Platform Pricing in Mixed Two-Sided Markets*

Ming Gao†

May 2015

Abstract

When a consumer can appear on both sides of a two-sided market (2SM), such as a user who both buys and sells on eBay, the platform may want to bundle the services provided to two sides. Other examples include renewable energy power grids where households can both use and generate electricity and text-processing software that allows users to both read and create files. We develop a general model for such "mixed" 2SMs, and find that a monopolist platform's incentive to bundle and its optimal pricing strategy are determined by simple and testable formulas using familiar price elasticities of demand, which embody the bundling effect, and price-cost margins adjusted for network externalities, which incorporate "two-sidedness". Contrary to existing results on bundling using non-price elasticities, we show when elasticities are measured with the price for either side, bundling may be profitable even when the bundle demand is less elastic than the demand for either side. The impacts from different sides on the incentive to bundle are separable in a "seesaw" pattern. The optimal pricing rule for mixed 2SMs generalizes the familiar Lerner formula that applies in non-mixed 2SMs, where the optimal adjusted price-cost margin on each side may be higher or lower than the inverse price elasticity of demand, contingent on a comparison between how well the platform performs on the opposite side and in the intersection of two sides. Comparative statics and potential applications are discussed.

Key Words: two-sided market, platform, bundling, price elasticity of demand

JEL **Classification**: D42, L11, L12, L22.

---

*I thank Mark Armstrong, Jean-Pierre Benoît, Glen Weyl, Alex White and two referees for very valuable comments. I also thank Emeric Henry, Eric Maskin, Preston McAfee, David P. Myatt, Marco Ottaviani, Emre Ozdenoren and Leonard Waverman for comments on my PhD dissertation which is the basis of this work. I thank Stefan Behringer, Paul Belleflamme, Yong Chao, Jaimie Lien, Xiao Liu (Tracy), Kathy Spier, Ruqu Wang, Ruhai Wu, Wen Zhou and other participants in IIOC 2015, the 6th Workshop on the Economics of ICTs, EARIE 2014, the 3rd Workshop on Industrial Organization: Theory, Empirics and Experiments, AMES 2014, the 2014 Annual Conference of the American Law and Economics Association, the 2014 Annual Conference of the RES, and seminars at Tsinghua and Renmin. I acknowledge the financial assistance of Project 20121087916 supported by the Tsinghua University Initiative Scientific Research Program, Project 71203113 supported by the National Natural Science Foundation of China, and Project 20131029117 sponsored by SRF for ROCS, SEM.

†School of Economics and Management, Tsinghua University, Beijing 100084, China. Email: gaom@sem.tsinghua.edu.cn.

1 Introduction

Existing theories of two-sided markets in the literature have mostly focused on classic examples such as credit card, video game and media, where there are good reasons to model the two sides as two **distinct** groups of participants - card holders versus merchants, game players versus developers, and media viewers versus advertisers - because there is little possibility of overlap between these groups (see, for instance, Caillaud and Jullien (2003), Armstrong (2006), Rochet and Tirole (2003 and 2006) and Weyl (2010)).

However, in many other, perhaps less classic two-sided markets, one participant may act on **both** sides of the market. Consider for instance the consumer-to-consumer online marketplace mediated by eBay, which clearly exhibits "two-sidedness" as buyers (respectively, sellers) value eBay more when they expect to have access to a larger group of sellers (respectively, buyers) on the platform.[1] More importantly, an eBay user can quite freely buy and sell as he or she pleases, and therefore there exists an overlap between the buyer side and seller side of the market. In telecommunications markets, if we consider the networks as platforms connecting call makers or message senders to receivers, there clearly exist network externalities across the two groups, which indicates two-sidedness. But many (if not most) subscribers both make and receive phone calls, and many both send and receive messages. Therefore the two groups have a conceivably large overlap. In the fast growing renewable energy market, power grids in many countries (e.g. Australia, Germany and UK) are offering households the option of installing and connecting home electricity generation systems that can both draw power from and feed power to the grid. In some regions, especially those with no or very limited power supply from traditional utility companies (e.g. parts of Alaska and Scotland), new self-sustainable power grids linking home solar/wind electric systems are built, where each household can be an electricity user, an electricity feeder, or both. Such power grids exhibit both two-sidedness and overlap between two sides. Other similar examples include many kinds of financial intermediation where consumers can both buy and sell, or both borrow and lend (such as securities brokerage and crowd funding), software that allows users to both create and view files in certain formats (such as text-processing software and computer-aided design software), and information exchange platforms that allow users to both post (or send) and view messages (such as bulletin boards, online forums, social networking websites and user-generated content platforms).

In this article, we provide a general model for two-sided markets where *a consumer can appear on different sides of the market*, which we call **mixed two-sided markets**. If no one can appear on both sides, we call it a **standard** two-sided market. The existence

[1] This analysis follows the definition of two-sidedness from an "indirect network externality" perspective, as in Armstrong (2006) and Rysman (2009). Alternatively, one could use the definition based on price structures by Rochet and Tirole (2006): If eBay were to charge a fixed total price to the seller and buyer in each transaction, how this price is divided between the two parties will most certainly affect their total volume of transactions.

of "double-side" consumers brings new pricing options to a mixed two-sided platform. In particular, in addition to setting a separate price for each side, it can also engage in **bundling** and charge all double-side consumers a price that is different than the sum of the separate prices for two sides. We study when such bundling is desirable, and how it affects the platform's optimal pricing strategy.

In real life, such bundling may be achieved by **two-part tariffs**. Solar/wind power grids, for instance, typically use a two-part tariff consisting of one common access fee applicable to all installed home electric systems, which allows for both power using and feeding, plus an additional fee for usage and a credit (i.e. a negative fee) for feeding in electricity. Most telecommunications networks also charge subscribers a monthly fee for access to both callers and receivers on the network, plus additional separate fees for making and answering calls[2]. Such two-part tariffs can generally be represented by three prices (A, a_1, a_2), where A is the common fixed fee applicable to all consumers, and a_i is the additional fee applicable to buyers of product i. Long (1984) has shown that such a two-part tariff is equivalent to a **mixed bundling** strategy, (p_1, p_2, p_b), where p_i is the final price for product i which is equal to $A + a_i$, and p_b is the final price for the bundle of two products which is equal to $A + a_1 + a_2$. Moreover, whenever there is a positive fixed fee A in the two-part tariff, there will be an equal positive discount for the bundle of two products in mixed bundling. The reason is that, under the two-part tariff, a buyer of both products pays the fixed fee A **once**, not twice. Therefore, compared to separate purchases of two products, the buyer of the bundle receives a discount that is exactly equal to the fixed fee A (as $p_1 + p_2 - p_b = A$). *Therefore the fixed fee that applies to all consumers of the platforms mentioned previously actually creates a discount for joining both sides, and thereby implements mixed bundling of the services for two sides.*

There are certainly also platforms that use separate pricing strategies for different sides, without bundling. For instance, Zopa.com, a crowd funding platform facilitating lending and borrowing among individuals, offers different accounts for lenders and borrowers, which involve completely different fees and none of them applies to both sides. eBay only charges sellers whilst providing free services to buyers. Table 1 summarizes the different choice of pricing strategies by these real-life mixed two-sided platforms.

[2] Mobile telephone networks in some countries, e.g. China and US, used to or still charge subscribers for answering calls, although in many other regions, e.g. Europe, it is free to receive calls. When there is a fee for answering calls, it may be different from that for making calls. For instance, in one tariff offered by China Mobile in 2003, the per-minute fees for answering and making calls were RMB 0.25 yuan and 0.6 yuan, respectively.

Table 1. Mixed Two-Sided Platforms and Their Pricing Strategies

Platform	Solar/Wind Power Grid	Telecom. Network	Stock Exchange	eBay	Zopa.com	Acrobat Software
Side 1	Feeder	Caller/Sender	Seller	Seller	Lender	Reader
Side 2	User	Receiver	Buyer	Buyer	Borrower	Writer
Strategy Choice	2-Part Tariff/ Bundling	2-Part Tariff/ Bundling	2-Part Tariff/ Bundling	Separate Pricing	Separate Pricing	Separate Pricing
Access to 1	One common fee	One common fee	One common fee	-	Fee	Fee
Access to 2				-/Fee	Fee	-
Usage on 1	Credit	Fee	Fee	-/Fee	Fee	-
Usage on 2	Fee	-/Fee	Fee	-	Fee	-

What determines a monopolist platform's incentive to bundle the services for two sides? What is the optimal pricing strategy when such bundling is feasible?

The existing two-sided market models are not equipped to answer these questions, as in these models any "double-side" consumer would be split into two "artificial" consumers - each on one side of the market only - who make consumption decisions **independently** of each other, although such decisions are in reality inter-dependent and made by the same consumer herself. Ignoring the existence of "double-side" consumers will inevitably result in a misrepresentation of their actual demand, and hence, in misleading implications for the platform's pricing strategies. As Armstrong (2006) noted in the discussion of his Proposition 1 (on page 673), "*for the analysis* (of two-sided markets) *to apply accurately* (to software markets)*, though, there need to be two disjoint groups of agents: those who wish to read files and those who wish to create files. It does not readily apply when most people wish to perform both tasks.*"

Using our model for mixed two-sided markets, we show in simple and testable formulas that both the incentive to bundle and the optimal pricing strategy are determined by familiar **price elasticities of demand**, which embody the bundling effect, and **price-cost margins adjusted for network externalities**, which incorporate the impact of two-sidedness. Price elasticities of demand are useful for gauging bundling profitability because they by definition measure the revenue gains due to a marginal price reduction (i.e. price multiplied by the marginal increase in demand) as a fraction of the corresponding revenue losses. Price-cost margins (i.e. the difference between price and cost divided by the relevant price) are also useful because, when multiplied to the relevant price elasticities, they convert such elasticities from revenue measures into net profit measures. Finally, two-sidedness affects the optimal price-cost margins because the value that a consumer on one side creates or destroys for the opposite side through network externalities ought to be accounted for in the calculation of the relevant profit margin earned from this consumer.

Before a dedicated discussion of the position of this paper in the literature, it is worthwhile to highlight several qualitatively different properties of our analyses and results, compared to the existing literatures on standard two-sided markets and on bundling in one-sided markets.

1. Long (1984) and Armstrong (2013) use a particular kind of non-price elasticity[3] to show that, in a one-sided market context, a bundle discount is profitable when the bundle demand is **more elastic** than the demand for either one product. Many people find this result intuitive as it closely resembles the familiar single-product Lerner formula where a more elastic demand implies a lower monopoly price. We use standard **price** elasticities, as Lerner (1934) did, and show that, contrary to the intuition shown by Long (1984) and Armstrong (2013), bundling may actually be profitable even if the bundle demand (i.e. number of "double-side" consumers in our model) is **less elastic** than the demand for either one product (i.e. number of consumers on one side). We also show that our result is robust in a one-sided context when network externalities are eliminated from our model.

Because we measure demand elasticities with the price for different sides, we are able to separate the impact from different sides on profitability when a discount is introduced for joining both sides (on top of the optimal separate prices). As explained previously, it turns out that the product between the elasticity of the bundle demand - measured with the price for side 1, say - and the adjusted price-cost margin of side 1 exactly captures the net gain from side 1 due to the discount, as a fraction of the cost of offering the discount. Therefore when the sum of two such fractions, each representing one side, exceeds 1, the discount increases profits.

Sometimes we can replace the optimal adjusted price-cost margin without bundling with the inverse of price elasticity of demand for the relevant side. This is true in the monopoly model of standard two-sided markets by Armstrong (2006), and also true in the special case with additivity in our model. In this case the previous fraction becomes the ratio between the elasticity of the bundle demand and the price elasticity of the demand for the relevant side. The incentive to bundle then depends on the sum of two such elasticity ratios, each representing one side. Therefore, in order for bundling to be profitable, we only need that the elasticity ratios for different sides are on average larger than one half, such that their sum will exceed 1. This means that the bundle demand only needs to be **more than half as elastic** as the demand for either side on average.

This finding reveals that the intuition shown by Long (1984) and Armstrong (2013) is actually very different from that of the familiar Lerner formula which links optimal prices to price elasticities of demand. In terms of standard price elasticities, the profitability of bundling generally does **not** rely on the bundle demand to be more elastic than the demand for either one product.

[3] Long (1984) and Armstrong (2013) defined elasticities with respect to "simultaneous and equal percentage increases in price-cost markups of both products", instead of prices. See section 3.4A for a more detailed discussion.

2. Our analysis also reveals a **seesaw pattern** in the determination of the incentive to bundle, where the impact from different sides are separable. That is, if the bundle demand is highly elastic with respect to the price for side 1, or if the optimal adjusted price-cost margin without bundling for side 1 is very large, then the constraint that overall profitability puts on side 2 will be more lax, and vice versa. We identify the accurate channels through which such "asymmetry" between the two sides affects the incentive to bundle.

3. The optimal pricing rule in mixed two-sided markets generalizes the Lerner formula for standard two-sided markets by Armstrong (2006). When the adjusted price-cost markup is positive on one side, the optimal adjusted price-cost margin for this side may be larger or smaller than the inverse of the price elasticity of demand, contingent on a comparison between how well the platform performs on the opposite side and in the intersection of two sides. A high margin on side 1, for instance, is necessary to compensate for a bad performance on side 2 relative to that in the intersection.

4. In some more tractable special cases of our general model, more interesting results emerge. When consumers' valuations of two sides are symmetrically distributed, for instance, a higher "degree of overlap" between the two sides without bundling becomes a favorable indicator of the profitability of bundling. Moreover, when a negative bundle discount is feasible (i.e. charging a premium for joining both sides), and the total number of single-side consumers is non-increasing in the price charged to either side, the optimal bundle discount is positive **if and only if** the bundle demand is more elastic than the total demand (i.e. the total number of consumers on the platform) given the optimal pricing strategy.

