14, 25]. Different from those works, our work generalizes their linear form of network effects, which enables the study of the impact of the "degree of network effects."

To capture the correlation of customers' valuations, we generalize the previous Gaussian valuation distribution of a single service to a multivariate Gaussian valuation distribution for multiple services. We choose the Gaussian form to better present our theoretical results. As a complement, for uniform distribution, other works [14, 25] conducted comprehensive studies about the impact of correlation on the profitability of bundling with network externalities (or network effects), although they look at a linear form of network effects. In the case of two products, it is a bivariate Gaussian valuation distribution. Formally, if $\theta \in \Theta$ is a random customer, then the intrinsic valuation of θ on the two products is a random vector $(v_1(\theta), v_2(\theta))$ following bivariate Gaussian distribution with mean vector (μ_1, μ_2) and covariance matrix $\Sigma = [\begin{smallmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{smallmatrix}]$, where $\rho \in [-1, 1]$ is the *correlation coefficient*. If we bundle these two products, then the distribution of customers' intrinsic valuation is still a Gaussian distribution

$$v_b(\theta) \triangleq v_1(\theta) + v_2(\theta) \sim \mathcal{N}\left(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2\right).$$

Hence, the cumulative distribution function of $v_b(\theta)$ is

$$H_B(x) \triangleq \Phi\left(\frac{x - (\mu_1 + \mu_2)}{\sqrt{\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2}}\right). \tag{21}$$

In this section, we study the services with the same network effects (i.e., $\alpha_1 = \alpha_2 = \alpha$), so we have $F_B(x; \delta_b) = H_B(x)\delta_b^\alpha$. Note that the marginal distribution of $v_1(\theta)$ and $v_2(\theta)$ are the same as the case without bundling. In particular, when $\rho = 0$, the valuations $v_1(\theta)$ and $v_2(\theta)$ are independent. As ρ increases from -1 to 1, the correlation varies from strictly negative to strictly positive.

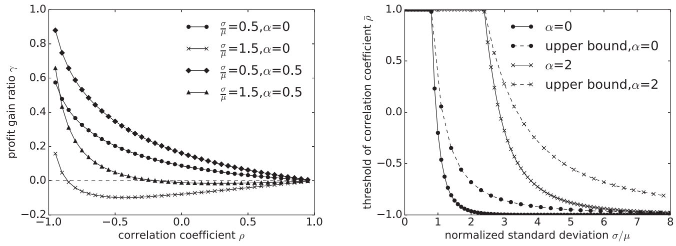
7.1 Impact of Correlation When $\sigma_1^2 = \sigma_1^2$

We study the impact of correlation, especially with the existence of network effects. Let us first consider when different services have the same variance of customers' valuations (i.e., $\sigma_1^2 = \sigma_2^2$). In Figure 9(a), we vary the correlation coefficient ρ and observe the profit gain ratio γ under different coefficient of variation σ/μ and network effects parameter α. When σ/μ is small and α is large (e.g., $\sigma/\mu = 0.5$ and $\alpha = 0.5$), γ monotonically decreases as ρ increases. The reason is that a stronger negative correlation of customers' valuations for the products indicates a larger reduction of variance of customers' valuations, and hence a higher profit. When σ/μ is large and α is small (e.g., $\sigma/\mu = 1.5$ and $\alpha = 0$), as ρ increases from -1 to 1, the profit gain ratio γ first decreases and then increases to zero. In addition, when σ/μ is large and α is small, bundling is more profitable than separate sales only when the correlation of different products is negative enough. We have observed that the impact of correlation on the profit of bundling is related to both the network effects and the coefficient of variation. We formally quantify the impact of correlation with the existence of network effects using the following theorem.

Theorem 7.1. *Consider two dependent services $S_1 = \langle(\mu, \sigma^2), 0, \alpha\rangle$ and $S_2 = \langle(\mu, \sigma^2), 0, \alpha\rangle$, where ρ is the correlation coefficient and $\alpha \geq 0$. Let $\gamma(\rho)$ be the profit gain ratio of the bundle with respect to ρ, and $\bar{\rho} = \sup_{\rho\in[0,1]}\{\rho|\gamma(\rho) > 0\}$. Then we have*

(1) $\gamma(\rho) > 0$ if and only if $\rho \in [-1, \bar{\rho})$;.
(2) $\gamma(\rho)$ decreases as ρ increases when $\rho \in [-1, \bar{\rho})$.
(3) If $\sqrt{2/\pi}(\alpha + 1) > \sigma/\mu$, then $\bar{\rho} = 1$.
(4) If $\sqrt{2/\pi}(\alpha + 1) \leq \sigma/\mu$, then $\bar{\rho} \leq \frac{4(\alpha+1)^2}{\pi(\sigma/\mu)^2} - 1 \leq 1$.

Proof. Please refer to the appendix. □



(a) Impact of correlation coefficient ρ (b) Correlation coefficient threshold $\bar{\rho}$

Fig. 9. Impact of correlation, homogeneous case, $\sigma_1 = \sigma_2, \mu_1 = \mu_2$.

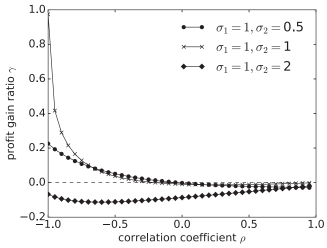


Fig. 10. Impact of correlation, heterogeneous case, $\sigma_1 \neq \sigma_2, \mu_1 = \mu_2 = 1$.

This theorem has the following physical meanings. First, there exists a unique threshold $\bar{\rho}$ on the correlation coefficient to differentiate whether bundling is either more or less profitable than separate sales. Second, in the profitable region $\rho \in [-1, \bar{\rho})$, the profit gain ratio is monotonically decreasing as correlation coefficient increases. This is consistent with the phenomena we observed in Figure 9(a). Third, for a fixed coefficient of variation, such a threshold is determined by the convexity of the network effect function represented by α. On the one hand, when the network effect function is highly convex (i.e., α is large), we have $\bar{\rho} = 1$, which means bundling is always more profitable for any correlation. On the other hand, when the network effect parameter α is small, bundling is more profitable only if the correlation is sufficient negative.

Let us use some numerical examples to further illustrate the impact of correlation. In Figure 9(b), we show how the variance of valuation affects the threshold of correlation coefficient. We vary the normalized standard deviation $\sigma/\mu \in [0, 8]$, and plot both $\bar{\rho}$ and the upper bound of $\bar{\rho}$, which is $\frac{4(\alpha+1)^2}{\pi(\sigma/\mu)^2} - 1$. When $\sigma/\mu < \sqrt{2/\pi}(\alpha + 1)$, $\bar{\rho} = 1$, as indicated by our theorem. When $\sigma/\mu \geq \sqrt{2/\pi}(\alpha + 1)$, we can see that the $\bar{\rho}$ is smaller than its upper bound. Moreover, comparing the cases $\alpha = 0$ and $\alpha = 2$, we could see the threshold $\bar{\rho}$ is larger when the network effect function is more convex. This is in accordance with our bound in Theorem 7.1. To summarize, we have the following observation.

Observation 7: The correlation of customers' valuation on different services is an important factor to determine how large the profit gain of bundling is. Bundling is more profitable when the coefficient of correlation is smaller than a certain threshold. The threshold value is allowed to be higher when the network effect function is more convex.

7.2 Impact of Correlation Coefficient When $\sigma_1^2 \neq \sigma_2^2$

We also consider the impact of correlation when two services have heterogeneous variances of customers' valuations. In Figure 10, we set $\mu_1 = \mu_2 = 1$ and show the impact of correlation coefficient ρ on the profit gain ratio γ for different values of σ_1 and σ_2. The profit gain first decreases and then increases as the correlation coefficient increases, which is similar to our observation in Figure 9(a). The difference for the case of heterogeneous variances is that bundling tends to be less profitable even if there is a strictly negative correlation. For example, in Figure 10, when $\sigma_1 = 1$ and $\sigma_2 = 2$, $\gamma < 0$ even if $\rho = -1$. This is because the profit gain of bundling reduces to negative when the variances of customers' valuations become heterogeneous, as discussed in Section 6.3, and a negative correlation in valuations is not able to alter this result.

