Ming Gao

Economic behaviour and decision making: theories of two-sided markets, multiproduct pricing and weighting for cumulative prospect theory

Thesis

Gao, Ming

(2010)

Economic behaviour and decision making: theories of two-sided markets, multiproduct pricing and weighting for cumulative prospect theory.

Doctoral thesis, University of London: London Business School.

DOI: https://doi.org/10.35065/JVKK8675

Economic Behaviour and Decision Making: Theories of Two-Sided Markets, Multiproduct Pricing and Weighting for Cumulative Prospect Theory

Ming Gao
Economics Department
London Business School
Regent's Park, London
NW1 4SA, UK
Email: mgao.phd2005@london.edu
Homepage: phd.london.edu/mgao/

July 2010

Contents

Introduction of Thesis

This thesis studies the microeconomic theories of two-sided markets, multiproduct pricing, and decision making in risky choice situations.

In the first part of the thesis, we focus on a special kind of two-sided markets, where participants can act on both the buying side and the selling side of the market, which we call "mixed two-sided markets". The literature on two-sided markets has assumed that buyers cannot sell and sellers cannot buy. In real life, however, many markets are mixed, with examples ranging from telecommunication to stock exchange. We provide a general model for mixed two-sided markets. We observe that in practice, platforms in such markets often use a "hybrid" bundling strategy: A common membership fee gives access to both buying and selling services, while the individual transaction fees are separated. The main impact of this strategy is what we call the "two-part-tariff effect": Imposing a small bundled membership fee on top of any transaction fees always leads to zero first-order losses in the demand of consumers who use both services, thus enabling the platform to extract more surplus from them. When this positive effect dominates the losses in demand from single-service users, hybrid bundling dominates unbundled sales. We present general conditions that guarantee such an outcome.

In the second part of the thesis, we show that the two-part-tariff effect still applies when such tariffs are used in a more general context of multiproduct pricing. We consider a monopolist provider of n (> 1) products who uses two-part tariffs consisting of a membership fee that is common to all consumers, and separate prices for different product bundles. We show that the change in demand for any bundle of $k \in [1, n]$ products caused by imposing an extra membership fee on top of any separate pricing strategy is proportional to the membership fee to the power of k. Therefore a small extra membership fee has no first-order impact on the demand for any multiproduct bundles, which enables the firm to extract more consumer surplus. When this positive effect dominates the loss of single-product demand, two-part tariff dominates separate pricing. We present conditions that guarantee such an outcome, which generalize McAfee, McMillan and Whinston (1989)'s result from two products to multiple products. The two-part-tariff effect provides a new multiproduct perspective for the wide application of two-part tariffs, complementary to the classical "single-product" efficiency-related explanation. Our results suggest that two-part tariffs can achieve multidimensional price discrimination and should be subject to similar regulatory scrutiny as other discriminatory strategies, such as mixed bundling.

The theories discussed in the first two parts address market situations where participants face deterministic decision problems. However, many if not most decision making processes involve uncertainty. The third part of the thesis focuses on people's economic behavior in risky choice situations. In this context, the cumulative prospect theory (CPT) by Tversky and Kahneman (1992) has been accepted as one of the best descriptive models that reconcile, within one unified model, the major phenomena that violate standard utility models. However, the inverse S-shaped weighting of cumulative probabilities posited in CPT causes difficulties in preference representation, which hinders its application in wider situations of risky choice. We propose a simplified weighting function for CPT, the (β, c) model, which plays a similar role in models with risky choice as that of the quasi-hyperbolic discounting function in models with intertemporal choice. The (β, c) model has a weighting function that is linear with slope smaller than 1 on the open interval $(0, 1)$, jumps down to 0 at end point 0, and jumps up to 1 at end point 1. It achieves highly tractable utility representation for CPT whilst preserving the basic tenets of CPT. It by construction can explain all four major phenomena of risky choice violating the standard model that CPT was developed to reconcile, including reference dependence and certainty effect. It also allows Bayesian updating (with distortions at certainty) which CPT cannot accommodate. We systematically examine the (β, c) representation of discrete and continuous lotteries, and provide four applications which illustrate that the (β, c) model is a useful work horse to analyze implications of preferences exhibiting certainty effect and reference dependence in standard models.

More detailed critical discussions are provided in each part of the thesis.

Part I

When to Allow Buyers to Sell? Bundling in Mixed Two-Sided Markets*

Ming Gao†

First version: January 2008. This version: July 2010.

Abstract

In many examples of two-sided markets, ranging from telecommunication to stock exchange, participants act on both sides of the market. Whereas the literature has assumed buyers cannot sell and sellers cannot buy, we provide a model for two-sided markets where users can appear on both sides, which we call "mixed two-sided markets". We observe that in practice, platforms in such markets often use a "hybrid" bundling strategy: A common membership fee gives access to both buying and selling services (in a "bundle"), while the individual transaction fees are separated. The main impact of this strategy is what we call the "two-part-tariff effect": Imposing a small bundled membership fee on top of any transaction fees always leads to zero first-order losses in the demand of consumers who use both services, thus enabling the platform to extract more surplus from them. When this positive effect dominates the losses due to single-service users, hybrid bundling dominates unbundled sales. We present general conditions that guarantee such an outcome.

Key Words: mixed two-sided market, hybrid bundling, platform pricing, telecommunication

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

---

*I thank Jean-Pierre Benoît, Ramon Casadesus-Masanell, Francesca Cornelli, James Dow, Dino Gerardi, Emeric Henry, Bruno Jullien, Eric Maskin, Preston McAfee, Joao Montez, Marco Ottaviani, Emre Ozdenoren, Alessandro Pavan, Stefano Sacchetto, Dezsoe Szalay, Jean Tirole, and Leonard Waverman for valuable discussions and comments. I also thank participants at CEPR Applied IO Conference 2009, EARIE Annual Conference 2009, La Pietra-Mondragone Doctoral Workshop 2009, ECORE/Université Libre de Bruxelles, International Industrial Organization Conference 2010, and the Second Brazilian Workshop of the Game Theory Society 2010. I am particularly indebted to my supervisors, Jean-Pierre Benoît and Leonard Waverman, and Emeric Henry for their guidance, encouragement and support.