This last result identifies a necessary and sufficient condition for profitable bundling, unlike most existing results in the literature and our first result discussed previously which are only sufficient. Because the elasticities are still measured by prices, this result is consistent with the familiar Lerner intuition that a more elastic demand implies a lower price. However, the correct comparison should be made between the bundle demand and the **total** demand.

Besides these conceptual and theoretical new findings, our results also have the potential advantage of being more convenient for both empirical testing and practical applications, compared to existing theories on monopolist mixed bundling (to be summarized shortly), for two reasons: 1) most of our results require virtually no information about the underlying distribution of consumers' valuations; 2) data on price elasticities and price-cost margins are easier to obtain or to estimate from usual accounting records than, say, data on non-price elasticities used in the literature.

The main limitation of our results is that they are derived from a monopoly model, and therefore do not apply in markets with competition among multiple platforms. However, they are relevant to markets mentioned previously with a dominant platform, such as solar/wind power grids in areas without traditional utility suppliers, stock exchanges with little or no competition due to financial regulation, and software or information exchange platforms with very large installed bases that allow for direct exploitation of monopoly power, etc. Several such potential applications are discussed after we present the relevant results.

Position in literature Our theory fills a gap in the literature and generalizes the canonical two-sided market model by Armstrong (2006) through incorporating features from models of mixed bundling in one-sided markets by Long (1984) and Armstrong (2013). Each consumer in our model can freely choose to join side 1 alone, side 2 alone, both sides, or neither. In the terminology of Rochet and Tirole (2006) and Weyl (2010), consumers are heterogeneous in the "membership benefits" they derive from joining each side or both sides. We assume such benefits from different sides may be **non-additive** for each consumer, and therefore the two sides may be complements, substitutes, or neither. Following Armstrong (2006), we assume consumers on the same side are homogeneous in the "interaction benefit" (i.e. network externality) they derive from every member of the opposite side, and model demand as functions of utility provided by the platform, unlike in Rochet and Tirole (2006) where demand depends on prices directly. The former approach allows us to distinguish between the effects of bundling and those due to network externalities, crucial for unifying multi-product pricing and two-sidedness in one general framework.

There are several existing models that study bundling in two-sided markets. For instance, Rochet and Tirole (2008) study **tying** - a special way to bundle two products by eliminating the buyers' option of purchasing one of them alone - in credit/debit card markets, where the two sides are merchants and card holders, and the payment card platform may tie-in their credit and debit cards on the merchant side. Chao and Derdenger (2013) look at mixed bundling by video game platforms which offer both hardware consoles and software content to game players. In these models, as there exists no overlap between two sides, the markets are standard instead of mixed two-sided markets, and the potential bundles are only offered to one side of the market. The bundling in our framework, however, takes place across two sides, and therefore our model, our results and the intended applications are very different.

Our model and results are "compatible" with one-sided markets as a degenerate case, because the impact of two-sidedness is isolated into the adjustments made to the standard price-cost markups and margins. In the large theoretical literature on **mixed bundling** of two products in one-sided markets, one strand of work focuses on the properties of the joint distribution of consumers' valuations for two products (especially their correlation or stochastic dependence), which includes Adams and Yellen (1976), Schmalensee (1984), McAfee, McMillan and Whinston (1989), and Chen and Riordan (2013) (who use copulas instead of distributions, to be more precise), among others. Several other works apply the law of large numbers to **pure bundling** of many products - a special way of mixed bundling by eliminating the option of purchasing subsets of all products - which includes Bakos and Brynjolfsson (1999) and Fang and Norman (2006), among others. Given specific distributions of consumer valuations, different special ways of mixed bundling (e.g. separate pricing, pure bundling, etc.) can usually be ranked in terms of profitability. For instance, Schmalensee (1984) used normal distribution, and Fang and Norman (2006) used log-concave distributions. One focus of this strand has therefore been to generalize the distributions and constraints on their behavior that will still allow for the ranking of different pricing strategies.

We follow another strand, i.e. Long (1984) and Armstrong (2013), that uses standard aggregate demand functions and links the incentive to bundle to elasticities of demand. However, we use **price elasticities** whereas their elasticities are defined with respect to "simultaneous and equal percentage increases in price-cost markups of both products", instead of prices. Long (1984) actually studied a monopolist using a two-part tariff that is equivalent to mixed bundling (as discussed earlier), and assumed additive values from different products. Armstrong (2013) used the same method as Long (1984) but generalized the analysis to non-additive values as well as bundling of products from different sellers. Our model is closer to Armstrong (2013)'s monopoly model as we consider mixed bundling with non-additive values. We however do not address bundling by different sellers or competition among multiple platforms, so our results do not apply in those situations. The main difference between their results and ours are summarized in the previous point 1. A more detailed comparison is postponed until section 3.4 after we present the relevant results.

To our knowledge, none of the articles mentioned previously presents the general properties of the optimal mixed bundling strategy. Hanson and Martin (1980) have solved the optimal mixed bundling problem with **discrete** consumer valuation distributions for two or more products in one-sided markets. Manelli and Vincent (2006) discuss the same problem with additive valuations and **continuous** distributions, and present results as constraints on the distributions. We use general non-additive valuations (of two products) with continuous distributions, and present properties of the optimal pricing strategy using price elasticities, which are also robust in one-sided markets. Although these results are not directly comparable, one sure difference is that, in the first-order conditions, the relevant optimal price-cost margin from each consumer in a one-sided market needs to be adjusted by the network externalities that a consumer in a mixed two-sided market creates for the opposite side, and what really matters is the resulting adjusted margin as mentioned previously. Armstrong (2006) actually shows that the optimal (separate) pricing strategy in standard two-sided markets satisfies a modified Lerner formula involving exactly such adjusted margins. Our results generalize Armstrong (2006)'s formula, but show in stark contrast that, when the market is mixed, the optimal margins may be larger or smaller than what the Lerner formula suggests.

In the remaining parts of this article, section 2 presents the basic model and discusses separate pricing as a benchmark; section 3 discusses the platform's incentive to bundle; section 4 derives the optimal pricing strategy; section 5 provides comparative statics analysis; section 6 discusses the welfare-maximizing strategy and the special case of one-sided market, and also provides a numerical example; and section 7 concludes.

2 Modelling Framework

2.1 Basic setting

A monopolist platform facilitates interaction between two market sides: $i = 1, 2$. A continuum of consumers may choose to join side 1 alone, side 2 alone, both sides, or neither. If a consumer joins side i only, we assume she obtains a total surplus of $u_i + t_i$, where u_i is a **common value** that the platform provides to all side-i members, and t_i represents some **idiosyncratic value** this consumer derives from side i.[4]

Common values Consumers on either side benefit from interaction with more members on the opposite side. Common values incorporate such benefits in the same way as in Armstrong (2006). In particular, if the platform attracts N_j members on side j, and charges a price p_i to side i, each consumer on side i will obtain a common value of

$$u_i \equiv \alpha_i N_j - p_i \tag{1}$$

where α_i measures the benefit a side-i consumer enjoys from interacting with each side-j consumer. The interaction between each pair of members of two sides therefore creates a total "interaction surplus" of $(\alpha_1 + \alpha_2)$. Parameters α_1 and α_2 represent the **strengths of network effects** in different directions across the platform, and are assumed to be exogenous.[5]

To study how the platform can exploit the fact that some consumers may want to join both sides, we allow the platform to offer a **discount** p_X to anyone who joins both sides, in addition to the separate prices, p_1 and p_2. This consumer therefore pays a total price

[4] Following most of the theorectical literature on two-sided markets, e.g. Armstrong (2006) and Rochet and Tirole (2006), both u_i and t_i here are defined as a consumer's utility pertaining to the *platform's* services and characteristics, net of any tansfers between consumers (if any).

[5] α_i corresponds to the "per-interaction benefit" of each member of side i in Rochet and Tirole (2006) and Weyl (2010). Here it is assumed to be the same for all members of side i, following Armstrong (2006). Although we will often interpret these alphas as positive numbers, consistent with most of the examples we discuss in section 1, our model and results do not depend on such an interpretation. A negative α_i, for instance, indicates how much members on side i dislike interaction with each side-j consumer. In the knife-edge case where $\alpha_1 = \alpha_2 = 0$, the market is simply one-sided.

of $p_1 + p_2 - p_X$ to the platform, and obtains a common value of

$$u_b \equiv (\alpha_1 N_2 + \alpha_2 N_1) - (p_1 + p_2 - p_X) = u_1 + u_2 + p_X$$

where the subscript b refers to the "bundle" of two sides, and the second equation follows from (1). Therefore there is an **extra common value** that the platform offers for joining both sides

$$u_X \equiv u_b - u_1 - u_2 = p_X \tag{2}$$

which is achieved exactly through the bundle discount p_X. Thus in our settings, $u_X = p_X$ always holds and they are interchangeable.

Idiosyncratic values Consumers may also derive different personal benefits from access to the platform's different sides, which are independent of the number of members on either side. For instance, mobile telecommunications network subscribers may derive different utility from having the option of making emergency calls. The idiosyncratic value that a consumer derives from joining side i alone is t_i. To keep the model general, we assume that joining both sides may bring an **extra idiosyncratic value** t_X, in addition to $t_1 + t_2$.[6]

A consumer's idiosyncratic values from two sides are called **additive** if $t_X = 0$. They are **complements** if $t_X > 0$, and are **substitutes** if $t_X < 0$. Each consumer can therefore be indexed with a **type** vector, $\mathbf{t} \equiv (t_1, t_2, t_X)$, which varies among different consumers following some *exogenous distribution $G(\cdot)$ known to the platform.*

Assumption 1 (Type) *The marginal distribution of (t_1, t_2) has a **continuous** density on support $\mathbb{T} \subseteq \mathbb{R}^2$, where $\mathbb{T}$ is **weakly convex** with **full dimension** on $\mathbb{R}^2$.*[7]

The regularity constraint on the marginal distribution of t_X is only such that the resulting demand functions to be defined later are differentiable. This allows for a wide range of cases, including a constant t_X for all consumers. The total surplus a consumer derives from joining both sides is the sum of all the common and idiosyncratic values: $u_1 + u_2 + u_X + t_1 + t_2 + t_X$. Not joining either side yields a total surplus of zero.

We use $\mathbf{u} \equiv (u_1, u_2, u_X) \in \mathbb{R}^3$ to denote the common values, and $\mathbf{p} \equiv (p_1, p_2, p_X) \in \mathbb{R}^3$ to denote the pricing strategy. Whenever the platform sets a discount $p_X = 0$, we say that it is using **separate pricing**. Whenever $p_X \neq 0$, we say that it is using **mixed**

[6] t_i corresponds to the "membership benefit" of each member of side i in Rochet and Tirole (2006) and Weyl (2010). Following Armstrong (2006), we assume consumers are heterogeneous in their membership benefits. However, unlike in Armstrong (2006) where each consumer only has one membership parameter as she can only appear on one side, in our model each consumer has three membership parameters (t_1, t_2, and t_X) and consumers may be heterogeneous in all of them.

[7] A two-dimensional support for the distribution of (t_1, t_2) is crucial here, as in all bundling models. Note that if $\mathbb{T} = (\mathbb{R} \times \varnothing) \cup (\varnothing \times \mathbb{R})$ so that each consumer obtains an idiosyncratic value from only one of the two sides but never from both (and $\mathbb{T}$ is one-dimensional in this case), the market becomes a standard two-sided market and our model reduces to the monopoly model of Armstrong (2006).

bundling, as in this case the price for the bundle of two sides together ($p_1 + p_2 - p_X$) is different than the sum of the prices for two sides.

To keep the model general, we allow all $\mathbf{u}$, $\mathbf{p}$ and $\mathbf{t}$ to take negative values whenever feasible. $u_1 < 0$, for instance, means that the platform offers a "disutility" to side 1 (but consumers with high t_1 may still find it worthwhile to join side 1). Negative prices may emerge, for instance, because it may be worthwhile for the platform to subsidize participation by some consumers in order to take advantage of the large network benefits they create. A negative bundle discount (i.e. a premium) is also allowed should it be feasible and the platform find it profitable. Some consumers may even have negative idiosyncratic values, which represent some personal "disutility" from being associated with the platform[8], in which case we require that the distribution of $\mathbf{t}$ should not be "too concentrated" on the negative values, such that there still exists some pricing strategy $\mathbf{p}$ that yields the platform positive profits.