Observation 8: When two services have different variances in customers' valuations, the impact of correlation is similar to the homogeneous setting. However, the impact of heterogeneous variances might be significant, making bundling less profitable even if there is a strictly negative correlation.

8 RELATED WORK

Bundling strategy has been discussed in economic community. Adams and Yellen [1] took the first step in analyzing the commodity bundling in a monopoly market, and use graphical and numerical analysis to show the reason for bundling and its inefficiency in social welfare. Later works [22] and Guiltinan [15] discussed *mixed bundling* and provided a framework for the firm to choose the appropriate bundling strategy. Matutes and Regibeau [21] extended the results from monopoly market to duopoly. Schmalensee [26] used a Gaussian distribution on users' reservation price and compared various bundling and separate sale using graphical illustration and numerical results. Bakos and Brynjolfsson [2] discussed bundling strategies for digital goods where unit cost could be ignored. Later works were extended to competitive markets under non-zero [19] or zero [3] unit cost. In the area of networking, Niyato et al. [23] studied bundling strategies for smart data pricing of IoT technologies. The rationale for bundling to be more profitable is that customers' valuations become more concentrated for the bundle so that the seller's pricing strategy can be more efficient. The concentration property for a sum of independent distributions is also applied in inventory/risk pooling [5, 10, 13], where the aggregation of demands across locations allows a reduced inventory. For inventory/risk pooling, the negative correlation among items are known to increase the value of pooling, which is similar to the its impact on the profitability of bundling.

Various factors that determine the profitability of bundling are studied. First, the impact of marginal costs was investigated. Bakos and Brynjolfsson [2] found out that bundling is profitable when marginal cost is small. Fang and Norman [11] found that "the higher the marginal cost and the lower the mean valuation, the less likely that bundling dominates separate sales," which was also stated by Schmalensee [26]. Second, the distribution of valuations is also an important factor. Most of the studies of bundling focused on the uniform or Gaussian distributions [2, 14, 21, 26, 27]. In the case of thin-tailed valuation distributions, Bakos and Brynjolfsson [2] showed that bundling any greater number of goods will further increase the seller's profits. Other works considering thin-tailed distributions [8, 11, 16] also showed that bundling can be more profitable. However, under heavy-tailed valuations, Ibragimov and Walden [17] pointed out that separate sale can be more profitable than bundling sale. Since risk pooling could be regarded as a form of bundling, it is not surprising to see that risk pooling has lower benefits for heavy-tailed distributions, as shown in the work of Bimpikis and Markakis [5]. Third, the impact of correlation on the profitability of bundling was also studied by many previous works by Schmalensee [26], McAfee et al. [22], and Bakos and Brynjolfsson [2]. They showed that bundling is profitable when the valuations of bundled items are independent or negatively correlated. Banciu and Ødegaard [4] pointed out that if customers' valuations are positively correlated, the profitability of bundling can be arbitrarily low. The impact of other factors, such as the complementarity of valuations on different products [27], were also analyzed. All the preceding works did not consider the network effects.

Network effects [9] (or network externalities) have also been discussed by researchers. In particular, the work of Katz and Shapiro [18] might be the earliest influential work that defined and discussed the impacts of network effects. Liebowitz and Margolis [20] concluded that direct and indirect network effects may result in market failure. Candogan et al. [7] discussed optimal pricing on divisible goods with positive network effects. There are a number of works on specific application analysis. For example, Wang et al. [28] discussed the impact of network effects in instant messaging service; in another work, Wang et al. [29] focused on network externality in mobile telecommunication innovation.

Although network effects and bundling sale have been both extensively studied, there are very few works that combine them. We find two recent works [24, 25] closely related to ours. Prasad et al. [24] discussed bundling strategy of technological products with network externality. Their work presented interesting findings, but they mainly relied on numerical and graphical explanations, and their analysis was restricted to some special cases. Meanwhile, the linear and additive form of network externality applied in their work is in a special form that does not capture all important features of online services. Hence, we need a more accurate model on the network effects to capture today's online market. In the work of Guérin et al. [25] and the follow-up journal version [14], the authors provided insightful results exploring how correlation in customers' valuation affects the profitability of bundling with network effects. For uniform distribution, they described the equilibrium by equations for general correlations, and they also considered customers' extra utilities from the network effects. In their work, this customers' extra utilities was implicitly assumed to be linear with respect to the fraction of users who have adopted the service. In our model of network effects, the degree of network externalities is captured by the parameter α that allows us to study the impact of various degrees of network externalities. More importantly, their results revealed practical insights by thorough discussions and graphical explanations, whereas our work emphasizes on the analytical framework.

Our work differs from previous works in that (1) we build a formal optimization framework that captures the optimal separate sale and bundling strategies; (2) we give rigorous analytical results based on the multiplication form of network effects; (3) with the existence of network effects, we characterize various important factors, including the network effect itself, operating cost, the variance and correlation of costumers' valuations; and (4) we further analytically show how these factors impact the profitability of bundling.

9 CONCLUSION

In this article, we discuss the bundling sale strategy for online service markets that exhibit network effects. In such a market, a customer's purchasing decision is influenced by other customers' purchasing decisions. We formulate a formal optimization framework to characterize the optimal sale strategies, which allows the service providers to determine their best sale strategies. Based on this, we analyze and quantify the impact of the key factors. Our important findings include the following. First, when the network effect function is more convex, the profit gain of bundling over separate sales becomes larger. Second, the operating cost usually plays a negative role toward bundling, but when the two services have different network effects, a moderate operating cost on a particular service may increase the profitability of bundling. Third, the variance of customers' valuation is significant in determining whether bundling is more profitable than separate sales. In particular, bundling is less profitable when the variance is larger than some threshold determined by the degree of network effects. Fourth, the correlation of customers' valuation determines the extent to which bundling can be more profitable than separate sales; a negative correlation indicates a larger profit gain, provided that bundling is more profitable than separate sales. Fifth, the asymmetry in operating costs, and the heterogeneity in valuation distributions or network effects, reduces the profitability of bundling and can make bundling even less profitable than separate sales. We believe that these findings provide valuable insights for online service providers to design effective pricing schemes, and we plan to better explore bundling sales via real data analytics.

Here we also state the limitations of our work and hope that this serves as the basis for future work. Due to the complex nature of bundling sale, our modeling approach can not cover all aspects and has several limitations. First, the complementarity/substitutability of the bundled products represented by the parameter Δv is assumed to be homogeneous for every customer. Second, our results do not cover *mixed bundling*, where the provider allows the customers to either buy products separately or to buy the bundle. Third, our model does not consider the competitions among different providers. Last but not least, this work only considers some specific forms of distributions of customers' valuations (e.g., uniform and Gaussian) to make the analysis tractable.

APPENDIX

Proof of Lemma 2.4. Since for any $\delta_i > 0$ we have $\rho_i(\delta_i) > 0$, so the condition $v_i(\theta)\rho_i(\delta_i) \geq p_i$ is equivalent to $v_i(\theta) \geq \frac{p_i}{\rho_i(\delta_i)}$. According to the definition of equilibrium, we have $\delta_i = \int_{\theta\in\Theta} \mathbf{1}_{\{v_i(\theta)\geq \frac{p_i}{\rho_i(\delta_i)}\}} f(\theta)d\theta$. By noting the preceding equation and recalling the definition of $H_i(x)$ in Equation (2), we can prove the separate sale part of the lemma.

Now we discuss the equilibria for the bundling case. Recall that in (6) we have defined $F_b(\cdot, \delta_b)$ as the cumulative distribution function of customers' valuation of the bundle when the fraction of adopters of the bundle is δ_b. When the price of the bundle is p_b, in the equilibrium the following equation holds:

$$F_b(p_b; \delta_b) = 1 - \delta_b. \tag{22}$$

It means that the fraction of customers whose valuations are less than the price p_b is $1 - \delta_b$, or in other words the fraction of customers whose valuations are higher than the price p_b is δ_b. □

Proof of Theorem 3.2. We start with the case of separate sale. According to Corollary 2.5, δ_i is an equilibria if and only if it satisfies

$$p_i = \delta_i^{\alpha_i} H_i^{-1}(1 - \delta_i), \ (i = 1, 2),$$

where $H_i^{-1}(1 - \delta_i) = V_i(1 - \delta_i)$.