†Economics Department, London Business School, Regent's Park, London NW1 4SA, UK. Email: `mgao.phd2005@london.edu`. Homepage: `phd.london.edu/mgao/`.

Contents

1 Introduction

In many two-sided markets, participants act on both sides of the market. In online trading, for instance, people can quite freely buy and sell as they please; in telecommunication networks, most subscribers both make and receive phone calls; in many if not most kinds of financial intermediation, traders (or account-holders) are allowed to both buy and sell (or to both borrow and lend), such as in stock exchange, securities brokerage and social lending.[1]

However, a common feature of the existing models in the literature on two-sided markets is that there is no overlap between the two sides (see, for instance, Caillaud and Jullien (2003), Armstrong (2006) and Rochet and Tirole (2003 and 2006)). While this assumption suits the classic examples studied in the literature, such as credit card, video games and media, it does not apply to the examples we mentioned previously.

In this paper, we study 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 the two-sided market **standard**. We provide a model which extends the standard model of Rochet and Tirole (2006) to the mixed case. Figure 1 illustrates the difference between these two kinds of markets.

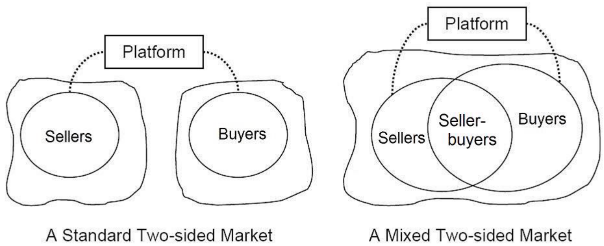


Figure 1: Standard and Mixed Two-Sided Markets

Since some consumers may want to use the services a mixed two-sided platform provides to *both* sides, the platform can employ *multiproduct* pricing strategies that are irrelevant in standard two-sided markets. In particular, it can bundle its selling and buying services and provide them to all potential users.[2] "Mixedness" therefore brings a multi-product/bundling perspective to two-sided markets and links the two literatures.

We study a new kind of bundling strategy that has not been studied in the literature but is widely used by mixed two-sided platforms. We name it **hybrid bundling** and it consists of two parts: a *bundled membership fee*, which gives access to both selling

[1] Other examples include software that allows users to both create and view files in certain formats (such as text-processing software, PDF software, computer-aided design software, etc.) and information exchange platforms that allow users to both post (or send) and view messages (such as bulletin boards, online forums, social networking websites, user-generated content platforms, etc.).

[2] This is what we mean by bundling in this paper. It does not include, for instance, an eBay seller selling two products in bundles on the platform, or eBay bundling two kinds of services designed for sellers but not buyers.

and buying services, and two *separate transaction fees*, which apply to the two parties involved in each transaction. For instance, a typical pay-monthly phone tariff consists of a monthly line rental for connection to both calling and answering services, plus separate per-minute fees for making and receiving calls; stock exchanges (and many other financial intermediaries) also usually charge users an annual membership fee for access to both buying and selling services, plus separate commissions on transactions.

This strategy is different from the two classic categories of bundling strategies in the literature: *pure bundling* (only selling multiple goods in bundles but not separately) and *mixed bundling* (providing two goods both separately and in "one-to-one" bundles). Schmalensee (1984), Fang and Norman (2006) and Banal-Estanol and Ottaviani (2007) show that pure bundling can yield higher profit than unbundled sales because consumers' valuation of the bundle is generally *less dispersed* than their valuation of each separate good and hence pure bundling can increase the probability of trade at certain prices. Schmalensee (1984) and McAfee, McMillan and Whinston (1989) show that mixed bundling can increase profit because by lowering the price of the bundle below the sum of the prices of separate goods (that is, lowering *one* price) a monopolist can increase demands for both goods (that is, increase *two* demands).[3]

Hybrid bundling is a more subtle way of bundling. If we divide both selling and buying services into two "stages", obtaining membership and making transactions, hybrid bundling can be interpreted as *pure bundling of memberships of two sides and unbundled transactions of two sides*. This is not the same as classic pure bundling, as transactions on different sides are separated and discretionary. It is also different from mixed bundling, as each user may make *multiple* transactions on either side, and the number of her sales need *not* equal that of her purchases (that is, they need *not* match "one to one").[4]

Rochet and Tirole (2006) show that a standard two-sided platform can maximize profits by an unbundled-sales type strategy involving separate membership fees and separate transaction fees for two sides. We show that when the market is mixed, hybrid bundling can further increase platform profits, because the bundled membership fee achieves multiproduct price discrimination.

We uncover the key mechanism of hybrid bundling, which we call the **two-part-tariff effect**: *Imposing a small bundled membership fee on top of any transaction fees always leads to zero first-order decrease in the demand of consumers who use both services* (whom we call **seller-buyers**). This effect enables the platform to extract more surplus from seller-buyers with hybrid bundling. When such gains dominate the losses due to

[3]Some mixed two-sided platforms use these classic strategies, such as Microsoft's Word software which is a pure bundle of both editing and reading functions that could potentially be separated as in the case of PDF software.

[4]The classic models of mixed bundling mentioned previously assume that each bundle consists of *one* unit of each product, and each consumer consumes at most *one* unit of each product. In section 6 we provide a detailed discussion of these differences.

single-service users, hybrid bundling dominates unbundled sales.