Demand and profit For model transparency and tractability, we define demand as functions of the common values $\mathbf{u}$, instead of prices $\mathbf{p}$.[9] Given $\mathbf{u} = (u_1, u_2, u_X)$ provided by the platform, a consumer chooses the largest amongst the following four options: $\{u_1 + t_1, u_2 + t_2, u_1 + t_1 + u_2 + t_2 + u_X + t_X, 0\}$. The demand functions in this discrete choice problem are defined as follows:

$$\begin{aligned}
\text{Side 1 alone} \quad &: \quad D_1(\mathbf{u}) \equiv \Pr[u_1 + t_1 \geq \max\{u_2 + t_2, u_1 + t_1 + u_2 + t_2 + u_X + t_X, 0\}] \\
\text{Side 2 alone} \quad &: \quad D_2(\mathbf{u}) \equiv \Pr[u_2 + t_2 \geq \max\{u_1 + t_1, u_1 + t_1 + u_2 + t_2 + u_X + t_X, 0\}] \\
\text{Both sides} \quad &: \quad D_b(\mathbf{u}) \equiv \Pr[u_1 + t_1 + u_2 + t_2 + u_X + t_X \geq \max\{u_1 + t_1, u_2 + t_2, 0\}] \\
\text{Side } i \quad &: \quad N_i(\mathbf{u}) \equiv D_i(\mathbf{u}) + D_b(\mathbf{u}) \\
\text{Total} \quad &: \quad N(\mathbf{u}) \equiv D_1(\mathbf{u}) + D_2(\mathbf{u}) + D_b(\mathbf{u})
\end{aligned}$$

These functions are uniquely defined once the consumer type distribution $G(\cdot)$ is given. The maximized aggregate consumer surplus is then

$$V(\mathbf{u}) \equiv \mathbf{E}_{\mathbf{t}}[\max\{u_1 + t_1, u_2 + t_2, u_1 + t_1 + u_2 + t_2 + u_X + t_X, 0\}]. \tag{3}$$

By the envelope theorem, we have

$$\frac{\partial V}{\partial u_1} = N_1, \frac{\partial V}{\partial u_2} = N_2, \frac{\partial V}{\partial u_X} = D_b.$$

[8] For instance, some consumers may prefer not having to remember or to keep records of their account numbers and/or passwords that are required to gain access to some platforms.

[9] Armstrong (2006) also models demand as functions of utitlities. The negative side of this modeling choice, however, is that the resulting demand becomes *implicit* functions of the parameters for network effects, α_1 and α_2, which only directly affect $\mathbf{u}$ (see definition (1)) but not other variables. In some parts of the following analyses, when we do not write out the composition of $\mathbf{u}$, features of two-sidedness may not show directly in equations. They however still exist and are implicit in $\mathbf{u}$.

To lighten notation, we use superscripts to denote derivatives. In particular, for $i \in \{1, 2, b\}$ and $j \in \{1, 2, X\}$, let

$$N_i^j(\mathbf{u}) \equiv \frac{\partial}{\partial u_j} N_i(\mathbf{u}),\ D_i^j(\mathbf{u}) \equiv \frac{\partial}{\partial u_j} D_i(\mathbf{u}), \text{ and } N^j(\mathbf{u}) \equiv \frac{\partial}{\partial u_j} N(\mathbf{u}).$$

By the symmetry of second-order derivatives of $V(\mathbf{u})$, we have

$$N_1^2 = N_2^1, N_1^X = D_b^1, N_2^X = D_b^2. \tag{4}$$

Lemma 1 *For $i, j \in \{1, 2\}$, $i \neq j$, the signs of first order derivatives of demand are: $N_i^i > 0$, $N_i^X = D_b^i > 0$, $D_i^i > 0$, $D_i^j < 0$, $D_i^X < 0$, $D_b^X > 0$, $N^i > 0$, $N^X > 0$. Also we have $N_i^i > \left| N_j^i \right|$.*[10]

(All omitted proofs are provided in the Appendix.) If the platform incurs a **fixed cost** f_i for each member on side i and there is no other costs, its profit is

$$\underbrace{N_1(p_1 - f_1)}_{\text{profit from side 1}} + \underbrace{N_2(p_2 - f_2)}_{\text{profit from side 2}} - \underbrace{D_b p_X}_{\text{discount for joining both sides}}$$

which can be written using (1) and (2) as a function of $\mathbf{u}$

$$\Pi(\mathbf{u}) \equiv \underbrace{(\alpha_1 + \alpha_2) N_1 N_2}_{\text{value created in interactions}} - \underbrace{N_1(u_1 + f_1)}_{\text{loss due to side 1}} - \underbrace{N_2(u_2 + f_2)}_{\text{loss due to side 2}} - \underbrace{D_b u_X}_{\text{additional loss through discount}} \tag{5}$$

In the remaining parts of the article, we will maximize $\Pi(\mathbf{u})$ with respect to different $\mathbf{u}$ that corresponds to either separate pricing (where $u_X = 0$) or mixed bundling (with no restriction on u_X), where the following concepts play a key role.

Adjusted price-cost markup To the platform, each member of side i brings a profit of $p_i - f_i$. Due to two-sidedness, each consumer also creates a value of α_j for each member on side j. Denote v_i the **price-cost markup** that the platform earns from each side-i member, adjusted by the total network benefits that each member creates for the opposite side, which we will also refer to as the **economic value** earned from this consumer:

$$\begin{aligned} v_1 &\equiv p_1 - f_1 + \alpha_2 N_2 = (\alpha_1 + \alpha_2) N_2 - u_1 - f_1 \\ v_2 &\equiv p_2 - f_2 + \alpha_1 N_1 = (\alpha_1 + \alpha_2) N_1 - u_2 - f_2 \end{aligned} \tag{6}$$

[10] The sign of N_j^i depends on the distribution of t_X and may also change at different $\mathbf{u}$. For instance, when $t_X = 0$ for all consumers, $N_j^i(u_X > 0) > 0$, $N_j^i(u_X < 0) < 0$, and $N_j^i(u_X = 0) = 0$; when $t_X > 0$ for all consumers, $N_j^i(u_X > 0) \geq 0$; when $t_X < 0$ for all consumers, $N_j^i(u_X < 0) \leq 0$.

The corresponding **adjusted price-cost margin** on side i is denoted

$$r_i \equiv \frac{p_i - (f_i - \alpha_j N_j)}{p_i} = \frac{v_i}{p_i} \tag{7}$$

These adjusted markups and margins incorporate two-sidedness into their usual one-sided counterparts. It is sometimes useful to rewrite profit using v_i as

$$\Pi = \underbrace{N_1 v_1 + N_2 v_2}_{\text{economic value from two sides}} - \underbrace{(\alpha_1 + \alpha_2) N_1 N_2}_{\text{network benefits enjoyed by all consumers}} - \underbrace{D_b u_X}_{\text{total discount}} \tag{8}$$

Price elasticities of demand We define the following general price elasticities of demand. For $i, j \in \{1, 2\}$, $i \neq j$, *given that the platform attracts N_j members on side j* (such that changes in u_i and p_i are one-to-one according to (1)), denote

$$E_i^i \equiv \frac{N_i^i \cdot p_i}{N_i} \text{ the } \textit{price elasticity of demand for side } i, \tag{9a}$$

$$E_j^i \equiv \frac{N_j^i \cdot p_i}{N_j} \text{ the } \textit{elasticity of demand for side } j \textit{ with respect to } p_i, \tag{9b}$$

$$E_b^i \equiv \frac{D_b^i \cdot p_i}{D_b} \text{ the } \textit{elasticity of demand for joining both sides with respect to } p_i, \tag{9c}$$

$$E_N^i \equiv \frac{\partial N}{\partial u_i} \frac{p_i}{N} \text{ the } \textit{elasticity of total demand for the platform with respect to } p_i, \tag{9d}$$

$$E_i^X \equiv \frac{N_i^X \cdot p_X}{N_i} \text{ the } \textit{elasticity of demand for side } i \textit{ with respect to } p_X, \tag{9e}$$

$$E_b^X \equiv \frac{D_b^X \cdot p_X}{D_b} \text{ the } \textit{elasticity of demand for joining both sides with respect to } p_X. \tag{9f}$$

By Lemma 1 we know the signs of all these elasticities are the same as those of the prices used to define them. E_i^i and E_N^i have the following general property.

Lemma 2 *For $i \in \{1, 2\}$, we have*
i) $E_N^i < E_i^i$ if and only if $p_i > 0$;
ii) $\left|E_N^i\right| < \left|E_i^i\right|$ if and only if $p_i \neq 0$.

An increase in the common value offered to side i generally increases the demand for side i but decreases the demand for side j alone, and as a result, the total demand for the platform does not increase as much as the demand for side i does. Therefore, in absolute value, *the total demand for the platform is generally less elastic than the demand for either one side.*

2.2 Separate pricing as benchmark

For distinction, we **relabel** demand under separate pricing with lower case letters as follows.

$$\begin{aligned}
\text{Demand for side } i \text{ alone} &: \quad d_i(u_1,u_2) \equiv D_i(u_1,u_2,0),\\
\text{Demand for both sides} &: \quad d_b(u_1,u_2) \equiv D_b(u_1,u_2,0),\\
\text{Demand for side } i &: \quad n_i(u_1,u_2) \equiv N_i(u_1,u_2,0) = d_i(u_1,u_2) + d_b(u_1,u_2),\\
\text{Total demand} &: \quad n(u_1,u_2) \equiv d_1(u_1,u_2) + d_2(u_1,u_2) + d_b(u_1,u_2).
\end{aligned}$$

Price elasticities under separate pricing are also relabeled accordingly as follows.

$\epsilon_i^i \equiv E_i^i(p_X = 0)$ the price elasticity of demand for side i, (10a)

$\epsilon_j^i \equiv E_i^j(p_X = 0)$ the elasticity of demand for side j with respect to p_i, (10b)

$\epsilon_n^i \equiv E_N^i(p_X = 0)$ the elasticity of total demand with respect to p_i, (10c)

$\epsilon_b^i \equiv E_b^i(p_X = 0)$ the elasticity of demand for both sides with respect to p_i. (10d)

By (8), the platform's profit under separate pricing is

$$\pi(u_1,u_2) \equiv n_1 v_1 + n_2 v_2 - (\alpha_1 + \alpha_2) n_1 n_2 \tag{11}$$

where v_i is the adjusted price-cost markup defined in (6), only now replacing N_i with the corresponding n_i. Therefore the first-order conditions for the profit-maximizing separate pricing strategy (p_1, p_2) are

$$\begin{cases} n_1^1 v_1 + n_2^1 v_2 = n_1 \\ n_1^2 v_1 + n_2^2 v_2 = n_2 \end{cases} \tag{12}$$

Using the elasticities in (10), the adjusted price-cost margins in (7), and the symmetric property (4) of derivatives of demand functions, we can rewrite these first-order conditions in the following way.

Proposition 1 *The optimal separate pricing strategy* (p_1, p_2) *satisfies*

$$\begin{cases} \epsilon_1^1 r_1 + \epsilon_1^2 r_2 = 1 \\ \epsilon_2^1 r_1 + \epsilon_2^2 r_2 = 1 \end{cases} \tag{13}$$

And the optimal adjusted separate price-cost margins of two sides (r_1, r_2) *are given by*[11]

$$\begin{cases} r_1 = \frac{\epsilon_2^2 - \epsilon_1^2}{\epsilon_1^1 \epsilon_2^2 - \epsilon_2^1 \epsilon_1^2} \\ r_2 = \frac{\epsilon_1^1 - \epsilon_2^1}{\epsilon_1^1 \epsilon_2^2 - \epsilon_2^1 \epsilon_1^2} \end{cases} \tag{14}$$

[11]Note that $\epsilon_1^1 \epsilon_2^2 \neq \epsilon_2^1 \epsilon_1^2$ as $n_i^i > \left| n_j^i \right|$ according to Lemma 1.

Special case with additive values Proposition 1 applies generally when values from two sides are non-additive. In the knife-edge case when nobody derives any extra idiosyncratic value from joining both sides ($t_X = 0$ for all consumers), the optimal rules for separate pricing in (13) can be simplified.

Corollary 1 *With additive values from two sides, the optimal separate pricing strategy* (p_1, p_2) *satisfies, for* $i = 1, 2,$

$$\epsilon_i^i \cdot r_i = 1. \tag{15}$$

This representation is the same as the familiar "Lerner formula" for monopoly pricing in standard two-sided markets by Armstrong (2006). This is not surprising as separate pricing plus additivity implies separable pricing decisions for two sides, such that $\epsilon_1^2 = \epsilon_2^1 = 0$,[12] and therefore the best monopoly price is chosen for each side. We can then interpret the formulas (13) as *generalizations of the Lerner formula* (15), from the case of additive values to a context with non-additive values. We will show in section 4 that the formulas (13) for optimal separate pricing can be further generalized to derive the optimal pricing strategy.

Proposition 1 and Corollary 1 imply that, forgoing bundling, the platform in a mixed two-sided market can do as well as it could if the two sides had no overlap. This convenient result not only makes separate pricing a natural benchmark for studying the incentive to use bundling when the market is mixed (in section 3), but also facilitates a clear comparison between the optimal pricing strategies in standard and mixed two-sided markets (in section 4).

3 Incentive to Bundle

When would the platform have an incentive to bundle the services for different sides, by charging all consumers a common fixed fee, or equivalently by introducing a discount for joining both sides? Starting from separate pricing, if the platform can increase profit by changing u_X, which is equivalent to changing bundle discount p_X according to (2), we will know for sure that it is worthwhile.

[12] With separate pricing ($u_X = 0$) and additivity ($t_X = 0$), the condition for a consumer to join side i (either alone or in addition to side j), $\max\{u_i + t_i, u_i + t_i + u_j + t_j\} \geq \max\{u_j + t_j, 0\}$, holds if and only if $u_i + t_i \geq 0$, independent of u_j and t_j. Therefore the demand for either side becomes independent of the price for the opposite side, such that $\epsilon_1^2 = \epsilon_2^1 = 0$ in Proposition 1.

3.1 General result

Proposition 2 *At the optimal separate pricing strategy, the platform has an incentive to introduce a discount for joining both sides if* ***either*** *of the following two conditions holds*

$$i)\ \epsilon_n^1 \cdot r_1 + \epsilon_n^2 \cdot r_2 \quad < \quad 1,\ or \tag{16a}$$

$$ii)\ \epsilon_b^1 \cdot r_1 + \epsilon_b^2 \cdot r_2 \quad > \quad 1. \tag{16b}$$

And if either condition is reversed, the platform has an incentive to charge a premium for joining both sides when feasible.