Let $g_i(x) = x^{\alpha_i} H_i^{-1}(1 - x)$, $i = 1, 2$. Then the equilibrium δ_i is a solution to

$$p_i = g_i(\delta_i). \tag{23}$$

By letting $g_i'(x) = 0$, $i = 1, 2$, we see there is a unique solution $x_i^* = \frac{\alpha_i}{\alpha_i+1}$ in $(0, 1]$.

Note that for $i \in \{1, 2\}$, $g_i(0) = 0$, $g_i(1) = 0$, $g_i(x) > 0$ if $0 < x < 1$, and that $g_i'(x) = 0$ has at most one solution in (0,1], we conclude that $g_i(x)$ has one and only one maxima when $x \in (0, 1)$. Let us denote this value as $\bar{p_i}$. Thus, when x increases from 0 to 1, $g_i(x)$ first increases from 0 to $\bar{p_i}$ and then decreases from $\bar{p_i}$ to 0. Therefore, when $p_i > \bar{p_i}$, Equation (23) has no solution; when $p_i = \bar{p_i}$ or 0, Equation (23) has only one solution; when $0 < p_i < \bar{p_i}$, Equation (23) has two solutions. In particular, we have $\bar{p_i} = g_i(\frac{\alpha_i}{\alpha_i+1}) = \frac{V_i}{\alpha_i+1}(\frac{\alpha_i}{\alpha_i+1})^{\alpha_i}$ for $i = 1, 2$, which completes the proof for the separate sale.

According to Equation (22), we know in the equilibrium that we have

$$p_b = F_b^{-1}(1 - \delta_b; \delta_b).$$

We observe that p_b is equal to a well-defined function of δ_b denoted as $P_b(\delta_b) \triangleq F_b^{-1}(1-\delta_b;\delta_b)$, where $\delta_b \in [0,1]$. Moreover, the function $P_b(\delta_b)$ has the following form:

$$P_b(\delta_b) = p_b = \begin{cases} \sqrt{2(1-\delta_b)(V_1\delta_b^{\alpha_1}V_2\delta_b^{\alpha_2})} & \text{if } 0 \le p_b \le V_1\delta_b^{\alpha_1}, \\ (1-\delta_b)V_2\delta_b^{\alpha_2} + V_1\delta_b^{\alpha_1}/2 & \text{if } V_1\delta_b^{\alpha_1} < p_b \le V_2\delta_b^{\alpha_2}, \\ V_1\delta_b^{\alpha_1} + V_2\delta_b^{\alpha_2} - \sqrt{2V_1V_2\delta_b^{(1+\alpha_1+\alpha_2)}} & \text{if } V_2\delta_b^{\alpha_2} < p_b \le V_1\delta_b^{\alpha_1} + V_2\delta_b^{\alpha_2}. \end{cases} \quad (24)$$

Let us provide some properties of the function $P_b(\delta_b)$. We first claim that $P_b(\delta_b)$ is a continuous function. In fact, the cumulative function $F_b(x;\delta_b)$ is a continuous function with respect to x and δ_b. Therefore, the inverse function $F_b^{-1}(y;\delta_b)$ is continuous with respect to y and δ_b. Since $1-\delta_b$ is continuous with respect to δ_b, the composition function $F_b^{-1}(1-\delta_b;\delta_b)$ is also continuous with respect to δ, which is our function $P_b(\delta_b)$. Second, one can verify that $P_b(0) = 0$. In fact, when $\delta_b = 0$, $P_b(\delta_b) = 0$ for all three equations in (24). This is because when $\delta_b = 0$, the valuations of customers are 0 and the only possible price should be 0. Third, we have $P_b(1) = 0$ from (24). The reason is that if we want all customers to be willing to adopt the service, the price should be as low as zero.

The continuous function $P_b(\delta_b)$ defined in the closed interval $\delta_b \in [0,1]$ has its maximal value. Let $\bar{p}_b = \max_{\delta_b\in[0,1]} P_b(\delta_b)$ be this maximal value, let and $\bar{\delta}_b = \arg\max_{\delta_b\in[0,1]} P_b(\delta_b)$ be the maximizer. Recall that $P_b(\delta_b) = p_b$ in the equilibrium. First, when $p_b > \bar{p}_b$, the formula $P_b(\delta_b) = p_b$ has no solution, and therefore there is no equilibrium. Second, when $p_b = \bar{p}_b$, the formula $P_b(\delta_b) = p_b$ has at least one solution (i.e., $\delta_b = \bar{\delta}_b$). Hence, there is at least one equilibrium. Third, when $0 \le p_b < \bar{p}_b$, there exists $\delta_b^1 \in [0,\bar{\delta}_b)$ such that $P_b(\delta_b^1) = p_b$, and there exists $\delta_b^2 \in (\bar{\delta}_b, 1]$ such that $P_b(\delta_b^2) = p_b$ by the intermediate value theorem for continuous functions. We have found two solutions $\delta_b^1 \neq \delta_b^2$ of the formula, and therefore the market has at least two equilibria. Note that for a given price of the bundle, there can be multiple equilibria for the fraction of adopters, but if we observe a certain fraction of adopters, we can derive the unique price of the bundle. □

Explanation on the stability of the largest equilibrium for bundling. Suppose that δ_b^* is the largest equilibrium for the bundling sale that satisfies $P_b(\delta_b^*) = p_b$. We claim that for any $\delta_b' \in (\delta_b^*, 1]$, $P_b(\delta_b') < p_b$. Otherwise, if there is any $\delta_b' > \delta_b^*$ such that $P_b(\delta_b') \ge p_b$, then we can find some $\delta_b'' \in [\delta_b', 1]$ such that $P_b(\delta_b'') = p_b$. In other words, δ_b'' is a larger equilibrium than "the largest equilibrium" δ_b^*, which leads to a contradiction. Moreover, one can verify that $P_b(\delta_b)$ defined in (24) is a differentiable function. The derivative $P_b'(\delta_b^*) = \lim_{\epsilon\to 0+} \frac{P_b(\delta_b^*+\epsilon)-P_b(\delta_b^*)}{\epsilon} \le 0$, which indicates the negative slope of the function at δ_b^*. For the case where $P_b'(\delta_b^*) = 0$, the point δ_b^* is the local maximum and mathematically can be an "unstable" equilibrium. But the local maximum of $P_b'(\delta_b^*) = 0$ is not likely to be the maximal solution of (15), which means that the case $P_b'(\delta_b^*) = 0$ rarely happens when the bundling strategy is selected by our optimization framework. In addition, in practice, the seller can always do some small adjustment to avoid such a "local maximum" case and the adjustment has little impact on the profit. Due to these reasons, we then focus on the case $P_b'(\delta_b^*) < 0$.

On the one hand, if the fraction of adopters increases to $\delta_b' \in [\delta_b^*, \delta_b^*+\epsilon)$ (ϵ is small), then we have $p_b' = P_b(\delta_b') < p_b$, where $1 - F_{\delta_b'}(p_b') = \delta_b'$. We can see that for $p_b > p_b'$, $1 - F_{\delta_b'}(p_b) < \delta_b'$, which means that the fraction of adopters when the price is still p_b will be less than δ_b'. Namely, the market will move down toward the equilibrium δ_b^* if the market is perturbed to have a larger fraction $\delta_b' \in [\delta_b^*, \delta_b^*+\epsilon)$ of adopters. On the other hand, if the fraction of adopters decreases to $\delta_b' \in (\delta_b^*-\epsilon, \delta_b^*]$, then we have $p_b' = P_b(\delta_b') > p_b$. Similarly, the market will move up to the equilibrium δ_b^* if the market is perturbed to have a smaller fraction $\delta_b' \in (\delta_b^*-\epsilon, \delta_b^*]$ of adopters. Combining these two scenarios, we have shown that the largest equilibrium δ_b^* is stable.