We present a series of conditions that guarantee such an outcome, where three main factors favoring hybrid bundling emerge: the **degree of mixedness**, measured by the proportion of seller-buyers amongst all users (depending on prices and consumers' valuations); the scope economies in provision of two services; and negative correlation between consumers' valuations of two services. We summarize the general insights from our results below:

i) The profitability of hybrid bundling depends critically on market mixedness. There needs to be a positive number of seller-buyers under unbundled sales in order for the two-part-tariff effect to increase platform profits.

ii) The two-part-tariff effect generally increases the degree of mixedness. Under certain technological assumptions that fit real-life financial intermediaries and telecommunication networks, the higher is the market's degree of mixedness under unbundled sales, the more likely that hybrid bundling will dominate unbundled sales.

iii) Hybrid bundling exploits scope economies better than unbundled sales, even if the platform saves equal costs under both of these strategies when providing two services to one same user compared to providing them to two different users. This is because hybrid bundling increases the proportion of seller-buyers among users, who are more cost-effective than single-service users.

iv) When consumers' valuations of two services are weakly negatively correlated (including independence), hybrid bundling always dominates.

v) The desirability of hybrid bundling does *not* depend particularly on users' ability to take account of the service they value *less* in their decisions to join the platform. That is, even if all users base their membership decisions solely on the higher one between their expected benefit from selling and that from buying, all the results above remain robust.

Point iv) confirms similar results on bundling found in "one-sided" markets.

Point i), ii), iii) and v) are new findings of this paper, which are consistent with platform pricing behaviors observed in real life. For instance, in different financial intermediation markets, the popularity of hybrid bundling tends to "correlate" with the degree of mixedness: stock exchanges normally face a very high proportion of users who both sell and buy, and they normally use hybrid bundling; while social lending platforms (which facilitate lending and borrowing among individuals, such as Zopa.com) usually have a much lower proportion of users who both lend and borrow, and they usually employ unbundled sales. We postulate that the distributions of valuations in these markets are such that, even if stock exchanges also used unbundled sales, the degree of mixedness they face would still be much higher than in social lending. Then point ii) tells us that they are more likely to find hybrid bundling a more attractive strategy.

The remainder of this paper is structured as follows. Section 2 describes the model. In section 3 we discuss unbundled sales as a benchmark. Section 4 examines hybrid bundling and compares it to unbundled sales, where we present the main results. In section 5 we discuss two extensions and check the "robustness" of results under a different production technology and under bounded agent rationality. Section 6 discusses in detail how our results relate to the existing literatures on bundling and multiproduct pricing. Section 7 concludes and discusses possible extensions.

2 Model

We study a market with one monopoly platform and a continuum $[0, 1]$ of agents. The agents may want to trade a certain kind of good, and the only way to do this is to use the buying and selling services provided by the platform.

The market works in three stages:

Stage 1 - The platform chooses a strategy (unbundled sales or hybrid bundling) and announces relevant prices;

Stage 2 - Each agent observes the prices and decides whether to use the platform only as a buyer, only as a seller, as both a buyer and a seller (by paying the relevant membership fees), or not to use it at all;

Stage 3 - Users who decide to use the platform trade.

2.1 The platform

The platform has two strategies:

Unbundled sales - It separates the services provided to sellers and buyers, and sets different prices for different sides. The relevant price vector is denoted $\mathbf{P}_U \equiv (a_U^S, a_U^B, A^S, A^B) \in \mathbb{R}^{+4}$, where A^S and A^B are the respective membership fees for each seller and buyer, and a_U^S and a_U^B are the respective usage fees per transaction borne by the seller and buyer involved.[5]

Anyone who wishes to use both services under this strategy must pay both membership fees.

Hybrid bundling - The platform offers every potential user access to both selling and buying services. The relevant price vector is denoted $\mathbf{P}_H \equiv (a_H^S, a_H^B, A) \in \mathbb{R}^{+3}$, where A is the **bundled membership fee** for access to both services, and a_H^S and a_H^B are usage fees per transaction borne by the seller and the buyer involved, respectively.

Notice that, under hybrid bundling, any user can make multiple numbers of transactions on either side of the market. This distinguishes hybrid bundling from mixed bundling strategies studied in the literature, which only allow unit-demand of either product by each consumer (e.g. McAfee, McMillan and Whinston (1989)).

[5]We use superscripts to denote market sides (S for Seller and B for Buyer) and subscripts to denote strategies (U for Unbundled sales, and H for Hybrid bundling).

As we discuss in detail in section 4.1, neither unbundled sales nor hybrid bundling can replicate the other unless the bundled membership fee A is zero. When A is positive, they will generally induce different demands.

We focus on these two strategies for three reasons:

i) Unbundled sales is an optimal strategy in standard two-sided markets (see Rochet and Tirole (2006)). Hybrid bundling appears to be the prevalent type of strategy employed by real-life mixed two-sided platforms, and it is relevant only if the market is mixed. Given the difficulties in fully characterizing the optimal multiproduct nonlinear pricing strategy in our setting (mainly due to the network effects that link the demands of two sides, which we specify shortly), hybrid bundling becomes a natural first step.

ii) Within the set of pure and mixed bundling strategies with two-part tariffs, since each service (buying or selling) can be divided into two stages (membership and transaction), and either pure bundling or mixed bundling could be done in each stage (or even across stages), the total number of ways to bundle is quite large. Since not all such strategies seem very practical (such as pure bundling of transactions of two sides and unbundled memberships), we have to draw the line somewhere. Thus, the two prominent strategies again provide a natural starting point.[6]

iii) These two strategies reflect the contrast between standard and mixed two-sided markets that we want to show: mixedness allows the platform to exploit multiproduct price discrimination.

Cost These two strategies may involve different cost structures for the platform. A dedicated analysis of the impact of such differences is provided in section 5. In the main parts of the paper, we ignore such differences and use the following assumption:

Assumption C ***(Cost structure)*** *In both strategies, if a user only uses the selling (resp. buying) service, the platform incurs fixed cost* F^S *(resp.* F^B*) for her; if a user uses both services, the platform incurs fixed cost* F *for her. In addition, the platform incurs a variable cost of* c *per transaction. We assume* $\min(F^S, F^B, F, c) \geq 0$*, and* $F \in [\max(F^S, F^B), F^S + F^B]$.