Under separate pricing, recall from (10) that ϵ_n^i is the elasticity of total demand with respect to p_i, ϵ_b^i is the elasticity of demand for joining both sides (i.e. bundle demand) with respect to p_i, and r_i is the optimal adjusted price-cost margin for side i, as given by (14). Proposition 2 implies that, given the optimal separate margins for two sides, *the platform would generally be more inclined to introduce a bundle discount if its total demand is less elastic, or if the demand for joining both sides is more elastic*, no matter if these elasticities are measured with the price for side 1 or side 2. The magnitudes of these elasticities are then "moderated" by the optimal separate price-cost margins to jointly determine the incentive to bundle.

Why do these price elasticities determine the platform's incentive to bundle? The reason is that, *ceteris paribus*, a less elastic total demand means that the loss from the discount would be smaller relative to the gain, whereas a more elastic demand for the bundle implies that the gain from introducing a bundle discount would be larger relative to the loss.

Price elasticity of bundle demand We first explain condition (16b) as its intuition is easier to understand. Given the optimal separate prices for two sides, when the platform introduces a discount of one unit for joining both sides, the direct loss is equal to the number of consumers who are already on both sides, d_b. If the demand for joining both sides is very elastic with respect to the price for side 1, say, then the discount will induce a large number of members of side 2 alone to also join side 1, because with the discount it costs them less to join side 1.[13] The elasticity ϵ_b^1 measures exactly the additional revenue brought by these new members of side 1, $d_b^1 \cdot p_1$, as a fraction of the total loss d_b. Similarly, when the demand for joining both sides is very elastic with respect to p_2, the discount will induce a high demand for side 2, and the additional revenue represented as a fraction of the total loss is exactly ϵ_b^2.

What remains to be done in order to gauge the profitability of introducing the discount is to convert these elasticities from "revenue measures" into "profit measures". The

[13] This argument invokes the symmetry of derivatives of demand functions in (4), where $n_1^X = d_b^1$.

adjusted price-cost margins defined in (7) and solved for separate pricing in (14), $r_i = \frac{v_i}{p_i}$, achieve exactly this. Multiplying the elasticities of different sides by the corresponding margins and adding them together, we get condition (16b), which intuitively says that *the total gains from both sides should outweigh the loss when introducing a discount for joining both sides.*

Price elasticity of total demand Now consider condition (16a) and note that the direction of the inequality is reversed compared to (16b). Though mathematically equivalent to (16b), the intuition of (16a) is easier to explain with the following alternative price manipulations:

1) Imposing an additional fixed fee of one unit on every consumer (e.g. as a membership fee); and

2) Lowering both p_1 and p_2, by one unit each.

After these steps, the final price for any consumer on either side alone does not change, but those who join both sides now pay one unit less. Therefore the outcome is equivalent to introducing a one-unit discount for joining both sides. The optimality of the original separate prices implies that step 2) has no impact on profits. Therefore, the impact on profitability from the introduction of a fixed fee in step 1) is exactly the same as that from offering a bundle discount.

The additional fixed fee in step 1) has two effects: a) there is a direct gain from fees collected, equal to the total number of consumers, n; and b) there is a loss in **total** demand, equal to the combined impact on n from raising both p_1 and p_2 by one unit each[14], as the fixed fee applies equally to all users of the platform no matter which side(s) they join.

Therefore, when the total demand n is very **inelastic** with respect to either p_1 or p_2, the fixed fee will induce a **small** loss in total demand. The elasticity ϵ_n^i measures exactly the lost revenue due to the equivalent rise in p_i, $n^i \cdot p_i$, as a fraction of the total gain n. To convert revenues into profits, we again multiply the elasticities by the corresponding adjusted price-cost margins in (14) and add them together to derive condition (16a), which intuitively says that *the total losses from both sides should not exceed the gain when introducing a discount for joining both sides* (or equivalently, when introducing a fixed fee for using the platform).

Flexibility in choice between total demand and bundle demand Aside from their theoretical implications, the conditions in Proposition 2 has several advantages over the existing results in the bundling literature. First of all, they require no information about the underlying distribution of consumers' valuations. And the data they do require are those on price elasticities and price-cost margins, likely obtainable or inferable from usual

[14] In this argument, when raising p_i, the size of side j is held constant so that the effect on total demand is equal to $n^i = \frac{\partial n}{\partial u_i}$.

accounting records. Moreover, these two conditions are equivalent, one using only total demand and the other using only bundle demand. We can therefore choose between them according to the data available. This flexibility makes Proposition 2 potentially more convenient for both empirical testing and practical applications, insofar as when one type of data and elasticity is easier to obtain than the other.

Consider Adobe Acrobat, a PDF software with different versions for readers and writers. Its current pricing strategy involves a freely downloadable version that only allows reading, and a priced full package for both reading and editing. Whereas the software provider may keep data on the sales of the full package, it is unlikely that it will be able to monitor precisely the number of users of its free "Reader" version. It is also virtually impossible for the firm to track interactions between PDF readers and writers, as each action of opening or editing any file with the software constitutes such an interaction. This means it is also unlikely that the firm can infer the number of readers from interactions. As it is probably safe to assume that all PDF writers also read PDFs, the firm can use its sales record as data for the bundle demand, but total demand data are likely inaccurate. Therefore, condition (16b) is preferable. There may also be other examples where (16a) is better, which does not require any information about the bundle (i.e. the intersection of two sides).

The seesaw pattern The dynamics on different sides interact in a "seesaw" pattern in conditions (16a) and (16b). Each condition is an inequality with a sum of two terms, where the impact from different sides are separated. Therefore, *if the constraint that these conditions put on one market side is more stringent, that on the other side will be more lax.*

For instance, in (16b), if the discount results in a *larger* gain from side 1, $\epsilon_b^1 \cdot r_1$, perhaps because the bundle demand is **more elastic** with respect to the price for side 1, then the threshold for side 2, $\epsilon_b^2 \cdot r_2$, will be **lower** for a bundle discount to be profitable, and vice versa.

This seesaw pattern may be particularly useful as it provides flexibility in applications where the two market sides are very **asymmetric** such that one component of the sum in either condition is very large and the other very small, in which case *the inequality may hold with just the component for one market side alone.* We discuss a particular such example after Corollary 2.

Impact of two-sidedness The direct impact of two-sidedness on the incentive to bundle in conditions (16a) and (16b) is isolated into $r_i (= \frac{v_i}{p_i})$, where v_i, the economic value earned from a member of side i as defined in (6), incorporates the network externalities each consumer creates for the opposite side. Therefore, *two-sidedness affects the platform's incentive to bundle through the "moderators" of the adjusted price-cost margins at the* *optimal separate prices.* Further discussion is postponed to the comparative statics analysis in section 5.

3.2 Special case with additive values

Proposition 2 applies generally no matter if consumers regard the two sides as complements, substitutes, or neither. In the simplest case with additivity (i.e. $t_X = 0$ for all consumers), the Lerner equation (15) in Corollary 1 allows us to further simplify conditions (16a) and (16b).

Corollary 2 *With additive values from two sides, at the optimal separate pricing strategy, the platform has an incentive to introduce a discount for joining both sides if* ***either*** *of the following two conditions holds*

$$i)\ \frac{\epsilon_n^1}{\epsilon_1^1} + \frac{\epsilon_n^2}{\epsilon_2^2} < 1,\ or \tag{17a}$$

$$ii)\ \frac{\epsilon_b^1}{\epsilon_1^1} + \frac{\epsilon_b^2}{\epsilon_2^2} > 1. \tag{17b}$$

And if either condition is reversed, the platform has an incentive to charge a premium for joining both sides when feasible.

With additivity, Corollary 2 shows accurately how the platform's incentive to bundle depends on four price elasticities of demand in the market. In particular, *bundling is profitable when,* ***on average****, either the total demand is* ***less than half as elastic*** *as the demand for each side, or the demand for joining both sides is* ***more than half as elastic*** *as the demand for each side.*

Recall that ϵ_i^i represents the price elasticity of demand for side i as defined in (10). In condition (17a), each ratio $\frac{\epsilon_n^i}{\epsilon_i^i}$ measures *how elastic the total demand is, relative to the demand for side i.* We already know from Lemmas 1 and 2 that $0 < \frac{\epsilon_n^i}{\epsilon_i^i} < 1$ for sure. Condition (17a) requires that, *on average, this ratio for each side should be smaller than* $\frac{1}{2}$ *so that their sum does not exceed 1.*

Condition (17b) is equivalent to (17a), but the direction of the inequality is reversed. In (17b) each ratio $\frac{\epsilon_b^i}{\epsilon_i^i}$ measures *how elastic the demand for joining both sides is, relative to the demand for side i.* Condition (17b) requires that, *on average, this ratio for each side should be larger than* $\frac{1}{2}$ *so that their sum is larger than 1.*[15]

The seesaw feature of (17b) may be particularly useful for monopolist power grids of renewable energy, such as the ones mentioned in section 1 in areas without traditional

[15] The existense of $\frac{\epsilon_b^i}{\epsilon_i^i}$ is implied by condition (17b) holding. As $\frac{\epsilon_b^i}{\epsilon_i^i} = \frac{d_b^i/d_b}{n_i'/n_i}$, where $n_i' > 0$, and $d_b^i > 0$, each ratio $\frac{\epsilon_b^i}{\epsilon_i^i}$ exists and is positive whenever $d_b > 0$ and $n_i > 0$. Even when $p_i = 0$, the value of the ratio $\frac{\epsilon_b^i}{\epsilon_i^i}$ can still be calculated by $\frac{d_b^i/d_b}{n_i'/n_i}$, as p_i is canceled out. This is also true for condition (17a).

utility providers. It is likely that the demand for joining such a grid as an electricity **user** is highly **inelastic**, not least because electricity has virtually no good substitute. Therefore the relative elasticity ratio for the electricity user side is likely large, and may even exceed 1. If this happens, it makes sense for the grid to introduce a discount for joining as both a user and a feeder, no matter what demand features are like on the electricity **feeder** side (as each ratio $\frac{\epsilon_b^i}{\epsilon_i^i}$ is positive by Lemma 1). The typical fixed fee to join such a grid that applies to every household regardless of which side(s) they are on is usually one way to implement such a discount[16], whether it is called a membership fee, an installation fee, or whatever other names.

Moreover, as power grids that support fed-in electricity from households are relatively new in most countries, there is likely insufficient historical data to support analysis of the feeders' behavior to begin with. However our conditions (17a) and (17b) may still be applicable as long as the grid can infer elasticities from data of the more familiar user side.

3.3 Special case with symmetric type distribution

Another special case that is relevant to some of the examples mentioned in section 1 is one where the two market sides are symmetric.[17] Consider a stock market where buyers and sellers are on the two sides of a stock exchange. If trading stocks (instead of holding them) is the primary purpose of consumers in the market, the distribution of their idiosyncratic values from the exchange's service facilitating buying stocks may be similar to the distribution of their idiosyncratic values from the selling service. More precisely, we say the consumers' types are distributed symmetrically if the following assumption holds.

Assumption 2 (Symmetric Distribution) *The marginal distribution G_{12} of (t_1, t_2) satisfies: $G_{12}(t_1, t_2) = G_{12}(t_2, t_1)$ for any $(t_1, t_2) \in \mathbb{T}$; and the platform incurs the same cost for consumers on different sides: $f_1 = f_2$.*

Given symmetric distributions, we call a pricing strategy $\mathbf{p} = (p_1, p_2, p_X)$ **symmetric** if it *induces the same common value for both sides according to* (1). As α_1 and α_2 may still be different, the prices for different sides, p_1 and p_2, may also be different. We however know by (6) that a symmetric pricing strategy must yield the same economic value from both sides. With symmetric separate pricing, we also know at least one of the optimal separate prices (p_1 or p_2) must be positive, otherwise the platform makes non-positive profit.

Corollary 3 *Suppose the distribution of types is symmetric such that Assumption 2 holds. At the optimal symmetric separate pricing strategy, suppose without loss of generality that*

[16] See section 1 for a discussion on how two-part tariffs can implement mixed bundling.

[17] In this part we still keep the assumption that the values from both sides are non-additive.

$p_i > 0$ for side i. Then the platform has an incentive to introduce a discount for joining both sides if

$$\epsilon_b^i > \epsilon_n^i. \tag{18}$$

And if this condition is reversed, the platform has an incentive to charge a premium for joining both sides when feasible.

With symmetric consumer types, *bundling is profitable as long as the **bundle** demand is **more elastic** than the **total** demand.* Corollary 3 illustrates that, using real *price* elasticities, our more general result in Proposition 2 is indeed consistent with the intuition of a typical Lerner formula: *A more elastic bundle demand relative to the total demand implies a lower optimal price for the bundle.*

Degree of mixedness We now introduce a measure of "how mixed" a two-sided market is, defined as *the proportion of consumers who join both sides amongst all users of the platform.* Formally, given $\mathbf{u}$, the degree of mixedness is

$$M(\mathbf{u}) \equiv \frac{D_b(\mathbf{u})}{N(\mathbf{u})}.$$

Corollary 4 *Suppose the distribution of types is symmetric such that Assumption 2 holds, and the optimal symmetric separate pricing strategy induces the same common value u^S for each side. Define $\epsilon_{\{i\}}^i \equiv d_i^i \cdot p_i/d_i$ as the price elasticity of demand for side i alone, given the size of the opposite side n_j. Then the platform has an incentive to introduce a discount for joining both sides if*

$$M(u^S, u^S, 0) > \frac{\epsilon_{\{1\}}^1}{\epsilon_1^1} - 1.$$

With symmetric consumer types, *the more **mixed** the market is without bundling, the more likely that introducing a discount for joining both sides is profitable.* There is a definitive threshold for the degree of mixedness without bundling, which is determined by the elasticity of demand for either one side and that for the same side alone. Given the optimal symmetric separate prices, M is equal to $\frac{d_b}{n}$, and $M + 1$ is equal to $\frac{2n_1}{n}$ because $d_b + n = n_1 + n_2 = 2n_1$. This means when M increases, the *proportion of consumers joining either side* also increases. And when the difference between M and the threshold $(\frac{\epsilon_{\{1\}}^1}{\epsilon_1^1} - 1)$ increases, that is, when $(M + 1 - \frac{\epsilon_{\{1\}}^1}{\epsilon_1^1})$ is **larger**, introducing one unit of bundle discount u_X will result in a **larger** gain from consumers on either side than the loss from the discount paid to consumers joining both sides.