Proof of Theorem 4.1. We first prove the second part of the theorem. By applying the optimization framework, we see the optimal separate sale is a solution to

$$\begin{aligned} \max_{\delta_i} \quad & U_i(\delta_i) = \sum_{i=1}^{2} \delta_i \left(\delta_i^{\alpha} - \delta_i^{\alpha+1}\right), \\ \text{subject to} \quad & 0 \le \delta_i \le 1. \end{aligned} \tag{25}$$

We can easily obtain the solution as $\delta_i^* = \frac{\alpha+1}{\alpha+2}$ and have $p_i^* = \frac{V}{\alpha+2}(\frac{\alpha+1}{\alpha+2})^{\alpha}, U_s^* = 2p_i^*\delta_i^*$. Similarly, we can solve the optimal bundling sale. By noting the definition of γ and taking the forms of U_s^* and U_b^* into Equation (16), we can derive the profit gain ratio $\gamma(\alpha)$ as desired. By taking the derivative of $\log\gamma(\alpha)$ with respect to α, it is easy to show that $\gamma(\alpha)$ is increasing in α.

Now we prove the first part of the theorem. For any given α, we have $\delta_1^* = \delta_2^* = \frac{\alpha+1}{\alpha+2}$, and we denote this value as δ^*. Obviously, $\delta^* > 1/2$, and we have

$$U_s^* = \delta^*\rho_1(\delta^*)H_1^{-1}(1-\delta^*) + \delta^*\rho_2(\delta^*)H_2^{-1}(1-\delta^*). \tag{26}$$

We consider a bundling sale with price p_b such that the largest equilibrium is δ^*. For the case where $\alpha_1 = \alpha_2 = \alpha$, we define

$$H_B(x) \triangleq \begin{cases} 0 & \text{if } x < 0, \\ x^2/(2V_1V_2) & \text{if } 0 \le x \le V_1, \\ (2x - V_1)/(2V_2) & \text{if } V_1 < x \le V_2, \\ 1-(V_1+V_2-x)^2/(2V_1V_2) & \text{if } V_2 < x \le V_1+V_2, \\ 1 & \text{if } x > V_1 + V_2. \end{cases} \tag{27}$$

One can verify that $F_B(x;\delta_b) = H_B(x)\delta_b^{\alpha}$ for any x, δ_b. The service provider's utility in the optimal bundling sale satisfies

$$U_b^* \ge \delta^*\rho_b(\delta^*)H_B^{-1}(1-\delta^*). \tag{28}$$

Given the form of $H_i(x)$ and $H_B(x)$, one can easily verify that $\frac{H_i^{-1}(1-\delta^*)}{V_i} < \frac{H_B^{-1}(1-\delta^*)}{V_1+V_2}$ for $\delta^* > 1/2$. Therefore, we have $H_1^{-1}(1-\delta^*) + H_2^{-1}(1-\delta^*) < H_B^{-1}(1-\delta^*)$. Since $\rho_1(\delta) = \rho_2(\delta) = \rho_b(\delta) = \delta^{\alpha}$, we have

$$\sum_{i=1}^{2} \delta^*\rho_i(\delta^*)H_i^{-1}(1-\delta^*) < \delta^*\rho_b(\delta^*)H_B^{-1}(1-\delta^*). \tag{29}$$

Combining inequalities (26), (28), and (29), we have $U_b^* > U_s^*$ and therefore the profit gain ratio $\gamma > 0$. □

Proof of Theorem 5.1: Let us denote $\frac{m_1}{V_1} = \frac{m_2}{V_2} = m$. According to Lemma 3.4, the equilibria of separate sales satisfy $\delta_i^* = \delta_2^*$ and we denote it as δ^*. Based on the optimization formulation, we know that δ^* is a solution to the following optimization:

$$\begin{aligned} \max_{\delta} \quad & U(\delta) = \delta^{\alpha+1} - \delta^{\alpha+2} - m\delta, \\ \text{subject to} \quad & 0 \le \delta \le 1. \end{aligned} \tag{30}$$

We next show $\delta^* = 0$ or $\delta^* > 1/2$. Note that $U(\delta) = \delta g(\delta)$, where $g(\delta) = \delta^{\alpha} - \delta^{\alpha+1} - m$. Let us consider that if $\delta^* = 0$ is not the unique solution, then $g(\delta^*) \ge 0$. By taking first- and second-order derivatives, we can derive that $g(\delta)$ achieves the unique maximal value in $[0,1]$ when $\delta = \frac{\alpha}{\alpha+1}$. Let us suppose that $\delta^* < \frac{\alpha}{\alpha+1}$, then $0 \le g(\delta^*) < g(\frac{\alpha}{\alpha+1})$, so $\delta^* g(\delta^*) < \frac{\alpha}{\alpha+1}g(\frac{\alpha}{\alpha+1})$, which is $U(\delta^*) < U(\frac{\alpha}{\alpha+1})$. This contradicts our assumption that $U(\delta^*)$ is the maximal value in [0,1]. Therefore, we have $\delta^* = 0$ or $\delta^* \ge \frac{\alpha}{\alpha+1} > 1/2$.

If $\delta^* = 0$, then the optimal separate sale achieves a profit of zero, which is obviously no larger than the optimal bundling sale; if $\delta^* > 1/2$, then using the same approach in the proof of Theorem 4.1, we can prove that the optimal bundling sale is more profitable than the optimal separate sale. Combining the preceding two cases, we prove the theorem. □

Proof of Theorem 5.2. We first analyze the bundling sale with price p_{bi} such that the equilibrium $\delta_{bi} = \delta_i^*, i = 1, 2$. Given the forms of $H_i(\cdot)$ and $H_B(\cdot)$ where H_B is defined in (27) and $F_B(x; \delta_b) = H_B(x)\delta_b^\alpha$), we have $\frac{H_i^{-1}(1-\delta_i^*)}{V_i} < \frac{H_B^{-1}(1-\delta_i^*)}{V_1+V_2}$. Since $p_{bi} = \delta_i^{*\alpha} H_B^{-1}(1-\delta_i^*)$ and $p_i^* = \delta_i^{*\alpha} H_1^{-1}(1-\delta_i^*)$, we have $p_{bi} > p_i^* \frac{V_1+V_2}{V_i}$, so the service provider's utility under this setting satisfies

$$U_{bi} > (p_i^*(V_1+V_2)/V_i - \beta_1 m_1 - \beta_2 m_2)\, \delta_i^*. \tag{31}$$

Since the preceding settings $(\delta_{b1}, \delta_{b2})$ are two realizations in the bundling strategy, we have that the optimal bundling utility satisfies

$$2U_b^* \geq U_{b1} + U_{b2} > p_1^*\delta_1^*(V_1+V_2)/V_1 + p_2^*\delta_2^*(V_1+V_2)/V_2 - \beta_1(\delta_1^*+\delta_2^*)m_1 - \beta_2(\delta_1^*+\delta_2^*)m_2. \tag{32}$$

Therefore, we have

$$\begin{aligned} 2(U_b^* - U_s^*) > &\,(V_2 - V_1)\,(p_1^*\delta_1^*/V_1 - p_2^*\delta_2^*/V_2) \\ &-((\beta_1-2)\delta_1^* + \beta_1\delta_2^*)m_1 - ((\beta_2-2)\delta_2^* + \beta_2\delta_1^*)m_2. \end{aligned} \tag{33}$$

Let us consider another service $S_0 = \langle 1, 0, \alpha_i \rangle$, and assume that its optimal equilibrium is δ_0^*. Since the increase of unit cost reduces the value of the optimal equilibrium, and that $\delta_2^* \leq \delta_1^*$, we have $\delta_0^* \geq \delta_1^* \geq \delta_2^*$. Since δ_1^*, and δ_2^* are also two sale strategies of S_0, and that δ_1^* is nearer to the optimal separate sale, we have $\frac{p_1^*\delta_1^*}{V_1} \geq \frac{p_2^*\delta_2^*}{V_2}$. Recalling $V_2 \geq V_1$, we have

$$(V_2 - V_1)\,(p_1^*\delta_1^*/V_1 - p_2^*\delta_2^*/V_2) \geq 0. \tag{34}$$

Since $\beta_1 \leq 1, \beta_2 \leq 1, \delta_1^* \geq \delta_2^*$, we have $((\beta_1-2)\delta_1^* + \beta_1\delta_2^*)m_1 \leq 0$. Since $\frac{2-\beta_2}{\beta_2}\delta_1^* \leq \delta_2^*$, we have $((\beta_2-2)\delta_2^* + \beta_2\delta_1^*)m_2 \leq 0$. Therefore, we have

$$((\beta_1-2)\delta_1^* + \beta_1\delta_2^*)m_1 + ((\beta_2-2)\delta_2^* + \beta_2\delta_1^*)m_2 \leq 0. \tag{35}$$