This is a standard assumption except that it allows for the possibility of scope economies in the platform's provision of two services. Note that this cost structure applies to both

[6] We do not consider any kind of bundling of transactions. We do not study formally the case of mixed bundling of memberships, either. However we can show that our model does accommodate the latter, the caveat being that in our model there always exist strategies of pure bundling of memberships that replicate the demand and profit under mixed bundling of memberships. As such replication is not a general property in different models, the results we show are actually only about pure bundling of memberships, and should not be invoked in other circumstances. While our model explains platforms' incentive to use the hybrid bundling strategy, further work is needed to understand why they refrain from using mixed bundling of memberships in real life.

strategies and hence the possible scope economies do not necessarily favor either strategy *ex ante*.

The platform maximizes profits.

2.2 The agents

The whole set of agents is denoted Ω. An agent $\omega \in \Omega$, by using the platform, gets a constant benefit v_ω^S from each sale, and a constant benefit v_ω^B from each purchase.[7] The pair of valuations (v_ω^S, v_ω^B) is called agent ω's **type**.

Agents have different types. The distribution of types is given by G on a support $\mathbb{V} \subseteq \mathbb{R}^2$. Each agent knows her own type, but the platform only knows G.

Assumption 1 ***(Distribution)*** *The type space $\mathbb{V}$ is weakly convex and has full dimension on $\mathbb{R}^2$, and distribution G has no atoms on $\mathbb{V}$.*

This is a relatively weak version of the standard regularity assumption used in the literature of bundling and multiproduct pricing.[8] By using a two-dimensional type space, our model links the literature on two-sided markets to the one on bundling.[9]

We denote the joint density function g, the conditional distributions $G^{S|B}$ and $G^{B|S}$, marginal distributions G^S and G^B, and their respective density functions $g^{S|B}$, $g^{B|S}$, g^S and g^B.

In stage 2 of the game, each agent decides whether to use the platform's services or not. We call this a **membership decision**. We say an agent becomes a user (or member) of the platform by paying the relevant membership fee(s). The set of all users is denoted Y.

In stage 3 of the game, each user decides whether to just sell, to just buy, or to both sell and buy, by paying the relevant transaction fee(s).[10] We call this a **trading decision**. Any agent who uses the selling (respectively, buying) service is called a **seller** (respectively, **buyer**).

We denote the set of all sellers S and the set of all buyers B. If a user both sells and buys, she's called a **seller-buyer**, otherwise she's either a **pure seller** or a **pure**

[7] Notice v_ω^S (resp. v_ω^B) is the "net" benefit from a sale (resp. purchase) for seller (resp. buyer) ω, including any payments transferred from the buyer (resp. seller). Only the fees charged by the platform are not included in v_ω^S and v_ω^B. Similar "net valuation" parameters are used by Rochet and Tirole (2006).

[8] We do not restrict $\mathbb{V}$ to be $\mathbb{R}^{+2}$ as in McAfee, McMillan and Whinston (1989), or a closed set as in Armstrong (1996). For example, our Assumption 1 allows G to be the bivariate normal distribution, in which case $\mathbb{V}$ is the whole $\mathbb{R}^2$ space.

[9] In some models of standard two-sided markets (e.g. Rochet and Tirole (2003 and 2006)), an agent may have a *different* kind of "two-dimensional" type, e.g. (B^S, b^S), where B^S is her "membership benefit" as a seller, and b^S is her "usage benefit" as a seller. However, such a type still only pertains to one side of the market, and thus bears no relevance to bundling of services on two sides.

[10] Any user must trade, because having paid a non-negative membership fee, no-trade would give her non-positive net payoff which is at least weakly worse than her outside option (0 payoff), in which case we assume she would not have joined.

buyer. We denote the set of pure sellers PS, the set of pure buyers PB, and the set of seller-buyers SB.

These sets are also called **market segments**. The relationships among them are summarized as follows:

$$S = PS \cup SB; \quad B = PB \cup SB;$$
$$SB = S \cap B; \quad Y = S \cup B = PS \cup PB \cup SB.$$

Ex post, a mixed two-sided market is one where $SB \neq \varnothing$.

Now we define the number of agents in these sets as the probability measures of the relevant sets, including:

The number of sellers: $N^S \equiv \Pr[\omega \in S]$.

The number of buyers: $N^B \equiv \Pr[\omega \in B]$.

The number of seller-buyers: $N^{SB} \equiv \Pr[\omega \in SB]$.

The number of pure sellers: $N^{PS} \equiv \Pr[\omega \in PS] = N^S - N^{SB}$.

The number of pure buyers: $N^{PB} \equiv \Pr[\omega \in PB] = N^B - N^{SB}$.

The number of users: $N \equiv \Pr[\omega \in Y] = N^S + N^B - N^{SB}$.

To the platform, these numbers are **demands** from different market segments.

2.3 Transaction

The network effects between the two sides are crucial determinants of the platform's demands. We model them in the standard way used in the two-sided market literature, as summarized in the following assumption.[11]

Assumption 2 ***(Network effects)*** *There is a constant positive probability for each buyer to trade with each seller, which is simplified to 1.*

Therefore, each agent's surplus from using either service is increasing in the number of agents using the other service. In particular, suppose there are N^S sellers and N^B buyers, and the transaction fees are a^S and a^B, then Assumption 2 implies that an agent of type (v^S, v^B) will get total selling surplus (gross of any membership fee):

$$u^S \equiv (v^S - a^S)N^B$$

and total buying surplus (gross of any membership fee):

$$u^B \equiv (v^B - a^B)N^S$$

[11] Rochet and Tirole (2003 and 2006) use this same assumption. Caillaud and Jullien (2003) and Armstrong (2006) also use an equivalent assumption of a constant probability of matching between agents from two sides. We use this assumption here to facilitate a straightforward extension from earlier models and a direct comparison between results.

As mentioned in Rochet and Tirole (2006), Assumption 2 also implies that the total volume of transactions on the platform is $N^S \cdot N^B$, which we will use later in the profit function.[12]

Furthermore, we make the following standard assumption in the bundling literature:

Assumption 3 ***(Separability)*** *There exists no complementarity or substitutability in consumption of the two services. In particular, a seller-buyer's surplus gross of any membership fee is* $u^S + u^B$.