Consider the example of a symmetric stock market again, where it is likely that the nature of the market primitives is such that, under separate pricing, a very large proportion of consumers will chose to both buy and sell stocks. That is, M is likely very near

1. Therefore, according to Corollary 4, introducing a membership fee that applies to everyone in addition to the separate prices for buying and selling (which implements a discount for joining both sides) is likely profitable. This is consistent with the prevalence of such two-part tariffs among stock exchanges.

3.4 Detailed comparison with existing results

All elasticities in Proposition 2 and Corollaries 2 through 4 are defined in terms of the **price** on the relevant side, given the size of the opposite side at the optimal separate pricing strategy. In a context of one-sided markets, Long (1984) (on page S243) and Armstrong (2013) (in his Proposition 1) looked at the effect on bundling from alternative elasticities of demand, measured in terms of "*simultaneous and equal percentage increases in price-cost markups of both products*". Though arguably unusual, their carefully defined elasticities can reduce the first-order condition for the optimal separate price of each product to a simple expression of the marginal change in the product's demand, which is in turn useful for determining the sign of the marginal change in profits induced by introducing a bundle discount. Using such elasticities, they show that a bundle discount is profitable *when the bundle demand is more elastic than the demand for either one product.*

However, it is worth emphasizing that *the intuition of their result is very different from that of a typical Lerner formula*, such as (15) by Armstrong (2006), which many people find familiar. Price elasticities are essential to the Lerner intuition because it appeals to people's common sense about the word "elasticity" and the law of demand. *When elasticities are not measured with prices, they do not imply price differences in the same intuitive way.* Although neither Long nor Armstrong interpreted their results using the Lerner intuition, they did not distinguish from it either.[18] What we want to clarify is that, although the bundle discount acts as a price reduction from the sum of stand-alone product prices, its profitability does **not** rely on the **price** elasticity of demand for the bundle to be **greater** than the **price** elasticity of demand for either one product.

Contrary to the intuition that Long (1984) and Armstrong (2013) portray, our Corollary 2 clearly illustrates that, bundling can be profitable even when the bundle demand is **less elastic** than the demand for either one side. Consider, for instance, a market where $\frac{\epsilon_b^1}{\epsilon_1^1} = \frac{\epsilon_b^2}{\epsilon_2^2} = \frac{2}{3}$ given the optimal separate prices, which implies $\left|\epsilon_b^i\right| = \frac{2}{3}\left|\epsilon_i^i\right|$ for both sides, such that the demand for joining both sides is indeed less elastic than that for each side. However, as $\frac{\epsilon_b^1}{\epsilon_1^1} + \frac{\epsilon_b^2}{\epsilon_2^2} = \frac{4}{3} > 1$, condition (17b) implies that a bundle discount will still be profitable.[19]

[18] It is natural to relate the results by Long (1984) and Armstrong (2013) to that of Lerner (1934). In our observation, however, it is not rare for scholars to go so far as equating these results, especially Armstrong's, to the Lerner intuition. This is why we consider it necessary to distinguish between them here.

[19] Using the definition of elasticities by Long (1984) and Armstrong (2013), however, in this situation the elasticity of demand for each side would be 1, and that for the bundle would be $\frac{4}{3}$, which means that,

Our results in Proposition 2 and Corollaries 2 and 3 show that, in determining the incentive to bundle, a more general benchmark price elasticity to compare with the price elasticity of the bundle demand is **not** that for either one product, but should be that for the **total demand**, which includes the demand for both stand-alone products and their bundle and is generally less elastic than the demand for either product according to Lemma 2. The relative magnitudes of these elasticities, measured from different sides, affect the incentive to bundle through the "moderators" of the price-cost margins.

Therefore, a **general** interpretation of the incentive to bundle using the Lerner intuition with **price elasticities** should be that: *A monopolist will be more inclined to lower the price for the bundle (by introducing a bundle discount or a fixed membership fee) either when the bundle demand is more elastic or when the total demand is less elastic.*

Note that all the results in Proposition 2 and Corollaries 2 through 4 remain applicable in the one-sided market context of Long (1984) and Armstrong (2013) (in the integrated supply situation) by setting $\alpha_1 = \alpha_2 = 0$. As there exists no network effect in this degenerate case, all the elasticities we define in (9) and (10) become the standard price elasticities for one-sided markets, without the need to hold the "size of the opposite side" fixed. This special case is discussed in section 6.2.

What would happen to the result by Armstrong (2013) when values were additive? With additivity, the elasticity of demand for either product at the optimal separate prices, with respect to "simultaneous and equal percentage increases in price-cost markups of both products", would **always** be 1; and his condition for a bundle discount to be profitable would simply reduce to: *the bundle demand is elastic*, i.e. the so-defined elasticity of bundle demand is larger than 1. His result would tell us no more than this. Our Corollary 2, however, provides further insights in this situation. Conditions (17a) and (17b) not only show the two ways in which his elasticity concept can be decomposed into two ratios between price elasticities, each for one side of the market, but also shows that there exists a seesaw pattern in the dynamics between different market sides, as described previously.

4 Optimal Pricing Strategy

4.1 General result

We now derive the optimal pricing strategy under mixed bundling, denoted (p_1^*, p_2^*, p_X^*). Using the general price elasticities of demand ($E_i^i, E_i^j, E_b^i, E_i^X$ and E_b^X) in (9), the adjusted price-cost margins in (7), and the symmetry of derivatives of demand functions in (4), we can write the first-order conditions for the optimal pricing strategy in a simple way.

in their terminology, the demand for the bundle would still be considered *more* elastic than that for each side, even though the price elasticities give the exact opposite conclusion. This contrast shows again that their elasticity concept should not be interpreted following the intuition of Lerner formulas.

Proposition 3 *The platform's optimal pricing strategy* (p_1^*, p_2^*, p_X^*) *and the corresponding adjusted price-cost margins* (r_1^*, r_2^*) *satisfy*

$$E_1^1 \cdot r_1^* + E_1^2 \cdot r_2^* - E_1^X = 1 \tag{19a}$$

$$E_2^1 \cdot r_1^* + E_2^2 \cdot r_2^* - E_2^X = 1 \tag{19b}$$

$$E_b^1 \cdot r_1^* + E_b^2 \cdot r_2^* - E_b^X = 1 \tag{19c}$$

Recall that equations (13) in Lemma 1 generalize the Lerner formula (15) by Armstrong (2006) from additive values to non-additive values, for separate pricing in standard two-sided markets. Proposition 3 now further generalizes Lemma 1 from separate pricing in standard two-sided markets to mixed bundling in mixed two-sided markets, with non-additive values from two sides.

Lerner formula (15) says that, with separate pricing, the optimal adjusted price-cost margin on each side, when multiplied by the price elasticity of demand for the same side, should equal 1. The system of three equations in Proposition 3 tells us that, with mixed bundling, the optimal price-cost margins for different market segments - side 1, side 2, and their intersection - adjusted for network benefits, should be such that their weighted sum will equal to 1 when evaluated in relation to each market segment, where each weight is the elasticity of demand for the segment in which the sum is being evaluated, defined with respect to the price of the margin being weighted.

Why does the optimal pricing rule for mixed bundling require that the sums of adjusted price-cost margins, when weighted by the relevant elasticities, should equal 1? The reason is that, with mixed bundling, consumers' demand for each market segment depends on all three price instruments (p_1, p_2, p_X). Whenever one price is altered, the demand for **every** market segment changes. *The elasticities capture the intricate revenue changes in the interrelated demand system, whereas the impact of two-sidedness is again isolated to the adjusted price-cost margins, which convert revenues into profits. Each equation in Proposition 3 represents the net effect of gains and losses from such demand changes, as fractions of the cost of the initial price change, which should equal 1 to balance all gains, losses and costs.*

In this interpretation, the "extra" price-cost margin for the consumers joining both sides is defined as $r_X^* = \frac{-p_X^*}{p_X^*} = -1$, which should be interpreted in incremental terms, as the margins that the platform earns from these consumers when they act on different sides are already accounted for in r_1^* and r_2^*. And r_X^* is negative because the relevant price instrument p_X for this segment is defined as a **deduction** from the sum of separate prices.

Consider equation (19a), for instance, where the platform offers one more util to everyone on side 1 (and as a result p_1 drops by one unit for a given N_2). This has four effects on profit:

1) The total cost of this price cut is equal to the number of members on side 1, N_1.

2) The demand for side 1 increases by N_1^1, bringing additional revenue $N_1^1 \cdot p_1^*$, which represented as a fraction of the cost of the price cut is exactly $E_1^1 = \frac{N_1^1 \cdot p_1^*}{N_1}$. Multiplying it by r_1^* to convert revenues into profits, we get $E_1^1 \cdot r_1^*$, the first term on the left-hand side of (19a).

3) The demand for side 2 increases by N_2^1, bringing additional revenue $N_2^1 \cdot p_2^* = N_1^2 \cdot p_2^*$, which represented as a fraction of the cost of the price cut is exactly $E_1^2 = \frac{N_1^2 \cdot p_2^*}{N_1}$. Multiplying it by r_2^* to convert revenues into profits, we get $E_1^2 \cdot r_2^*$, the second term.

4) The demand for joining both sides increases by D_b^1, where the gain brought by each new consumer has already been accounted for in effects 3) and 4) previously, but in addition the platform incurs a loss of p_X^* in discount paid out to each new consumer, resulting in a total additional **loss** of $D_b^1 \cdot p_X^* = N_1^X \cdot p_X^*$, which represented as a fraction of the cost of the price cut is exactly $E_1^X = \frac{N_1^X \cdot p_X^*}{N_1}$, the last term.

Therefore equation (19a) intuitively says that *the total gains and losses from the price cut induced by offering one more util to side 1 should exactly cancel out, such that these fractions add up to exactly 1.*

If we want to evaluate the weighted sum of the three adjusted price-cost margins, r_1^*, r_2^* and r_X^*, in relation to side 2 instead, the weights we should use are the elasticities of demand for side 2, with respect to p_1, p_2 and p_X, respectively. This sum should equal to 1 when these prices are optimal, which gives us equation (19b). The intuition of equation (19c) is similar. The optimal pricing strategy that solves the system of equations (19a) through (19c) does not look intuitive and is therefore provided in the Appendix.

The following result directly compares Proposition 3 with Lerner formula (15).

Proposition 4 *At the optimal pricing strategy* (p_1^*, p_2^*, p_X^*), *if* $v_i^* > 0$, *we have*

$$|r_i^*| = \left| \frac{p_i^* - (f_i - \alpha_j N_j)}{p_i^*} \right| \geq \frac{1}{|E_i^i|} \text{ if and only if } p_X^* \geq \frac{E_j^i}{E_b^i} \cdot \frac{N_j}{D_b} \cdot v_j^*. \tag{20}$$

Recall that $v_i^* = p_i^* - (f_i - \alpha_j N_j)$ represents the optimal economic value from each consumer on side i, as defined in (6), and E_i^i, E_i^j and E_b^i are the general price elasticities of demand defined in (9). Therefore, *whether the optimal adjusted price-cost margin on one side should be higher or lower than suggested by the familiar Lerner formula depends on a comparison between how well the platform performs on the opposite side and in the intersection of two sides.*

The latter condition in Proposition 4, $p_X^* \geq \frac{E_j^i}{E_b^i} \cdot \frac{N_j}{D_b} \cdot v_j^*$, compares the platform's performance on side j and in the intersection of two sides. It means that, when the platform offers one more util to everyone on side i (and as a result p_i drops by one unit for a given N_j), the total extra discount paid to new consumers who join both sides $(= E_b^i \cdot D_b \cdot p_X^*)$ exceeds the additional economic value earned from new members of side j

($= E_j^i \cdot N_j \cdot v_j^*$). Whenever this happens, the platform incurs a net loss from side j and the intersection of two sides by cutting p_i, and therefore it needs to raise the optimal adjusted price-cost margin on side i in order to compensate for this loss. On the other hand, should the discount paid out to new consumers who join both sides be outweighed by the incremental economic value earned on the opposite side, the platform will optimally set a margin on the current side that is lower than what the Lerner formula would suggest.[20]

The stark contrast between (20) and (15) precisely illustrates the difference between the optimal pricing strategies in mixed and standard two-sided markets.

4.2 Special case with symmetric type distribution

Proposition 5 *Suppose the distribution of types is symmetric such that Assumption 2 holds, and suppose the total number of* ***single-side*** *consumers is non-decreasing in* u_1*, i.e.* $D_1^1(\mathbf{u}) + D_2^1(\mathbf{u}) \geq 0$*, within a neighborhood of the optimal pricing strategy* (p_1^*, p_2^*, p_X^*)*. Then we have*

i) $v_1^* = v_2^* > 0$*;*
ii) $p_X^* > 0$ *if and only if* $|E_b^1| > |E_N^1|$.