Combining (33), (34), and (35), we conclude that $U_b^* - U_s^* > 0$. □

Proof of Theorem 6.2. Let $U(\delta, \langle(\mu,\sigma^2), c, \alpha\rangle) \triangleq [\delta^\alpha(\Phi^{-1}(1-\delta)\sigma + \mu) - c] \times \delta$ be the profit when δ is the equilibria for service $\langle(\mu,\sigma^2), c, \alpha\rangle$, and let

$$U^*(\langle(\mu,\sigma^2), c, \alpha\rangle) \triangleq \max_{\delta\in[0,1]} [\delta^\alpha(\Phi^{-1}(1-\delta)\sigma + \mu) - c] \times \delta$$

be the optimal profit for service $\langle(\mu,\sigma^2), c, \alpha\rangle$. The equilibrium achieving this optimal profit is

$$\delta^*(\langle(\mu,\sigma^2), c, \alpha\rangle) \triangleq \arg\max_{\delta\in[0,1]} U(\delta, \langle(\mu,\sigma^2), c, \alpha\rangle).$$

Let us first prove the following three lemmas. □

Lemma A.1. $\delta^* \frac{1}{\phi(\Phi^{-1}(1-\delta^*))} - (\alpha+1)\Phi^{-1}(1-\delta^*)$ *is a strictly monotonically increasing function of* δ^*.

Proof. Consider the function $g(t) = \frac{1-\Phi(t)}{\phi(t)} - (\alpha+1)t$, where $\phi(t) = \Phi'(t)$. We have $g'(t) = \frac{t(1-\Phi(t))}{\phi(t)} - (\alpha+2)$. When $t \leq 0$, obviously $g'(t) < 0$. When $t > 0$,

$$g'(t) = \frac{t\int_t^{+\infty}\phi(x)}{\phi(t)} - (\alpha+2) \leq \frac{\int_t^{+\infty} x\phi(x)}{\phi(t)} - (\alpha+2) = \frac{-\phi(+\infty)+\phi(t)}{\phi(t)} - (\alpha+2) = -(\alpha+1) < 0. \tag{36}$$

Hence, $g(\cdot)$ is monotonically strictly decreasing in $(-\infty, +\infty)$.

Let $t(\delta^*) = \Phi^{-1}(1-\delta^*)$ be a strictly monotonically decreasing function of δ^*. Then the following function of δ^*,

$$g(t(\delta^*)) = g(\Phi^{-1}(1-\delta^*)) = \delta^* \frac{1}{\phi(\Phi^{-1}(1-\delta^*))} - (\alpha+1)\Phi^{-1}(1-\delta^*),$$

is strictly monotonically increasing with respect to δ^*, which concludes the lemma. □

Lemma A.2. *For any given $\alpha \geq 0$, $\delta^*(\langle(\mu,\sigma^2), 0, \alpha)$ is a strictly decreasing function with respect to $\frac{\sigma}{\mu}$.*

Proof. To simplify the notation, we denote

$$\delta^* \triangleq \delta^*(\langle(\mu,\sigma^2), 0, \alpha\rangle).$$

Obviously, $U^*(\langle(\mu,\sigma^2), 0, \alpha\rangle) > 0$ (in fact, letting $\delta_0 = 1 - \Phi(\frac{1-\mu}{\sigma})$, we have $U(\delta_0, \langle(\mu,\sigma^2), 0, \alpha\rangle) = \delta_0^{\alpha+1} > 0$). Note that when $\delta = 0$, $U(\delta, \langle(\mu,\sigma^2), 0, \alpha\rangle) = 0$. When $\delta = 1$, $U(\delta, \langle(\mu,\sigma^2), 0, \alpha\rangle) = -\infty$. Therefore, $\delta^* \neq 0$ and $\delta^* \neq 1$, so $\delta^* \in (0,1)$. Write down the optimality condition for the following optimization problem:

$$\delta^* = \arg\max_{\delta\in(0,1)} U(\delta, \langle(\mu,\sigma^2), 0, \alpha\rangle).$$

A necessary and sufficient condition for δ^* is

$$0 = U'(\delta^*, \langle(\mu,\sigma^2), 0, \alpha\rangle) = (\alpha+1)(\delta^*)^{\alpha}[\Phi^{-1}(1-\delta^*)\sigma + \mu] - (\delta^*)^{\alpha+1}\left[\frac{1}{\phi(\Phi^{-1}(1-\delta^*))}\sigma\right].$$

Then we have

$$\frac{\sigma}{\mu} = \frac{\alpha+1}{\delta^* \frac{1}{\phi(\Phi^{-1}(1-\delta^*))} - (\alpha+1)\Phi^{-1}(1-\delta^*)} \geq 0. \tag{37}$$

We can see that $\frac{\sigma}{\mu}$ is represented as a function of δ^*. From Lemma A.1, we know that $\delta^* \frac{1}{\phi(\Phi^{-1}(1-\delta^*))} - (\alpha+1)\Phi^{-1}(1-\delta^*)$ as the denominator is strictly monotonically increasing in δ^*.

Let us regard δ^* as the hidden function of $\frac{\sigma}{\mu}$. From (37), we know that $\frac{\sigma}{\mu}$ is strictly monotonically decreasing in δ^* for all δ^* in the natural domain, which satisfies

$$\delta^* \frac{1}{\phi(\Phi^{-1}(1-\delta^*))} - (\alpha+1)\Phi^{-1}(1-\delta^*) > 0.$$

Reversely, δ^* is also monotonically decreasing in $\frac{\sigma}{\mu}$, which concludes the lemma. □

Lemma A.3. *In the region $\sigma \in [0, \mu\sqrt{2/\pi}(\alpha+1)]$, $U^*(\langle(\mu,\sigma), 0, \alpha\rangle)$ is a strictly monotonically decreasing function of σ. In the region $\sigma \in [\mu\sqrt{2/\pi}(\alpha+1), +\infty)$, $U^*(\langle(\mu,\sigma), 0, \alpha\rangle)$ is a strictly monotonically increasing function of σ.*

Proof. Let $\delta^* = \frac{1}{2}$ in Equation (37), then $\frac{\sigma}{\mu} = \sqrt{\frac{2}{\pi}}(\alpha+1)$. Because of the monotonicity in Lemma A.2, we know that when $\sigma < \mu\sqrt{\frac{2}{\pi}}(\alpha+1)$, $\delta^* > \frac{1}{2}$. Similarly, when $\sigma > \mu\sqrt{\frac{2}{\pi}}(\alpha+1)$, $\delta^* < \frac{1}{2}$.

Let $\delta_1^* = \delta^*(\langle(\mu,\sigma_1^2), 0, \alpha\rangle)$ and $\delta_2^* = \delta^*(\langle(\mu,\sigma_2^2), 0, \alpha\rangle)$.

First, suppose that $\sigma_1 < \sigma_2 < \mu\sqrt{2/\pi}(\alpha+1)$. According to the strict monotonicity in Lemma A.2, $\delta_1^* > \delta_2^* > \frac{1}{2}$. Then we have $\Phi^{-1}(1-\delta_2^*) < 0$. Hence,

$$\begin{aligned} U^*(\langle(\mu,\sigma_2^2), 0, \alpha\rangle) &= U(\delta_2^*, \langle(\mu,\sigma_2^2), 0, \alpha\rangle) = (\delta_2^*)^{\alpha+1}(\Phi^{-1}(1-\delta_2^*)\sigma_2+\mu) < (\delta_2^*)^{\alpha+1}(\Phi^{-1}(1-\delta_2^*)\sigma_1+\mu) \\ &= U(\delta_2^*, \langle(\mu,\sigma_1^2), 0, \alpha\rangle) \leq U(\delta_1^*, \langle(\mu,\sigma_1^2), 0, \alpha\rangle) = U^*(\langle(\mu,\sigma_1^2), 0, \alpha\rangle). \end{aligned} \tag{38}$$

We see that $U^*(\langle\langle\mu,\sigma_2^2),0,\alpha\rangle) < U^*(\langle\langle\mu,\sigma_1^2),0,\alpha\rangle)$. By definition, $U^*(\langle\langle\mu,\sigma),0,\alpha\rangle)$ is a strictly monotonically decreasing function with respect to σ in the region $\sigma \in [0, \mu \times \sqrt{2/\pi}(\alpha+1)]$.