A user's net surplus is then the total surplus she gets from trading minus the membership fee(s) she pays. For example, a seller-buyer's net surplus is $u^S + u^B - A^S - A^B$ under unbundled sales, while it is $u^S + u^B - A$ under hybrid bundling.

2.4 Demand

We make the following normalization assumption:

Assumption 4 ***(Normalization)*** *Any agent's outside option (i.e. not joining the platform) gives zero net surplus. Whenever an agent is indifferent between joining and not joining, she does not join.*

Given the platform's choice of strategy and the relevant prices announced, each agent makes her membership decision by comparing her net expected surplus from trading with the outside option, taking account of the *expected* number of people on either side of the market. Individual agent's membership decisions then aggregate into the market demand. Like in Rochet and Tirole (2006), we assume that, given the platform's pricing strategy, all the agents have the same "rational" expectation of the induced market demand. If such demand exists, we call it the **equilibrium demand** (or equilibrium outcome).[13]

Assumption 4 combined with the definitions of u^S and u^B implies that, *whenever one of* N^S *and* N^B *is zero, the other must also be zero.* It is due to the assumption that all the value that a user can get on the platform comes from the interaction with the opposite side of the market - there is no value in just being on the platform when there is no one to trade with. Therefore, $(N^S = 0, N^B = 0)$ *is always an equilibrium outcome regardless of the pricing strategy chosen and the prices announced.* This "trivial" equilibrium outcome is quite common in models of two-sided markets and it is of little interest to our goal in this paper, as it always yields zero platform profit. In the remaining parts of the paper we focus without loss of generality on the more interesting equilibrium outcome with

[12] Concerns might arise regarding self-trade of a seller-buyer. However this is not a problem in our model, since each individual agent is infinitesimal and her self-trade has zero measure in the accounting of the volume of transactions based on probability measures.

[13] Given the platform's pricing strategy, the equilibrium demand is actually the "rational expectations equilibrium" of the subgame that all agents play.

positive demand. Our task is then to identify such demands[14] under different strategies and compare the platform's profits.

We start with unbundled sales.

3 Unbundled Sales

Suppose the platform uses unbundled sales and announces $\mathbf{P}_U = (a_U^S, a_U^B, A^S, A^B)$.[15] Then agents allocate into market segments in the following way:

$$\begin{aligned} PS_U &= \{\omega \mid u^S > A^S, u^B \leq A^B\} \\ PB_U &= \{\omega \mid u^S \leq A^S, u^B > A^B\} \\ SB_U &= \{\omega \mid u^S > A^S, u^B > A^B\} \end{aligned}$$

This means the demands from the two sides of the market at $\mathbf{P}_U$ are:

$$\begin{aligned} N_U^S &= \Pr[\omega \in PS_U \cup SB_U] = \Pr[u^S - A^S > 0] = 1 - G^S(\tfrac{A^S}{N_U^B} + a_U^S) \\ N_U^B &= \Pr[\omega \in PB_U \cup SB_U] = \Pr[u^B - A^B > 0] = 1 - G^B(\tfrac{A^B}{N_U^S} + a_U^B) \end{aligned} \qquad (1)$$

Given $\mathbf{P}_U$, N_U^S and N_U^B are simultaneously determined by system (1).

Notice that neither N_U^S nor N_U^B depends directly on the prices applicable to the opposite side - this is due to consumption separability in Assumption 3. This means that given $\mathbf{P}_U$, an agent actually makes *two separate membership decisions*, each regarding one side. Moreover, the criterion she uses in each decision is exactly the *average per-transaction charge* on the relevant side: $\frac{A^S}{N_U^B} + a_U^S$ for selling and $\frac{A^B}{N_U^S} + a_U^B$ for buying.

Figure 2 illustrates agent allocation under unbundled sales.[16]

[14] Existence of such positive equilibrium demands is guaranteed by Assumption 5 below.

[15] The analyses and results under unbundled sales are very similar to those in Rochet and Tirole (2006). This is no surprise because the unbundled sales strategy "treats" a mixed two-sided market exactly the same way as a standard two-sided market.

[16] The distribution shown in Figure 2 has support $\mathbb{V} = [0, \bar{v}^S] \times [0, \bar{v}^B]$. The same support is also used in Figures 3, 4 and 5.

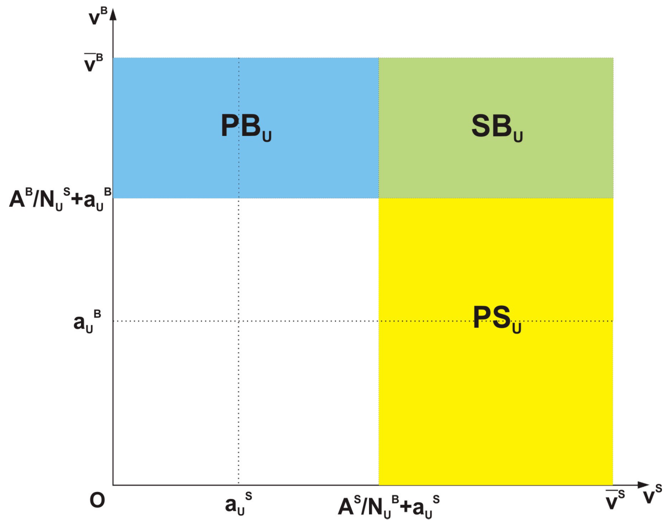


Figure 2: Agent Allocation under Unbundled Sales

The platform's profit at $\mathbf{P}_U$ is then:

$$\Pi_U(\mathbf{P}_U) \equiv \underbrace{(A^S - F^S)N_U^{PS}}_{\text{pure sellers' membership}} + \underbrace{(A^B - F^B)N_U^{PB}}_{\text{pure buyers' membership}} + \underbrace{(A^S + A^B - F)N_U^{SB}}_{\text{seller-buyers' membership}} + \underbrace{(a_U^S + a_U^B - c)N_U^S N_U^B}_{\text{transactions}}$$