E_b^i is the elasticity of demand for joining both sides with respect to p_i, E_N^i is the elasticity of total demand for the platform with respect to p_i as in (9), and v_i^* is the optimal adjusted price-cost markup for side i. With symmetric consumer types, Proposition 5 illustrates a case in which the platform will earn **positive** economic values from both sides, and a sufficient condition for this is that the total number of **single-side** consumers weakly increases in the common value offered to either side.[21] Lemma 1 tells us that the demand for either one side alone generally increases in the common value for the same side and decreases in the common value for the opposite side. $D_1^1 + D_2^1 \geq 0$ therefore requires that when raising the common value for one side, the demand for this side alone increases faster than the demand for the opposite side alone decreases, such that the total number of single-side consumers increases.

When consumer types are symmetric and $D_1^1 + D_2^1 \geq 0$ holds, Proposition 5 also tells us that what determines the sign of the optimal discount for joining both sides is exactly the comparison between the elasticity of the bundle demand and that of the total demand, both measured from the same side and in absolute value. This echoes condition (18) in Corollary 3. The latter says that the platform has an incentive to introduce a

[20] A caveat of this interpretation is that it requires p_i^* to be positive. When $p_i^* < 0$, we have $E_b^i < 0$ and therefore $p_X^* \geq \frac{E_j^i}{E_b^i} \cdot \frac{N_j v_j^*}{D_b}$ is equivalent to $E_b^i D_b p_X^* \leq E_j^i N_j v_j^*$, in which case the interpretation should be reversed. However, Proposition 4 does not depend on a positive p_i^*, as in the ratio expression $\frac{E_j^i N_j v_j^*}{E_b^i D_b}$, p_i^* is canceled out in E_j^i and E_b^i. Also note that because the sign of E_j^i depends on the signs of both N_j^i and p_i^*, and the sign of N_j^i varies at different $\mathbf{u}$ (see footnote 10), it is much more concise to express the second inequality as a constraint on p_X^* than one on v_j^*.

[21] Given symmetric type distribution, $D_1 + D_2$ increasing in u_1 is equivalent to it increasing in u_2.

bundle discount if the bundle demand is more elastic than the total demand given the optimal separate pricing strategy. Proposition 5 says that *this bundle discount will indeed be positive if and only if the bundle demand is* ***still*** *more elastic than the total demand at the optimal mixed bundling strategy.*

5 Comparative Statics

Now we discuss the properties of the equilibrium strategy and outcome in relation to the parameters for the strength of network effects, α_1 and α_2, which embody two-sidedness in our model.[22]

Recall that an interaction between **any** pair of members of two sides creates a total surplus of $\alpha_1+\alpha_2$, where α_i is the share enjoyed by the consumer on side i. Let $\alpha \equiv \alpha_1+\alpha_2$. As the mass of all consumers is normalized to 1 (see section 2.1), the number of all potential pairs of members of different sides is also 1. Therefore α actually measures the *total interaction surplus in the market* that can be created on the platform.[23]

Lemma 3 *The platform's maximized profit is determined by the magnitude of α, the sum of α_1 and α_2, but not by its composition; and the maximized profit strictly increases in α.*

Proof. Inspection of the profit function (5) reveals that α_1 and α_2 affect profit only through their sum α. Substituting α in (5) and rewriting profit as an explicit function of α, we have

$$\hat{\Pi}(\mathbf{u}, \alpha) \equiv \alpha N_1 N_2 - N_1(u_1 + f_1) - N_2(u_2 + f_2) - D_b u_X.$$

Therefore the optimal mixed-bundling common values $\mathbf{u}^*$ can be written as an explicit function of α only,

$$\mathbf{u}^*(\alpha) = (u_1^*(\alpha), u_2^*(\alpha), u_X^*(\alpha)) \equiv \arg\max_{\mathbf{u}} \hat{\Pi}(\mathbf{u}, \alpha). \tag{21}$$

The maximized profit is therefore $\hat{\Pi}(\mathbf{u}^*(\alpha), \alpha)$, a function of α only. By the envelope theorem we have $\frac{d}{d\alpha}\hat{\Pi}(\mathbf{u}^*(\alpha), \alpha) = \frac{\partial}{\partial\alpha}\hat{\Pi}(\mathbf{u}^*(\alpha), \alpha) = N_1 N_2$. ■

It is not surprising that the platform directly benefits from stronger network effects across two market sides. The result that its maximized profit depends on the total interaction surplus α, instead of how it is shared between two sides, is due to the assumption

[22] The generality of our model makes it difficult to conduct a standard comparative statics analysis through total differentiation of the first-order conditions (19). The problem remains rather intractable even when we impose additivity or symmetry. Therefore here we only present analysis of the impact of the relative magnitudes of α_1 and α_2 when their sum is held constant. For some specific applications, however, it may be useful to impose more restrictions that are justifiable given the context, e.g. on the behavior of the demand and/or profit functions, which may make the problem tractable.

[23] This way of defining parameters is common in models that use probability measures as demand, as in Armstrong (2006), Rochet and Tirole (2003 and 2006), and Weyl (2010). Weyl discusses several examples in his section I.B. where such "aggregate" parameters can be applied in practice.

that everyone on the same side has the **same constant probability** of interacting with everyone on the opposite side, independent of the benefit that each consumer derives from each interaction. This assumption is also used by Armstrong (2006) and Rochet and Tirole (2003 and 2006).

One implication of Lemma 3 is that, the platform may still earn a high profit when the network effect is **weak or even negative** in **one** direction, as long as the network effect in the **opposite** direction is **strong** enough. This is likely true for information exchange platforms including email and text messaging systems, where the receivers usually impose a stronger network effect on senders than conversely. Especially when there are many "junk message" senders (e.g. advertisers), the share of surplus that a receiver obtains from an average message might even be negative, in which case the platform may need to subsidize receivers to maintain their participation. However, as long as senders enjoy high network benefits from interactions, the platform can still make a profit. This may help to explain the existence of messaging systems abounding with junk messages.

Given the total interaction surplus α, although the relative magnitudes of network effects in different directions do not affect profit, they however **do** affect the optimal prices for different market segments, as summarized in the following result.

Proposition 6 *Given the total interaction surplus* α*,*

i) the optimal bundle discount p_X^* *is* ***invariant*** *in* α_1 *and* α_2*;*

ii) the final price that a consumer who joins both sides pays, $(p_1^* + p_2^* - p_X^*)$*,* ***increases*** *in* α_i *if and only if* i *is the* ***smaller*** *side in equilibrium;*

iii) the optimal price for side i*,* p_i^**, increases in* α_i*.*

For a given level of total interaction surplus, Proposition 6 first shows an interesting result that *the optimal discount those who join both sides receive (or the premium they pay, if feasible) does not change no matter how the surplus is shared between two sides.* This is again due to the assumption of "constant probability of interaction" between two sides. It does not mean that two-sidedness is irrelevant here, though. As shown in (21), *what determines the optimal level of discount is the level of joint network effects in both directions, not their structure.* This result shows the extent to which two-sidedness changes the traditional mixed bundling strategy.

Proposition 6 also says that the final price those who join both sides pay to the platform **increases** in the share of the surplus enjoyed by the **smaller** side of the market (and decreases in the share enjoyed by the larger side). The sizes of different sides matter here because they act as multipliers to the share of surplus a consumer enjoys per interaction when calculating the total network benefit obtained from each side according to (1). For the benefit that the marginal consumer who joins both sides receives on the **smaller** side, an **increase** in the share of surplus for the smaller side is **doubly good** because this higher share is then multiplied by the size of the opposite side, which is also **larger**.

As a result, the benefit she receives on the smaller side then increases at a **higher** rate than the decrease in the benefit she receives on the larger side. Hence her total pre-price utility increases, enabling the platform to charge a higher price without affecting her participation. Similarly, for the benefit that she receives on the **larger** side, a **decrease** in the share of surplus per interaction for that side is **doubly bad**, as this lower share will be multiplied by the size of the opposite side which is also smaller. Therefore the resulting pre-price utility decreases, implying a **lower** final price. Point iii) is rather straightforward.

Possible application Proposition 6 is perhaps useful for understanding the difference in price structures between similar platforms that monopolize different geographic areas. As we do not have a specific real-life example in mind, let us continue to use the text messaging example mentioned previously, and consider two such platforms in two hypothetical countries, A and B. Suppose, for illustration only, there are reasons to believe that the market conditions are very similar in these countries such that they have the same potential total interaction surplus. The only difference is that the receivers in country A enjoy a **smaller** share of surplus per message than their counterparts in B, maybe because there are more advertisers sending out annoying junk messages in A than in B.

If both platforms use two-part tariffs of the kind discussed in section 1 that implement mixed bundling, point i) of Proposition 6 implies that the fixed fee they charge for access to their systems should be **identical** (which implements the same bundle discount). Point iii) implies that senders in country A will pay a **higher** price than their counterparts in country B, and receivers in A pay a **lower** price (or receive a larger subsidy) than their counterparts in B. And if in equilibrium there are **more senders than receivers** in both countries, for instance, point ii) then implies that consumers who join both sides in country A pay a **lower** final price than their counterparts in B, as the smaller side (i.e. receivers) in A enjoys a smaller share of surplus per message. Therefore, despite the fact that sending a message in A brings a **larger** share of surplus, those who both send and receive messages in A will pay an overall **lower** final price than their counterparts in B, as the final price for receivers is much lower in A.

6 Welfare, One-Sided Market, and Numerical Example

6.1 Welfare maximization

The welfare in this market, as measured by the unweighted sum of profit (5) and consumer surplus (3), is $W(\mathbf{u}) \equiv \Pi(\mathbf{u}) + V(\mathbf{u})$. The first-order conditions for the welfare-maximizing common values can be expressed using the corresponding welfare-maximizing mixed-bundling strategy (p_1^W, p_2^W, p_X^W) as

$$\begin{aligned} E_1^1 \cdot r_1^W + E_1^2 \cdot r_2^W &= E_1^X \\ E_2^1 \cdot r_1^W + E_2^2 \cdot r_2^W &= E_2^X \\ E_b^1 \cdot r_1^W + E_b^2 \cdot r_2^W &= E_b^X \end{aligned}$$

where $r_i^W = \frac{p_i^W - (f_i - \alpha_j N_j)}{p_i^W}$ represents the adjusted price-cost margin for side i that maximizes welfare. Because given (p_1^W, p_2^W, p_X^W), $E_i^X = \frac{N_i^X}{N_i} \cdot p_X^W$ by (9), the unique solution to the system of equations above is $p_X^W = r_1^W = r_2^W = 0$, which gives us the following result.

Proposition 7 *The welfare-maximizing mixed bundling strategy* (p_1^W, p_2^W, p_X^W) *entails zero net economic values from consumers on either side and zero discount for consumers joining both sides, that is,*

$$p_1^W = f_1 - \alpha_2 N_2,\ p_2^W = f_2 - \alpha_1 N_1,\ p_X^W = 0.$$

Not surprisingly, the socially optimal pricing strategy involves no bundling, and each consumer is charged according to her "net cost" to the system, just as in the standard two-sided markets studied by Armstrong (2006).

6.2 One-sided market

In a "degenerate" case of our model where $\alpha_i = 0$, we have $u_i = -p_i$ according to (1), so that without network effects, any common value the platform provides for side-i consumers must be achieved at the cost of an exactly equal price cut. Moreover, $v_i = p_i - f_i$ so that the net economic value earned from side i is simply the standard price-cost markup.

When $\alpha_1 = \alpha_2 = 0$, the market falls back to a standard one-sided market, and our model becomes a standard one for mixed bundling with two generic products. The platform has profit $\Pi = N_1 v_1 + N_2 v_2 - D_b p_X$, which is exactly the standard one-sided market profit function given mixed bundling strategy (p_1, p_2, p_X), except that each demand function is defined with respect to $(u_1 = -p_1, u_2 = -p_2, u_X = p_X)$. All of the previous analyses and results still hold in this case.

6.3 Numerical example

Suppose the strengths of network effects are $\alpha_1 = \alpha_2 = 0.25$, the consumer type (t_1, t_2) is **uniformly** distributed on unit square $[-0.3, 0.7] \times [-0.3, 0.7]$, and $t_X = 0$. $f_1 = f_2 = 0$. This is therefore an example with both additive values from two sides and symmetric consumer types. All calculations are done via Scientific WorkPlace 5.0 and the results are presented in Table 2.

Table 2. Numerical Example
($\alpha_1 = \alpha_2 = 0.25$, $(t_1, t_2) \sim \mathbf{U}[-0.3, 0.7] \times [-0.3, 0.7]$, $t_X = 0$, $f_1 = f_2 = 0.$)

	Separate Pricing		Mixed Bundling	Comment
u_i	-0.233	$>$	-0.463	Negative values to either side
u_X	0		0.735	
u_b	-0.467	$<$	-0.192	
p_i	0.350		0.633	
p_X	0		0.735	M.B. discount large
"Transaction" fee	0.75	$>$	-0.149	M.B. subsidizes transactions
"Membership" fee	0		0.735	M.B. membership fee high
N_i	0.467	$<$	0.680	
D_i	0.249	$>$	0.007	
D_b	0.218	$<$	0.673	
N	0.716	$>$	0.687	M.B. attracts fewer users overall
$M = \frac{D_b}{N}$	30.4%	$<$	98.0%	M.B. market close to fully mixed
Π	0.327	$<$	0.367	
v_i	0.467	$<$	0.804	
v_b	0.933	$>$	0.872	
$r_i = \frac{v_i}{p_i}$	1.333		1.269	
ϵ_i^i or E_i^i	0.75		0.931	
$\frac{1}{\epsilon_i^i}$ or $\frac{1}{E_i^i}$	1.333		1.074	M.B. $\left\|\frac{v_i}{p_i}\right\| > \left\|\frac{1}{E_i^i}\right\|$
ϵ_b^i	2.637		–	S.P. $\left\|\epsilon_i^i\right\| < 2\left\|\epsilon_b^i\right\|$
$\frac{\epsilon_b^1}{\epsilon_1^1} + \frac{\epsilon_b^2}{\epsilon_2^2}$	7.031		–	S.P. $\frac{\epsilon_b^1}{\epsilon_1^1} + \frac{\epsilon_b^2}{\epsilon_2^2} > 1$

Optimal mixed bundling At the optimal mixed bundling strategy, the platform is providing a negative common value ($u_i = -0.463$) to consumers on both sides, but the consumers who join both sides get a huge extra value ($u_X = 0.735$). As some consumers have very asymmetric idiosyncratic values (t_i may be as high as 0.7 and t_j may be as low as -0.3), some of them still choose a single side despite the low common value. However their proportion is very small. Almost all consumers join both sides, and the degree of mixedness is 98%.