Second, suppose that $\mu\sqrt{2/\pi}(\alpha+1) < \sigma_1 < \sigma_2$, then $\frac{1}{2} > \delta_1^* > \delta_2^*$ according to Lemma A.2. Hence, $\Phi^{-1}(1-\delta_1^*) > 0$. Similarly,

$$\begin{aligned} U^*(\langle\langle\mu,\sigma_1^2),0,\alpha\rangle) &= U(\delta_1^*,\langle(\mu,\sigma_1^2),0,\alpha\rangle) = (\delta_1^*)^{\alpha+1}(\Phi^{-1}(1-\delta_1^*)\sigma_1+\mu) < (\delta_1^*)^{\alpha+1}(\Phi^{-1}(1-\delta_1^*)\sigma_2+\mu) \\ &= U(\delta_1^*,\langle(\mu,\sigma_2^2),0,\alpha\rangle) \le U(\delta_2^*,\langle(\mu,\sigma_2^2),0,\alpha\rangle) = U^*(\langle\langle\mu,\sigma_2^2),0,\alpha\rangle). \end{aligned} \tag{39}$$

We see that $U^*(\langle\langle\mu,\sigma_2^2),0,\alpha\rangle) > U^*(\langle\langle\mu,\sigma_1^2),0,\alpha\rangle)$. By definition, $U^*(\langle\langle\mu,\sigma),0,\alpha\rangle)$ is a strictly monotonically increasing function of σ in the region $\sigma \in [\mu \times \sqrt{2/\pi}(\alpha+1), +\infty)$.

The maximal profit of separate sale is $U_s^* = 2U^*(\langle\langle\mu,\sigma^2),0,\alpha\rangle)$, and the maximal profit after bundling is $U_b^* = U^*(\langle\langle 2\mu, 2\sigma^2),0,\alpha\rangle) = 2U^*(\langle\langle\mu,(\frac{\sqrt{2}}{2}\sigma)^2,0,\alpha\rangle\rangle)$:

(1) When $\frac{\sigma}{\mu} < \sqrt{\frac{2}{\pi}}(\alpha+1)$, $\frac{\sqrt{2}}{2}\sigma < \sigma < \mu\sqrt{\frac{2}{\pi}}(\alpha+1)$. According to Lemma A.3, $U_s^* < U_b^*$, so $\gamma > 0$.

(2) When $\frac{\sigma}{\mu} > \frac{2}{\sqrt{\pi}}(\alpha+1)$, $\mu\sqrt{\frac{2}{\pi}}(\alpha+1) < \frac{\sqrt{2}}{2}\sigma < \sigma$. According to Lemma A.3, $U_b^* < U_s^*$, so $\gamma < 0$. □

Proof of Theorem 6.3. For ease of presentation, let us define the following functions:

$$h_\delta(\sigma) = U(\delta,\langle(1,\sigma^2),0,\alpha\rangle); \qquad h(\sigma) = U^*(\langle\langle 1,\sigma^2),0,\alpha\rangle) = \max_{\delta\in[0,1]} h_\delta(\sigma).$$

Note that we regard α as an constant here. Let us prove the following four lemmas first. □

Lemma A.4. *$h(\sigma)$ is a convex function with respect to σ.*

Proof. Note that $h_\delta(\sigma) = \delta^{\alpha+1}(\Phi^{-1}(1-\delta)\sigma+\mu)$ is an affine function of σ. We have

$$h_\delta(t\sigma_1+(1-t)\sigma_2) = th_\delta(\sigma_1)+(1-t)h_\delta(\sigma_2), \forall t\in[0,1]. \tag{40}$$

Due to maximum property, we have $h_\delta(\sigma_1) \le h(\sigma_1)$ and $h_\delta(\sigma_2) \le h(\sigma_2)$. Hence,

$$h_\delta(t\sigma_1+(1-t)\sigma_2) \le th(\sigma_1)+(1-t)h(\sigma_2), \forall\delta\in[0,1]. \tag{41}$$

Since $[0,1]$ is a compact set, maximum is attainable. We have

$$h(t\sigma_1+(1-t)\sigma_2) = \max_{\delta\in[0,1]} h_\delta(t\sigma_1+(1-t)\sigma_2) \le th(\sigma_1)+(1-t)h(\sigma_2). \tag{42}$$

According to definition, $h(\sigma)$ is a convex function of σ. □

Lemma A.5. *$h(\sigma+\Delta_H)+h(\sigma-\Delta_H) \ge h(\sigma+\Delta_L)+h(\sigma-\Delta_L)$.*

Proof. We can see that

$$\sigma-\Delta_L = \frac{\Delta_H+\Delta_L}{2\Delta_H}(\sigma-\Delta_H)+\frac{\Delta_H-\Delta_L}{2\Delta_H}(\sigma+\Delta_H), \quad \sigma+\Delta_L = \frac{\Delta_H-\Delta_L}{2\Delta_H}(\sigma-\Delta_H)+\frac{\Delta_H+\Delta_L}{2\Delta_H}(\sigma+\Delta_H),$$

where $\frac{\Delta_L+\Delta_H}{2\Delta_H}, \frac{\Delta_H-\Delta_L}{2\Delta_H} \in [0,1]$.

Because of the convexity of $h(\cdot)$ proven in Lemma A.4, we have

$$\begin{aligned} h(\sigma-\Delta_L) &\le \frac{\Delta_H+\Delta_L}{2\Delta_H}h(\sigma-\Delta_H)+\frac{\Delta_H-\Delta_L}{2\Delta_H}h(\sigma+\Delta_H), \\ h(\sigma+\Delta_L) &\le \frac{\Delta_H-\Delta_L}{2\Delta_H}h(\sigma-\Delta_H)+\frac{\Delta_H+\Delta_L}{2\Delta_H}h(\sigma+\Delta_H). \end{aligned} \tag{43}$$

By summing them up, we have $h(\sigma+\Delta_H)+h(\sigma-\Delta_H) \ge h(\sigma+\Delta_L)+h(\sigma-\Delta_L)$. □

Lemma A.6. *If $\gamma_H > 0$, then $\sqrt{\frac{1}{2}(\sigma^2+\Delta_H^2)} \le \sqrt{\frac{2}{\pi}}(\alpha+1)$.*

Proof. We prove it by contradiction. Let us suppose that $\sqrt{\frac{1}{2}(\sigma^2 + \Delta_H^2)} > \sqrt{\frac{2}{\pi}}(\alpha + 1)$, and we know that $\sigma \geq \sqrt{\frac{1}{2}(\sigma^2 + \Delta_H^2)}$ because $\Delta_H < \sigma$. Now we have

$$\sigma \geq \sqrt{\frac{1}{2}\left(\sigma^2 + \Delta_H^2\right)} > \sqrt{\frac{2}{\pi}}(\alpha + 1). \tag{44}$$

According to Lemma A.3,

$$h\left(\sqrt{\frac{1}{2}\left(\sigma^2 + \Delta_H^2\right)}\right) = U^*\left(\left\langle\left(1, \frac{1}{2}\left(\sigma^2 + \Delta_H^2\right)\right), 0, \alpha\right\rangle\right) \leq U^*\left(\langle(1, \sigma^2), 0, \alpha\rangle\right) = h(\sigma). \tag{45}$$

Because of convexity, we also have

$$h(\sigma + \Delta_H) + h(\sigma - \Delta_H) \geq 2h(\sigma). \tag{46}$$

Therefore, we have $h(\sigma + \Delta_H) + h(\sigma - \Delta_H) \geq 2h\left(\sqrt{\frac{1}{2}(\sigma^2 + \Delta_H^2)}\right)$ from (45) and (46). This means that bundling is less profitable than separate sale—for instance, $\gamma_H \leq 0$, which is a contradiction to $\gamma_H > 0$. This concludes our lemma. □