Lemma 1 *(**Redundancy of unbundled-sales membership fees**) For any* $\mathbf{P}_U = (a_U^S, a_U^B, A^S, A^B) \in \mathbb{R}^{+4}$*, there exists a degenerate strategy* $\mathbf{P}_{U0} = (a_U^{S\prime}, a_U^{B\prime}, 0, 0) \in \mathbb{R}^{+4}$*, such that* $\mathbf{P}_{U0}$ *exactly replicates the demand and profit under* $\mathbf{P}_U$*.*[17]

Proof. For any $\mathbf{P}_U = (a_U^S, a_U^B, A^S, A^B) \in \mathbb{R}^{+4}$, let N_U^S and N_U^B denote the demands induced by $\mathbf{P}_U$ which must solve (1) above. Now define

$$a_U^{S\prime} \equiv a_U^S + \frac{A^S}{N_U^B}$$
$$a_U^{B\prime} \equiv a_U^B + \frac{A^B}{N_U^S}$$

Then $\mathbf{P}_{U0} = (a_U^{S\prime}, a_U^{B\prime}, 0, 0)$ is in $\mathbb{R}^{+4}$ and N_U^S and N_U^B also solve (1) at $\mathbf{P}_{U0}$. It is easy to check that $\Pi_U(\mathbf{P}_{U0}) = \Pi_U(\mathbf{P}_U)$. ■

The redundancy in the unbundled-sales strategy is due to Assumption 2. As shown in Figure 2, we only need *one* threshold on each side of the market to determine all market segments, thus one price for each side will suffice.

From now on, we focus without loss of generality on the degenerate vector $\mathbf{P}_{U0}$ as the effective unbundled-sales strategy. And we make the following regularity assumption:

Assumption 5 *(**Existence of optimal unbundled-sales strategy**) There exists an optimal degenerate unbundled-sales strategy denoted by* $\mathbf{P}_0^* \equiv (a^{S*}, a^{B*}, 0, 0)$ *on the interior of support* $\mathbb{V}$*, at which the unbundled-sales profit is maximized and positive.*[18]

[17] Rochet and Tirole (2006) also mentioned the redundancy in the unbundled-sales strategy.

[18] This assumption simply serves the same purpose as the standard regularity assumptions used in the

4 Hybrid Bundling

Suppose the platform uses hybrid bundling strategy $\mathbf{P}_H = (a_H^S, a_H^B, A)$. In this section we assume agents are fully rational, and use "backward induction" while making membership and trading decisions.

Thus, given $\mathbf{P}_H = (a_H^S, a_H^B, A)$, each agent's membership decision is based on a comparison between $\max(u^S, u^B, u^S + u^B)$ and the membership fee A. This means the set of users at $\mathbf{P}_H$ is:[19]

$$Y_H \equiv \{\omega \mid \max(u^S, u^B, u^S + u^B) > A\}$$

And they allocate into different market segments in the following way:

$$\begin{aligned} PS_H &\equiv \{\omega \in Y_H \mid \max(u^S, u^B, u^S + u^B) = u^S\} = \{\omega \mid u^S > A, u^B \leq 0\} \\ PB_H &\equiv \{\omega \in Y_H \mid \max(u^S, u^B, u^S + u^B) = u^B\} = \{\omega \mid u^S \leq 0, u^B > A\} \\ SB_H &\equiv \{\omega \in Y_H \mid \max(u^S, u^B, u^S + u^B) = u^S + u^B\} \\ &= \{\omega \mid u^S > 0, u^B > 0, \text{ and } u^S + u^B > A\} \end{aligned}$$

where $u^S = (v^S - a_H^S)N_H^B$, and $u^B = (v^B - a_H^B)N_H^S$.

Because the same bundled membership fee A applies to both services, it will affect every agent's membership decision, and in turn affect the demands in all market segments under hybrid bundling. Therefore the bundled membership fee A is *not* redundant. Figure 3 illustrates agent allocation at $\mathbf{P}_H$.

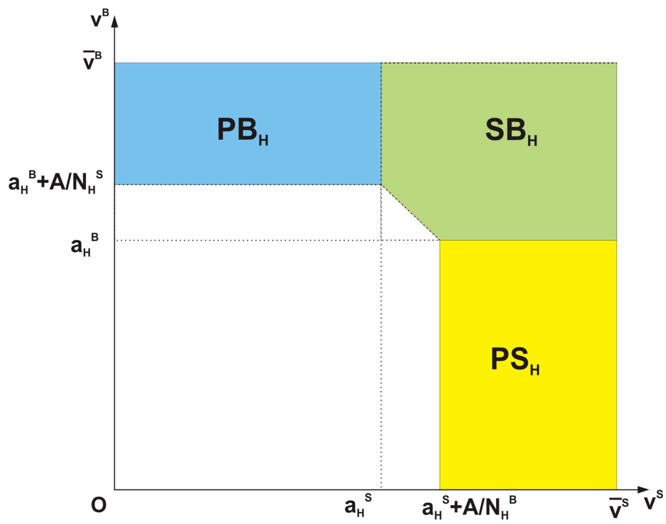


Figure 3: Agent Allocation under Hybrid Bundling (with Full Rationality)

The demands from two market sides at $\mathbf{P}_H = (a_H^S, a_H^B, A)$ are determined simultane-

bundling literature. We provide a detailed discussion of it in appendix.

[19] We use subscript H for all results derived under hybrid bundling with full rationality.

ously by the following two equations:[20]

$$N_H^S = \Pr[\omega \in PS_H \cup SB_H] \equiv n_H^S(a_U^S, a_U^B, A)$$
$$N_H^B = \Pr[\omega \in PB_H \cup SB_H] \equiv n_H^B(a_U^S, a_U^B, A)$$

The platform's profit under hybrid bundling is then:

$$\Pi_H(\mathbf{P}_H) \equiv (A - F^S)N_H^{PS} + (A - F^B)N_H^{PB} + (A - F)N_H^{SB} + (a_H^S + a_H^B - c)N_H^S N_H^B$$

4.1 Demand comparison

Lemma 2 ***(Replication with degenerate strategies)*** *At any transaction fees $(a^S, a^B) \in \mathbb{R}^{+2}$, the degenerate unbundled-sales strategy $\mathbf{P}_{U0} = (a^S, a^B, 0, 0)$ and the degenerate hybrid bundling strategy $\mathbf{P}_{H0} = (a^S, a^B, 0)$ produce exactly the same demand and profit.*

Proof. See appendix. ■

This lemma is quite intuitive. When all membership fees are set to zero, a unbundled-sales strategy and a hybrid bundling strategy with exactly the same usage fees will look exactly the same to agents. Thus each agent will make the same membership and trading decisions under these strategies, resulting in the same demand and profit in each market segment.