The strategy that achieves this outcome is quite interesting. The platform charges members of each side a price of $p_i = 0.633$, but offers those who join both sides a huge discount of $p_X = 0.735$, which in fact results in a lower final price for the bundle than that for either side ($p_1 + p_2 - p_X = 0.532 < p_i$). This strategy would only seem feasible in practice if consumers have no "free disposal" of their idiosyncratic values t_i, which is true in our model and in Rochet and Tirole (2006) and Weyl (2010), where negative "membership benefits" may represent some "fixed cost" involved in gaining access to the platform.

However, when implemented with an equivalent two-part tariff in "two stages", the assumption of "no free disposal" appears less problematic. The platform can charge a membership fee $A = 0.735$ for access to the platform (including both sides), and then offers a subsidy of 0.149 for each actual transaction or interaction made. Using the notation discussed in section 1, here the "additional fee" for side i is $a_i = -0.149 \times N_j = -0.102 (= p_i - A)$, which is actually a subsidy. For consumers who will "incur" negative t_i once they make a transaction on side i, sticking to only side j may be preferable despite the subsidy that could be obtained from side i.

The transaction subsidy essentially enhances the positive network effects ($\alpha_i = 0.25$) and makes the market slightly "more two-sided". Enforceability of such a two-part tariff requires the platform's ability to monitor users' membership status and interactions or transactions, which works in telecommunications and crowd funding markets, for instance.

At the optimal mixed bundling strategy, the platform earns high and relatively "even" economic values from all consumers, in that $v_i = 0.804$ and $v_1 + v_2 - p_X = 0.872$ are rather large and quite close in magnitude, relative to the parameters. The two inequalities in Proposition 4 both hold. The adjusted price-cost margin is higher than the inverse of the price elasticity of demand for either side.

Optimal separate pricing The optimal separate pricing strategy involves a relatively low price ($p_i = 0.35$) for either side, which still results in a negative common value ($u_i = -0.233 < 0$) offered to both sides. The platform earns twice more economic value from double-side consumers ($v_1 + v_2 = 0.933$) than from single-side consumers ($v_i = 0.467$). The adjusted price-cost margin at optimality on side i, $\frac{v_i}{p_i}$, is exactly equal to the inverse of the price elasticity of demand for side i, as the Lerner formula (15) suggests. The equivalent two-part tariff involves a positive fee for each interaction (0.35) on either side, contrary to the optimal transaction subsidy under mixed bundling.

Separate pricing attracts more users overall, but the degree of mixedness is much lower (30.4%). At the optimal separate pricing strategy, $\frac{\partial \Pi}{\partial u_X} = 0.218 > 0$, indicating that offering a positive bundle discount (or charging a positive membership fee if a two-part tariff is used) will be profitable. Alternatively, it can be calculated that condition (17b) holds with ">" as $\frac{\epsilon_b^1}{\epsilon_1^1} + \frac{\epsilon_b^2}{\epsilon_2^2} = 7.031$; and the condition in Corollary 4 becomes $30.4\% > 0$ as $\epsilon_{\{1\}}^1 = \epsilon_1^1 = 0.75$. Indeed we see that the profit of the platform is higher under mixed bundling ($0.367 > 0.327$).

7 Conclusion

The platform pricing problem in mixed two-sided markets has two prominent features: **network effect** from "two-sidedness" and **multiple products** due to "mixedness". Our model and results show that: 1) the impact of two-sidedness can be isolated into the price-cost margins of two sides, in which the network benefits that each consumer creates for the opposite side need to be accounted for; 2) these adjusted price-cost margins and various familiar price elasticities of demand suffice to explain both the platform's incentive to bundle and the optimal pricing strategy; and 3) the optimal pricing rule for mixed two-sided markets generalizes the familiar Lerner formulas for one-sided markets and standard two-sided markets.

We have endeavored to provide a theoretical framework that is as general as possible, as a first step toward understanding platforms' pricing behavior in mixed two-sided markets. Based on this model, we find the following three directions interesting for future work. First, it may be useful to work out the theoretical implications of the general model in several special cases, such as one where the two sides exhibit substitutability or complementarity, depending on the specific applications intended. Second, it would be ideal to extend the monopoly model to one that includes platform competition. Third, it would be superb to test empirically whether these results fit market data in some of the examples we have discussed.

For a mixed two-sided platform in real life, the choice among different theoretical pricing strategies may entail much more than just a change in pricing. It may also involve completely different product/system design of the platform. Note that mobile phone networks do not have to offer a SIM card that has both calling and receiving functions. It is technologically feasible to make separate devices that only have either one function. In fact, in the latter half of the twentieth century, telecommunications networks in many countries offered "pagers" which are mobile devices that can only receive messages. Nor does eBay have to provide all users with both buyer and seller services. It is up to the platform whether or not to design the product and/or system to allow the callers to receive calls, to allow the lenders to borrow money, or to allow the buyers to sell. The choice of optimal pricing strategy may be the result of a particular design, and can certainly also be the reason why the platform is designed the way it is. For burgeoning mixed two-sided markets such as renewable energy power grids, understanding the problem of platform pricing should be a necessary step in the process of platform design, not least because profitability will foster sustainability as the industry moves forward.

8 References

Adams, W. J. and J. L. Yellen (1976), "Commodity Bundling and the Burden of Monopoly", *Quarterly Journal of Economics*, XC, pp. 475-98.

Armstrong, M. (2006): "Competition in Two-Sided Markets", *Rand Journal of Economics*, 37(3), pp. 668-691.

Armstrong, M. (2013): "A More General Theory of Commodity Bundling", *Journal* *of Economic Theory*, 148, pp. 448-472.

Bakos, Y. and E. Brynjolfsson (1999): "Bundling Information Goods: Pricing, Profits, and Efficiency", *Management Science*, 45(12), pp. 1613-1630.

Caillaud, B. and Jullien, B. (2003): "Chicken and Egg: Competition Among Intermediation Service Providers", *Rand Journal of Economics*, 34(2), pp. 309-328.

Chao Y. and T. Derdenger (2013): "Mixed Bundling in Two-Sided Markets in the Presence of Installed Base Effects", *Management Science*, 59(8), pp. 1904-1926.

Chen, Y. and M. H. Riordan (2013), "Profitability of Product Bundling," *International Economic Review*, 54(1), pp. 35-57.

Fang, H. and P. Norman (2006): "To bundle or not to bundle", *Rand Journal of Economics*, 37(4), pp. 946-963.

Gao, M. (2010): "When to Allow Buyers to Sell: Bundling in Mixed Two-Sided Markets", Chapter 1 of Ph.D. dissertation, London Business School.

Hanson, W. and R. K. Martin (1990): "Optimal Bundle Pricing", *Management Science*, 36(2), pp. 155-174.

Lerner, Abba P. (1934): "The Concept of Monopoly and the Measurement of Monopoly." *Review of Economic Studies*, 1(3): pp. 157-75.

Long, J.B. Jr. (1984): "Comments on Gaussian Demand and Commodity Bundling", *Journal of Business*, 57(1), Part 2, pp. S235-S246.

Manelli, A. M. and Vincent, D. (2006): "Bundling as an optimal selling mechanism for a multiple-good monopolist", *Journal of Economic Theory*, 127, pp. 1-35.

McAfee, R. P., McMillan, J. and Whinston, M. D. (1989): "Multiproduct Monopoly, Commodity Bundling, and Correlation of Values", *Quarterly Journal of Economics*, 104, pp. 371-383.

Rochet, J.C. and Tirole, J. (2003): "Platform Competition in Two-Sided Markets." *Journal of the European Economic Association*, 1, pp. 990–1029.

Rochet, J.C. and Tirole, J. (2006): "Two-Sided Markets: A Progress Report", *Rand Journal of Economics*, 37(3), pp. 645-667.

Rochet, J.C. and Tirole, J. (2008): "Tying in two-sided markets and the honor all cards rule", *International Journal of Industrial Organization,* 26(2008), pp. 1333–1347.

Rysman, M. (2009): "The Economics of Two-Sided Markets", *Journal of Economic Perspectives*, 23(3), pp. 125-143.

Schmalensee, R. (1984): "Gaussian Demand and Commodity Bundling", *Journal of Business*, 57(1), Part 2, pp. S211-S230.

Weyl, G. (2010): "A Price Theory of Multi-Sided Platforms." *American Economic Review*, 100, pp. 1642–1672.

9 Appendix - Proofs

Proof of Lemma 1

First consider $i = 1$, $j = 2$. And the proof for the case where $i = 2$ and $j = 1$ is similar. Given the distribution of (t_1, t_2, t_X),

$$\begin{aligned} N_1(\mathbf{u}) &= \Pr[\max\{u_1 + t_1, u_1 + t_1 + u_2 + t_2 + u_X + t_X\} \geq \max\{u_2 + t_2, 0\}] \\ &= \Pr[u_1 + t_1 + \max\{u_2 + t_2 + u_X + t_X, 0\} \geq \max\{u_2 + t_2, 0\}] \end{aligned}$$

When u_1 or u_X increases, N_1 also increases (in the set-theoretic sense), such that $N_1^1 > 0$ and $N_1^X = D_b^1 > 0$.

$$D_1(\mathbf{u}) = \Pr[u_1 + t_1 \geq \max\{u_2 + t_2, u_1 + t_1 + u_2 + t_2 + u_X + t_X, 0\}]$$

For values of (t_1, t_2, t_X) such that $u_1 + t_1 + u_2 + t_2 + u_X + t_X \geq \max\{u_2 + t_2, 0\}$, we have $D_1(\mathbf{u}) = \Pr[0 \geq u_2 + t_2 + u_X + t_X]$ such that D_1 is invariant in u_1 and decreasing in u_2 and u_X; and for all other values of (t_1, t_2, t_X), D_1 is given by $\Pr[u_1 + t_1 \geq \max\{u_2 + t_2, 0\}]$ which is increasing in u_1, decreasing in u_2, and invariant in u_X. Therefore we conclude that $D_1^1 > 0$, $D_1^2 < 0$, and $D_1^X < 0$. Similarly, we have $D_2^1 < 0$. Therefore we have

$$N_1^1 - N_2^1 = D_1^1 + D_b^1 - D_2^1 - D_b^1 = D_1^1 - D_2^1 > 0.$$

$$N(\mathbf{u}) = \Pr[\max\{u_1 + t_1, u_2 + t_2, u_1 + t_1 + u_2 + t_2 + u_X + t_X\} \geq 0]$$

When u_1 or u_X increases, N also increases, such that $N^1 > 0$ and $N^X > 0$.

$$D_b(\mathbf{u}) = \Pr[u_1 + t_1 + u_2 + t_2 + u_X + t_X \geq \max\{u_1 + t_1, u_2 + t_2, 0\}]$$

When u_X increases, D_b also increases, such that $D_b^X > 0$.

As we also have $N_1^1 + N_2^1 = N^1 + D_b^1 > 0$ which implies $N_1^1 > -N_2^1$, therefore we can say $N_1^1 > \left|N_2^1\right|$.■

Proof of Lemma 2

First consider $i = 1$, and the proof for the case when $i = 2$ is similar.