Lemma A.7. *If $\gamma_H > 0$ and $\Delta_L < \Delta_H$, then $h(\frac{\sqrt{2}}{2}\sqrt{\sigma^2 + \Delta_L^2}) > h(\frac{\sqrt{2}}{2}\sqrt{\sigma^2 + \Delta_H^2})$.*

Proof. Note that $\frac{\sqrt{2}}{2}\sqrt{\sigma^2 + \Delta_L^2} < \frac{\sqrt{2}}{2}\sqrt{\sigma^2 + \Delta_H^2} \leq \sigma$. According to Lemma A.6, we have

$$\frac{\sqrt{2}}{2}\sqrt{\sigma^2 + \Delta_L^2} < \frac{\sqrt{2}}{2}\sqrt{\sigma^2 + \Delta_H^2} \leq \sqrt{\frac{2}{\pi}}(\alpha + 1). \tag{47}$$

According to Lemma A.3,

$$\begin{aligned} h\left(\frac{\sqrt{2}}{2}\sqrt{\sigma^2 + \Delta_H^2}\right) &= U^*\left(\left\langle\left(1, \frac{1}{2}\left(\sigma^2 + \Delta_H^2\right)\right), 0, \alpha\right\rangle\right) < U^*\left(\left\langle\left(1, \frac{1}{2}\left(\sigma^2 + \Delta_L^2\right)\right), 0, \alpha\right\rangle\right) \\ &= h\left(\frac{\sqrt{2}}{2}\sqrt{\sigma^2 + \Delta_L^2}\right). \end{aligned} \tag{48}$$

Here we have $h(\frac{\sqrt{2}}{2}\sqrt{\sigma^2 + \Delta_L^2}) > h(\frac{\sqrt{2}}{2}\sqrt{\sigma^2 + \Delta_H^2})$, which concludes our lemma. □

Now, let us write down the profit gain ratios:

$$\gamma_L = \frac{2 \times h(\frac{\sqrt{2}}{2}\sqrt{\sigma^2 + \Delta_L^2})}{h(\sigma + \Delta_L) + h(\sigma - \Delta_L)} - 1, \qquad \gamma_H = \frac{2 \times h(\frac{\sqrt{2}}{2}\sqrt{\sigma^2 + \Delta_H^2})}{h(\sigma + \Delta_H) + h(\sigma - \Delta_H)} - 1.$$

Thanks to the previous lemmas, we are able to compare γ_L and γ_H. In particular, we know that the numerator $h(\frac{\sqrt{2}}{2}\sqrt{\sigma^2 + \Delta_L^2}) > h(\frac{\sqrt{2}}{2}\sqrt{\sigma^2 + \Delta_H^2})$ based on Lemma A.7, and according to Lemma A.5, the denominator $h(\sigma + \Delta_L) + h(\sigma - \Delta_L) \leq h(\sigma + \Delta_H) + h(\sigma - \Delta_H)$. As a result, $\gamma_L > \gamma_H$.

Proof of Theorem 7.1. For services $S_1 = S_2 = \langle(\mu, \sigma^2), 0, \alpha\rangle$, the bundle could be regarded as a single service $S_b = \langle(2\mu, 2(1 + \rho)\sigma^2), 0, \alpha\rangle$. With the definitions in the proof of Theorem 6.2, we first define the profit gain ratio γ as a function of the correlation coefficient ρ:

$$\gamma(\rho) = \frac{U^*(\langle(2\mu, 2(1 + \rho)\sigma^2), 0, \alpha\rangle)}{2 \times U^*(\langle(\mu, \sigma^2), 0, \alpha\rangle)} - 1.$$

Denote $\delta_b^*(\rho) \triangleq \delta^*(\langle(2\mu, 2(1+\rho)\sigma^2), 0, \alpha\rangle)$ as the optimal equilibria for the bundle. Let us first prove the following two lemmas. □

Lemma A.8. *For correlation coefficient* $\rho_0 \in [-1, 1]$, *if* $\gamma(\rho_0) > 0$, *then* $\frac{\sqrt{2(1+\rho_0)}}{2}\sigma < \mu\sqrt{\frac{2}{\pi}}(\alpha+1)$.

Proof. Suppose otherwise that $\frac{\sqrt{2(1+\rho_0)}}{2}\sigma \geq \mu\sqrt{\frac{2}{\pi}}(\alpha+1)$. Then, $\sigma \geq \frac{\sqrt{2(1+\rho_0)}}{2}\sigma \geq \mu\sqrt{\frac{2}{\pi}}(\alpha+1)$. According to Lemma A.3,

$$U^*\left(\langle(2\mu, 2(1+\rho_0)\sigma^2), 0, \alpha\rangle\right) = 2 \times U^*\left(\left\langle\left(\mu, \left(\frac{\sqrt{2(1+\rho_0)}}{2}\sigma\right)^2\right), 0, \alpha\right\rangle\right) \leq 2 \times U^*\left(\langle(\mu, \sigma^2), 0, \alpha\rangle\right). \tag{49}$$

Therefore, bundle is less profitable and $\gamma(\rho_0) \leq 0$, which contradicts to $\gamma(\rho_0) > 0$. Hence, the lemma must hold. □

Lemma A.9. *If* $\gamma(\rho_0) > 0$ *for correlation coefficient* $\rho_0 \in [-1, 1]$, *then* $\gamma(\rho') > \gamma(\rho_0) > 0$ *for any* $\rho' \in [-1, \rho_0)$.

Proof. From Lemma A.8, we know that

$$\frac{\sqrt{2(1+\rho')}}{2}\sigma < \frac{\sqrt{2(1+\rho_0)}}{2}\sigma < \mu\sqrt{\frac{2}{\pi}}(\alpha+1).$$

Then by Lemma A.3, we have

$$U^*\left(\left\langle\left(\mu, \left(\frac{\sqrt{2(1+\rho')}}{2}\sigma\right)^2\right), 0, \alpha\right\rangle\right) > U^*\left(\left\langle\left(\mu, \left(\frac{\sqrt{2(1+\rho_0)}}{2}\sigma\right)^2\right), 0, \alpha\right\rangle\right). \tag{50}$$

The preceding equalities says that the profit under ρ' is higher than the profit under ρ_0. Furthermore, correlation does not affect the separate sale, so $\gamma(\rho') > \gamma(\rho_0) > 0$. Hence, the lemma holds. □

Now let us begin our proof to the theorem:

(1) According to Lemma A.9, we conclude that the correlation coefficient region for the bundle to be more profitable $\mathcal{S} = \{\rho|\gamma(\rho) > 0\}$ is a continuous region $[-1, \bar{\rho})$. Otherwise, let $\rho_{\max} = \sup\{\rho|\gamma(\rho) > 0\}$, then $[-1, \rho_{\max}) \subset \mathcal{S}$ because of Lemma A.9. If $[-1, \rho_{\max}) \neq \mathcal{S}$, then $\exists\hat{\rho} \geq \rho_{\max}$ such that $\hat{\rho} \in \mathcal{S}$, $\gamma(\hat{\rho}) > 0$. It is easy to verify that $\gamma(\cdot)$ is a continuous function. As a result, $\exists\epsilon > 0$ such that $\gamma(\hat{\rho}+\epsilon) > 0$, where $\hat{\rho}+\epsilon > \rho_{\max}$. Here, a contradiction to the maximum of $\rho_{\max}$. As a result, $\mathcal{S} = [-1, \bar{\rho})$.
(2) From Lemma A.9, we also know that γ is a decreasing function of ρ when $\rho \in [-1, \bar{\rho})$.
(3) When $\frac{\sigma}{\mu} < \sqrt{\frac{2}{\pi}}(\alpha+1)$, $\frac{\sqrt{2(1+\rho)}}{2}\sigma < \sigma < \mu\sqrt{\frac{2}{\pi}}(\alpha+1), \forall\rho \in [-1, 1)$. From Lemma A.3, we know that

$$\begin{aligned} U_s^* &= 2U^*\left(\langle(\mu, \sigma^2), 0, \alpha\rangle\right) < 2U^*\left(\left\langle\left(\mu, \left(\frac{\sqrt{2(1+\rho)}}{2}\sigma\right)^2\right), 0, \alpha\right\rangle\right) \\ &= U^*\left(\langle(2\mu, 2(1+\rho)\sigma^2), 0, \alpha\rangle\right) = U_b^*. \end{aligned}$$

We can see from the preceding inequalities that when $\rho \in [-1, 1)$, $\gamma(\rho) > 0$. Moreover, when $\rho = 1$, $\gamma = 0$. Hence, $\mathcal{S} = \{\rho|\gamma(\rho) > 0\} = [-1, 1)$, which means that $\bar{\rho} = 1$.
(4) Let us first prove the following lemma.