When membership fees are positive, however, an agent may make different decisions under different strategies.

Suppose we set the same a^S, a^B in both strategies and $A^S = A^B = A > 0$. It is easy to see the different shapes of each market segment from Figures 2 and 3. In particular, there exist non-users, pure sellers, and pure buyers in the unbundled-sales system who become seller-buyers under hybrid bundling. Their surplus from either buying or selling alone is not high enough to compensate the membership fee, but their combined surplus is. They value the savings from a lower combined membership fee for both services under hybrid bundling.

If we set $A^S = A^B = \frac{1}{2}A > 0$, still there exist pure sellers and pure buyers (among others) under unbundled sales who become seller-buyers under hybrid bundling. Take an agent α, for example, who expects the platform to have N^S sellers and N^B buyers, and has $v_\alpha^S \in (a^S, \frac{A^S}{N^B} + a^S)$ and $v_\alpha^B > \frac{A}{N^S} + a^B$. She will be a pure buyer under unbundled sales since $0 < u_\alpha^S < A^S$ and $u_\alpha^B > A > A^B$; but she will be a seller-buyer under hybrid bundling because $u_\alpha^S + u_\alpha^B > A$.[21]

[20] The specific demand formulae are shown in appendix, where N_H^S and N_H^B are determined at $\mathbf{P}_H$ by a simultaneous system of two equations (6). In case the solutions to that system are correspondences, we take the suprema of them. This is feasible because N_H^S and N_H^B are both bounded ($N_H^S, N_H^B \leq 1$).

[21] Notice our discussion in neither of the two cases here is exhaustive, and not all the differences in demands are represented in the figures. Whenever demand on one side changes (e.g. N^B), the *average per-transaction charge* on the other side (e.g. $\frac{A}{N^B} + a^S$) will be affected through the network effects. These

In general, unbundled sales and hybrid bundling will yield different demands, unless the bundled membership fee A is zero.

For the remainder of the paper, we will use a hybrid bundling strategy that is based on the degenerate strategies discussed in Lemma 2.

We start from a strategy with arbitrary transaction fees $(a^S, a^B) \in \mathbb{R}^{+2}$ and zero bundled membership fee, denoted $\mathbf{P}_{H0} \equiv (a^S, a^B, A = 0)$. By Lemma 2 we know it is equivalent to an unbundled-sales strategy with the same transaction fees (a^S, a^B) and no membership fees. Then we raise the bundled membership fee A by one unit from zero. This change breaks the equivalence between this strategy and any unbundled-sales strategy, and therefore can serve as a tool to study the differences between these two strategies. The results we show will be the marginal effects on demand and profits by applying such a strategy.

4.2 The two-part-tariff effect

The bundled membership fee applies to all users of the platform, therefore raising it by one unit results in a *simultaneous and commensurate* rise in the final prices of all users. Such a price rise will generally decrease demand, but its impact on different market segments is not the same. In particular, we have the following result.

Proposition 1 ***(The two-part-tariff effect)*** *For any transaction fees* $(a^S, a^B) \in \mathbb{R}^{+2}$, at $\mathbf{P}_{H0} = (a^S, a^B, 0)$ we have

$$\frac{\partial N_H^{SB}}{\partial A}(\mathbf{P}_{H0}) = 0.$$

That is, at ***any*** *transaction fees, imposing a small bundled membership fee on everyone has* ***no*** *first-order effect on demand of seller-buyers.*

Proof. See appendix. ■

We illustrate the intuition of Proposition 1 in Figure 4.

"finer" changes can be seen from (6) and (7) in appendix.

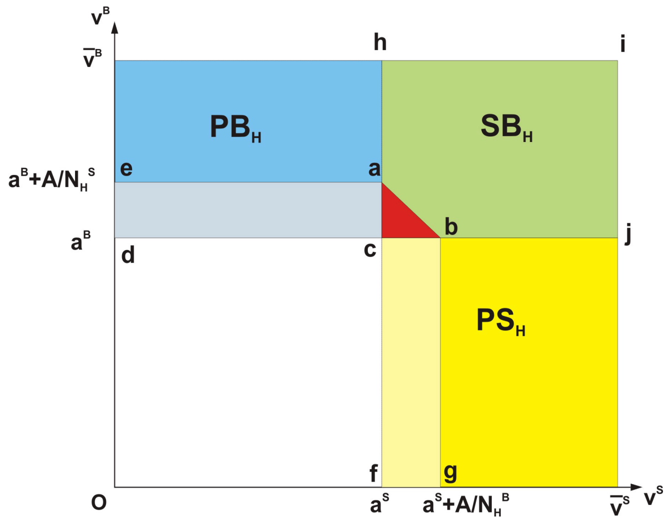


Figure 4: The Two-Part-Tariff Effect

At transaction fees (a^S, a^B), agent allocation in the market is fully characterized by two straight lines $v^S = a^S$ (line fh) and $v^B = a^B$ (line dj) in Figure 4, which split the type space $\mathbb{V}$ (assumed to be $[0, \bar{v}^S] \times [0, \bar{v}^B]$ here) into four segments:

Pure sellers PS_H $(cf\bar{v}^S j)$; pure buyers PB_H $(cd\bar{v}^B h)$; seller-buyers SB_H $(cjih)$; and non-users $(ofcd)$.