As $N = N_1 + D_2$, we have $N^1 = N_1^1 + D_2^1$. By Lemma 1 we know $N^1 > 0$, $N_1^1 > 0$, and $D_2^1 < 0$. Therefore

$$NN_1^1 - N^1 N_1 = (N_1 + D_2)N_1^1 - (N_1^1 + D_2^1)N_1 = D_2 N_1^1 - D_2^1 N_1 > 0$$

which means $\frac{N_1^1}{N_1} > \frac{N^1}{N} (> 0)$. Therefore $E_1^1 = \frac{N_1^1}{N_1} p_1 > E_N^1 = \frac{N^1}{N} p_1$ if and only if $p_1 > 0$, and $\left|E_1^1\right| = \frac{N_1^1}{N_1} |p_1| > \left|E_N^1\right| = \frac{N^1}{N} |p_1|$ if and only if $p_1 \neq 0$. (Note that when $p_i = 0$, we have $E_i^i = E_N^i = 0$.)■

Proof of Proposition 2

i) Condition (16a):

From (6) and (8) we know

$$\frac{\partial}{\partial u_X}\Pi(\mathbf{u}) = N_1^X v_1 + N_2^X v_2 - D_b - D_b^X u_X \tag{22}$$

(p_1, p_2) is the optimal separate pricing strategy, which induces common values ($u_1^{SP} = \alpha_1 n_2 - p_1, u_2^{SP} = \alpha_2 n_1 - p_2$). As $d_b = n_i - d_i$ implies $d_b^i = n_i^i - d_i^i$; $n = d_i + n_j$ implies $n^i \equiv \frac{\partial n}{\partial u_i} = d_i^i + n_j^i$; and $n_1^2 = n_2^1$, we have

$$\begin{aligned}
\frac{\partial}{\partial u_X}\Pi(u_1^{SP}, u_2^{SP}, 0) &= N_1^X \cdot v_1 + N_2^X \cdot v_2 - d_b \\
&= d_b^1 \cdot v_1 + d_b^2 \cdot v_2 - d_b \\
&= (n_1^1 - d_1^1)v_1 + (n_2^2 - d_2^2)v_2 - d_b \\
&= n_1^1 v_1 + n_2^2 v_2 - d_1^1 v_1 - d_2^2 v_2 - d_b \\
&= (n_1 - n_2^1 v_2) + (n_2 - n_1^2 v_1) - d_1^1 v_1 - d_2^2 v_2 - d_b \\
&= n_1 + n_2 - d_b - n_2^1 v_1 - d_1^1 v_1 - n_1^2 v_2 - d_2^2 v_2 \\
&= n - n^1 v_1 - n^2 v_2 \\
&= n(1 - \frac{n^1 \cdot p_1}{n} \cdot \frac{v_1}{p_1} - \frac{n^2 \cdot p_2}{n} \cdot \frac{v_2}{p_2}) \\
&= n(1 - \epsilon_n^1 r_1 - \epsilon_n^2 r_2)
\end{aligned}$$

where $n_1^1 v_1 = n_1 - n_2^1 v_2$ follows from the first-order condition (12) for p_1, and similarly for p_2. Therefore $\frac{\partial}{\partial u_X}\Pi(u_1^{SP}, u_2^{SP}, 0) \gtrless 0$ iff $\epsilon_n^1 r_1 + \epsilon_n^2 r_2 \lessgtr 1$.

ii) Condition (16b):

$$\begin{aligned}
\frac{\partial}{\partial u_X}\Pi(u_1^{SP}, u_2^{SP}, 0) &= N_1^X \cdot v_1 + N_2^X \cdot v_2 - d_b && \text{(23a)} \\
&= d_b^1 \cdot v_1 + d_b^2 \cdot v_2 - d_b && \text{(23b)} \\
&= d_b(\frac{d_b^1 \cdot p_1}{d_b} \cdot \frac{v_1}{p_1} + \frac{d_b^2 \cdot p_2}{d_b} \cdot \frac{v_2}{p_2} - 1) && \text{(23c)} \\
&= d_b(\epsilon_b^1 \cdot r_1 + \epsilon_b^2 \cdot r_2 - 1) && \text{(23d)}
\end{aligned}$$

where (23a) follows from (22) by setting $u_X = 0$; (23b) follows from (4); and finally in (23d), ϵ_b^i is defined in (10).■

Proof of Corollary 3

Given symmetric distribution, denote $u^S (= u_1 = u_2)$ the same optimal separate common value for each side, as induced by the optimal symmetric separate pricing strategy.

By (6) and (12) we know

$$v_1 = v_2 = \frac{n_1}{n_1^1 + n_2^1} \tag{24}$$

As $n_1^1 + n_2^1 = n^1 + d_b^1 > 0$, we have $v_1 = v_2 > 0$. Suppose without loss of generality that $p_1 > 0$. Therefore

$$\begin{aligned}
\epsilon_b^1 &> \epsilon_n^1 \Leftrightarrow \frac{d_b^1 p_1}{d_b} > \frac{n^1 p_1}{n} \Leftrightarrow \\
\frac{d_b^1}{d_b} &> \frac{n^1}{n}(> 0) \Leftrightarrow \frac{d_b^1}{d_b} > \frac{n^1 + d_b^1}{n + d_b} = \frac{n_1^1 + n_2^1}{2n_1} = \frac{1}{2v_1} \Leftrightarrow \\
\frac{d_b^1 p_1}{d_b} &> \frac{p_1}{2v_1} \Leftrightarrow \epsilon_b^1 > \frac{1}{2r_1} \Leftrightarrow 2\epsilon_b^1 \cdot r_1 > 1
\end{aligned}$$

which is exactly condition (16b) with symmetric types.■

Proof of Corollary 4

Given symmetric distribution, we have (24), and therefore (23b) becomes

$$\frac{\partial}{\partial u_X}\Pi(u^S, u^S, 0) = \frac{2n_1^1 d_1 - d_1^1 n - d_b d_2^1}{n_1^1 + n_2^1}$$

where $n_1^1 + n_2^1 = n^1 + d_b^1 > 0$, and $d_2^1 < 0$. Therefore

$$\frac{\partial}{\partial u_X}\Pi(u^S, u^S, 0) > \frac{2n_1^1 d_1 - d_1^1 n}{n_1^1 + n_2^1}$$

As $M = \frac{d_b}{n} \geq \frac{\epsilon_{\{1\}}^1}{\epsilon_1^1} - 1$ implies $\frac{d_1^1/d_1}{n_1^1/n_1} = \frac{\epsilon_{\{1\}}^1}{\epsilon_1^1} \leq 1 + M = 1 + \frac{d_b}{n} = \frac{2n_1}{n}$, we have $\frac{d_1^1}{n_1^1} \leq \frac{2d_1}{n}$, which implies $2n_1^1 d_1 - d_1^1 n \geq 0$. Therefore $\frac{\partial}{\partial u_X}\Pi(u^S, u^S, 0) > 0$.■

Proposition 3: Analytical solution of the optimal mixed bundling strategy

Suppose the optimal mixed bundling strategy is (p_1^*, p_2^*, p_X^*). We will first solve the system of equations in Proposition 3 for the optimal r_1^*, r_2^*, and p_X^*.

Substituting $E_i^X = \frac{N_i^X \cdot p_X^*}{N_i}$ and $E_b^X = \frac{D_b^X \cdot p_X^*}{D_b}$ in the system of equations in Proposition 3, we have

$$\begin{aligned}
E_1^1 \cdot r_1^* + E_1^2 \cdot r_2^* - \frac{D_b^1}{N_1} p_X^* &= 1 \\
E_2^1 \cdot r_1^* + E_2^2 \cdot r_2^* - \frac{D_b^2}{N_2} p_X^* &= 1 \\
E_b^1 \cdot r_1^* + E_b^2 \cdot r_2^* - \frac{D_b^X}{D_b} p_X^* &= 1
\end{aligned}$$

The solution is therefore

$$\begin{aligned}
r_1^* &= \frac{1}{c}[\frac{D_b^X}{D_b}(E_1^2 - E_2^2) + \frac{D_b^1}{N_1}(E_2^2 - E_b^2) + \frac{D_b^2}{N_2}(E_b^2 - E_1^2)] \\
r_2^* &= \frac{1}{c}[\frac{D_b^X}{D_b}(E_2^1 - E_1^1) + \frac{D_b^1}{N_1}(E_b^1 - E_2^1) + \frac{D_b^2}{N_2}(E_1^1 - E_b^1)] \\
p_X^* &= \frac{1}{c}[E_b^1(E_1^2 - E_2^2) + E_b^2(E_2^1 - E_1^1) + E_1^1E_2^2 - E_2^1E_1^2] \\
\text{where } c &= \frac{D_b^X}{D_b}(E_1^2E_2^1 - E_1^1E_2^2) + \frac{D_b^1}{N_1}(E_b^1E_2^2 - E_2^1E_b^2) + \frac{D_b^2}{N_2}(E_1^1E_b^2 - E_1^2E_b^1)
\end{aligned}$$

As $r_i = \frac{p_i-(f_i-\alpha_j N_j)}{p_i}$ according to (7), the corresponding optimal price for side i is

$$p_i^* = \frac{\alpha_j N_j - f_i}{r_i^* - 1}.\blacksquare$$

Proof of Proposition 4

Consider $i = 1$ and $j = 2$. By (19a), when $v_1^* > 0$, we have $|r_1^*| \geq \frac{1}{|E_1^1|}$ if and only if $E_1^X \leq E_1^2 \cdot r_2^*$. By $N_1^X = D_b^1 > 0$, and $r_2^* = \frac{v_2^*}{p_2^*}$, we have $E_1^X \leq E_1^2 \cdot r_2^*$ if and only if $p_X^* \geq \frac{E_2^1 N_2 v_2^*}{E_b^1 D_b}$. The case when $i = 2$ and $j = 1$ is symmetric.$\blacksquare$

Proof of Proposition 5

i) $v_i^* > 0$:

Given symmetric distribution, we solve the system of equations in Proposition 3. The optimal symmetric mixed bundling strategy (u_1^*, u_2^*, u_X^*) must induce the same optimal common value for each side: $u_1^* = u_2^*$. By (6) and (19a) we know

$$v_1^* = v_2^* = \frac{N_1 + N_1^X p_X^*}{N_1^1 + N_2^1} = \frac{N_1 + D_b^1 p_X^*}{N^1 + D_b^1}$$

substituting in (19c) and we have $2\frac{D_b^1}{D_b}v_1^* = 2\frac{D_b^1}{D_b}\frac{N_1+D_b^1 p_X^*}{N^1+D_b^1} = 1 + E_b^X$, which after reduction gives us the solution

$$p_X^* = \frac{D_b^1 N - D_b N^1}{D_b^X N^1 - D_b^1 N^X} = \frac{D_b^1 N - D_b N^1}{D_b^X D_s^1 - D_b^1 D_s^X}$$

where $D_s \equiv D_1 + D_2$ (and therefore $N = D_s + D_b$). Thus

$$v_1^* = \frac{N_1 + D_b^1 p_X^*}{N^1 + D_b^1} = \frac{D_b^X D_1 - D_b D_1^X}{D_b^X N^1 - D_b^1 N^X}$$

where the numerator must be positive as $D_b^X > 0$ and $D_1^X < 0$. Therefore $v_1^* > 0$ if and only if $D_b^X N^1 - D_b^1 N^X = D_b^X D_s^1 - D_b^1 D_s^X > 0$. As $D_s^X = D_1^X + D_2^X < 0$, we know

$D_s^1 = D_1^1 + D_2^1 > 0$ implies $D_b^X D_s^1 - D_b^1 D_s^X > 0$ which in turn implies $v_1^* > 0$.

Note that by definition, $p_1^* = v_1^* + (f_1 - \alpha_2 N_2)$, so $v_1^* > 0$ does not guarantee that $p_1^* > 0$. The sign of p_1^* also depends on $(f_1 - \alpha_2 N_2)$.

ii) $p_X^* > 0$ if and only if $\left|E_b^1\right| > \left|E_N^1\right|$:

Recall the definitions $E_b^1 = \frac{D_b^1 p_1}{D_b}$ and $E_N^1 = \frac{N^1 p_1}{N}$.

As $p_X^* = \frac{D_b^1 N - D_b N^1}{D_b^X D_s^1 - D_b^1 D_s^X}$, where $D_s^1 = D_1^1 + D_2^1 > 0$ implies that the denominator is positive, therefore its sign depends on the numerator

$$D_b^1 N - D_b N^1 = \frac{D_b N}{p_1^*}(E_b^1 - E_N^1)$$

where $D_b N > 0$. Therefore p_X^* has the same sign as $\frac{1}{p_1^*}(E_b^1 - E_N^1)$.

Necessity: When $p_X^* > 0$ (a positive discount for joining both sides), we know at least one of the optimal separate prices (p_1^* or p_2^*) must be positive, otherwise the platform makes non-positive profit. Suppose without loss of generality that $p_1^* > 0$. Therefore $E_b^1 > 0$, $E_N^1 > 0$, and $E_b^1 > E_N^1$, which implies $\left|E_b^1\right| > \left|E_N^1\right|$.

Sufficiency: Suppose $\left|E_b^1\right| > \left|E_N^1\right|$, then $p_1^* \neq 0$. If $p_1^* > 0$, we have $E_b^1 > E_N^1 > 0$, and therefore $p_X^* > 0$. If $p_1^* < 0$, we have $E_b^1 < E_N^1 < 0$, and therefore $p_X^* > 0$.

Note that we clearly also have $p_X^* = 0$ if and only if $\left|E_b^1\right| = \left|E_N^1\right|$.■

Proof of Proposition 6

For a given level of $\alpha = \alpha_1 + \alpha_2$, the optimal mixed-bundling common values are given by $\mathbf{u}^*(\alpha) = (u_1^*(\alpha), u_2^*(\alpha), u_X^*(\alpha))$ according to (21), which also determine the demand for all market segments, $N_1(\mathbf{u}^*(\alpha))$, $N_2(\mathbf{u}^*(\alpha))$ and $D_b(\mathbf{u}^*(\alpha))$.

Point i) is immediate from (2) and (21), as the optimal bundle discount is uniquely determined by $p_X^* = u_X^*(\alpha)$.

Point i): Note that by (1) and (2) we have

$$\begin{aligned} p_1^* + p_2^* - p_X^* &= \alpha_1 N_2 - u_1^* + \alpha_2 N_1 - u_2^* - p_X^* \\ &= \alpha_1 N_2 - u_1^* + (\alpha - \alpha_1) N_1 - u_2^* - p_X^* \\ &= \alpha_1 (N_2 - N_1) + \alpha N_1 - u_1^* - u_2^* - u_X^* \end{aligned}$$

which implies that, given α, the final price increases in α_1 if and only if $N_2 - N_1 > 0$.

Similarly, as we also have $p_1^* + p_2^* - p_X^* = \alpha_2(N_1 - N_2) + \alpha N_2 - u_1^* - u_2^* - u_X^*$, the final price increases in α_2 if and only if $N_1 - N_2 > 0$.

Point iii): From (1) we have $p_1^* = \alpha_1 N_2 - u_1^*$ and $p_2^* = \alpha_2 N_1 - u_2^*$, which immediately implies point i).■