LEMMA A.10. *If* $\frac{\sigma}{\mu} > \frac{2}{\sqrt{\pi(1+\rho)}}(\alpha+1)$, *then* $\gamma(\rho) < 0$.

PROOF. Letting $\frac{\sqrt{2(1+\rho)}}{2}\frac{\sigma}{\mu} = \sqrt{\frac{2}{\pi}}(\alpha+1)$, we have $\frac{\sigma}{\mu} = \frac{2}{\sqrt{\pi(1+\rho)}}(\alpha+1)$. As a result, when $\frac{\sigma}{\mu} > \frac{2}{\sqrt{\pi(1+\rho)}}(\alpha+1)$, $\frac{\sqrt{2(1+\rho)}}{2}\sigma > \mu\sqrt{\frac{2}{\pi}}(\alpha+1)$. Hence, $\sigma > \frac{\sqrt{2(1+\rho)}}{2}\sigma > \mu\sqrt{\frac{2}{\pi}}(\alpha+1)$. Thus, by Lemma A.3, we have

$$U_b^* = U^*\left(\langle(2\mu, 2(1+\rho)\sigma^2), 0, \alpha\rangle\right) = 2U^*\left(\left\langle\left(\mu, \left(\frac{\sqrt{2(1+\rho)}}{2}\sigma\right)^2\right), 0, \alpha\right\rangle\right) < 2U^*\left(\langle(\mu, \sigma^2), 0, \alpha\rangle\right) = U_s^*. \tag{51}$$

From the preceding inequalities, we know that bundling is less profitable than separate sales, so $\gamma(\rho) < 0$. □

From Lemma A.10, we know that when $\rho > \frac{4(\alpha+1)^2}{\pi(\sigma/\mu)^2} - 1$, we have $\gamma(\rho) < 0$. Therefore, $\rho \leq \frac{4(\alpha+1)^2}{\pi(\sigma/\mu)^2} - 1$ is a necessary condition for $\gamma(\rho) > 0$. As a result, $\bar{\rho} = \sup\{\rho|\gamma(\rho) > 0\} \leq \frac{4(\alpha+1)^2}{\pi(\sigma/\mu)^2} - 1$.

REFERENCES

[1] W. J. Adams and J. L. Yellen. 1976. Commodity bundling and the burden of monopoly. *Quarterly Journal of Economics* 90, 3 (1976), 475–498.

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

[3] Y. Bakos and E. Brynjolfsson. 2000. Bundling and competition on the internet. *Marketing Science* 19, 1 (2000), 63–82.

[4] Mihai Banciu and Fredrik Ødegaard. 2016. Optimal product bundling with dependent valuations: The price of independence. *European Journal of Operational Research* 255, 2 (2016), 481–495.

[5] Kostas Bimpikis and Mihalis G. Markakis. 2015. Inventory pooling under heavy-tailed demand. *Management Science* 62, 6 (2015), 1800–1813.

[6] Chiranjeeb Buragohain, Divyakant Agrawal, and Subhash Suri. 2003. A game theoretic framework for incentives in P2P systems. In *Proceedings of the 3rd International Conference on Peer-to-Peer Computing (P2P'03)*. 48–56.

[7] O. Candogan, K. Bimpikis, and A. Ozdaglar. 2010. Optimal pricing in the presence of local network effects. In *Proceedings of the 6th International Conference on Internet and Network Economics*. 118–132.

[8] Chenghuan Sean Chu, Phillip Leslie, and Alan Sorensen. 2011. Bundle-size pricing as an approximation to mixed bundling. *American Economic Review* 101, 1 (2011), 263–303.

[9] D. Easley and J. Kleinberg. 2010. *Networks, Crowds, and Markets*. Cambridge University Press.

[10] Gary D. Eppen. 1979. Effects of centralization on expected costs in a multi-location newsboy problem. *Management Science* 25, 5 (1979), 498–501.

[11] Hanming Fang and Peter Norman. 2006. To bundle or not to bundle. *RAND Journal of Economics* 37, 4 (2006), 946–963.

[12] Neil Gandal, Sarit Markovich, and Michael Riordan. n.d. *Ain't It "Suite"? Strategic Bundling in the PC Office Software Market*. Technical Report.

[13] Yigal Gerchak and Qi-Ming He. 2003. On the relation between the benefits of risk pooling and the variability of demand. *IIE Transactions* 35, 11 (2003), 1027–1031.

[14] Roch Guérin, Jaudelice C. de Oliveira, and Steven Weber. 2014. Adoption of bundled services with network externalities and correlated affinities. *ACM Transactions on Internet Technology* 14, 2–3 (Oct. 2014), Article 13, 32 pages.

[15] J. P. Guiltinan. 1987. The price bundling of services: A normative framework. *Journal of Marketing* 51, 2 (1987), 74–85.

[16] Lorin M. Hitt and Pei-Yu Chen. 2005. Bundling with customer self-selection: A simple approach to bundling low-marginal-cost goods. *Management Science* 51, 10 (2005), 1481–1493.

[17] Rustam Ibragimov and Johan Walden. 2010. Optimal bundling strategies under heavy-tailed valuations. *Management Science* 56, 11 (2010), 1963–1976.

[18] M. L. Katz and C. Shapiro. 1985. Network externalities, competition, and compatibility. *American Economic Review* 75, 3 (1985), 424–440.

[19] J. O. Kephart and S. A. Fay. 2000. Competitive bundling of categorized information goods. In *Proceedings of the 2nd ACM Conference on Electronic Commerce (EC'00)*. ACM, New York, NY, 117–127.

[20] S. J. Liebowitz and S. E. Margolis. 1994. Network externality: An uncommon tragedy. *Journal of Economic Perspectives* 8, 2 (1994), 133–150.

[21] C. Matutes and P. Regibeau. 1992. Compatibility and bundling of complementary goods in a duopoly. *Journal of Industrial Economics* 40, 1 (1992), 37–54.

[22] R. P. McAfee, J. McMillan, and M. D. Whinston. 1989. Multiproduct monopoly, commodity bundling, and correlation of values. *Quarterly Journal of Economics* 104, 2 (1989), 371–383.

[23] D. Niyato, D. T. Hoang, N. C. Luong, P. Wang, D. I. Kim, and Z. Han. 2016. Smart data pricing models for the Internet of Things: A bundling strategy approach. *IEEE Network* 30, 2 (March 2016), 18–25. DOI: https://doi.org/10.1109/MNET.2016.7437020

[24] A. Prasad, R. Venkatesh, and V. Mahajan. 2010. Optimal bundling of technological products with network externality. *Management Science* 56, 12 (2010), 2224–2236.

[25] Roch Guérin, Jaudelice C. de Oliveira, and Steven Weber. 2014. Adoption of bundled services with network externalities and correlated affinities. *ACM Transactions on Internet Technology* 14, 2–3 (Oct. 2014), Article 13.

[26] R. Schmalensee. 1984. Gaussian demand and commodity bundling. *Journal of Business* 57, 1 (1984), 211–230.

[27] R. Venkatesh and W. Kamakura. 2003. Optimal bundling and pricing under a monopoly: Contrasting complements and substitutes from independently valued products. *Journal of Business* 76, 2 (2003), 211–231.

[28] C. C. Wang, Y. Hsu, and W. Fang. 2004. Acceptance of technology with network externalities. *Journal of Information Technology Theory and Application* 6, 4 (2004), 15–28.

[29] C. C. Wang, S. K. Lo, and W. Fang. 2008. Extending the technology acceptance model to mobile telecommunication innovation: The existence of network externalities. *Journal of Consumer Behaviour* 7, 2 (2008), 101–110.

Received August 2017; revised July 2018; accepted September 2018