When the bundled membership fee is raised to $A > 0$, all users' final prices are raised together by A. This means all the service options that the platform provides (selling, buying or both) become equally more expensive, whereas the price differences among them stay unchanged. Therefore raising A does not lead to any agent switching from single-service user to two-service user or reversely, but only causes the lowest-typed users to leave the platform. This is represented in Figure 4 as:

PS_H shrinking by $cfgb$; PB_H shrinking by $acde$; and SB_H shrinking by abc.

We focus on seller-buyers - SB_H, and the change in their demand - area abc.

Without the bundled membership fee, seller-buyers are characterized by $SB_H = \{\omega \mid u^S > 0, u^B > 0\} = \{\omega \mid v^S > a^S, v^B > a^B\}$. Since the bundled membership fee applies to both services that seller-buyers use, raising it must cause changes "along both dimensions" of set SB_H. In particular, a positive A imposes an additional constraint, $u^S + u^B > A$, on set SB_H, which requires that a seller-buyer get total benefits from using *two* services high enough to compensate the extra fee A.

Put differently, *the burden to compensate the one-dimensional rise in the bundled membership fee from zero to A is shared by increases in seller-buyers' valuations in two dimensions.* Therefore the change in SB_H (area abc) has a "triangular" shape, whose probability measure is of order greater than A (e.g. it is proportional to A^2 in the case of uniform distribution). This implies that the first-order decrease in the number of seller-buyers, $\frac{\partial N_H^{SB}}{\partial A}$, vanishes when $A = 0$ (e.g. it is proportional to A in the case of uniform distribution, and therefore vanishes when $A = 0$).[22]

In mathematical terms, the two-part-tariff effect reflects the difference in dimensionality between the type space $\mathbb{V}$ (two-dimensional) and the domain of A (one-dimensional). If the type space were one-dimensional (as in standard two-sided markets), there would be no such effect.[23]

In summary, Proposition 1 says that, starting from any transaction fees, imposing a small additional membership fee on everyone will only dissuade a negligible number of seller-buyers from continuing to use the services.

This effect does not exist in either pure bundling or mixed bundling strategies.[24] Only a *two-part tariff* with a *bundled* membership fee can achieve it. Therefore we call it the **two-part-tariff effect**.

4.3 The degree of mixedness

The seller-buyers are of special importance to the platform in a mixed two-sided market. They are more active - each of them acts as both a seller and a buyer and therefore brings "double" revenues; they are also potentially more cost-effective - their fixed cost to the platform is no more than the combination of two single-service users, since we allow the possibility of scope economies (see Assumption C). As we will show shortly, the conditions for hybrid bundling to dominate unbundled sales would not hold without them.

Given price $\mathbf{P}$ of either strategy, we use *the proportion of seller-buyers among all users* as a measure of how mixed a two-sided market is, which we call **the degree of mixedness**, denoted $m(\mathbf{P}) \equiv \frac{N^{SB}}{N}$. This measure turns out to be an important factor in many of the results we present later.

Lemma 3 *For any transaction fees $(a^S, a^B) \in \mathbb{R}^{+2}$, at $\mathbf{P}_{H0} = (a^S, a^B, 0)$ we have*

$$\frac{\partial}{\partial A} m(\mathbf{P}_{H0}) > 0 \text{ \textbf{\textit{if and only if}} } 0 < m < 1.$$

That is, the degree of mixedness is strictly increasing in the bundled membership fee if and only if there exists a positive measure of seller-buyers and the market is not fully mixed.

Proof. See appendix. ∎

As long as the transaction fees a^S and a^B induce some but not full mixedness in the market, imposing a small bundled membership fee strictly increases the degree of

[22] The argument here relies on a *non-empty* set of seller-buyers, which is relaxed in the formal proof. In the case that transaction fees a^S and a^B induce an empty SB_H, i.e. $N_H^{SB} = 0$, the area *abc* must also have a zero probability measure regardless of A, which means the marginal effect of raising A on N_H^{SB} is still zero.

[23] Note that the two-part-tariff effect relies on A starting from zero, so that the multiproduct monopolist is *not* already exploiting price discrimination in two dimensions before it raises A. From $A > 0$, further raising the bundled membership fee may lead to first-order decrease in the demand of seller-buyers.

[24] We present a detailed discussion of the differences between these strategies in section 6.

mixedness, because on the first order this strategy discourages some single-service users but none seller-buyers from using the platform.

If there are no seller-buyers at certain transaction fees ($N_H^{SB} = 0$), the degree of mixedness m is either zero (when $N_H > 0$) or non-existent (when $N_H = 0$). Imposing a bundled membership fee in either case cannot increase m, as it does not change the number of seller-buyers. Thus, if at some transaction fees hybrid bundling does increase m, it must be that the market is already mixed.

Since seller-buyers are more cost-effective compared to single-service users, and raising A generally increases their proportion among users, it will therefore save costs for the platform. We discuss this further in the next section.

4.4 When to allow buyers to sell?

In this section we answer the question of when the platform should use hybrid bundling instead of unbundled sales.

We present a series of conditions that guarantee the dominance of hybrid bundling. Starting from the optimal unbundled-sales strategy $\mathbf{P}_0^* = (a^{S*}, a^{B*}, 0, 0)$ (defined in Assumption 5), we impose an additional bundled membership fee on top, which together with a^{S*} and a^{B*} constitutes a hybrid bundling strategy. If this manipulation increases profits, we are sure that hybrid bundling dominates unbundled sales.

In Proposition 2 we present the most general (and least transparent) condition. It implies strict dominance of hybrid bundling in a wide range of situations, including independence and negative correlation between agents' valuations for two services (v^S and v^B). We discuss these more specific cases in the corollaries to the proposition.

Proposition 2 *Suppose $\mathbf{P}_0^* = (a^{S*}, a^{B*}, 0, 0)$ is an optimal unbundled-sales strategy, then hybrid bundling strictly dominates unbundled sales if:*

$$(F^S + F^B - F)\left[\frac{g^S(a^{S*})G^{B|S}(a^{B*}|a^{S*})(1-G^{B|S}(a^{B*}|a^{S*}))}{1-G^B(a^{B*})} + \frac{g^B(a^{B*})G^{S|B}(a^{S*}|a^{B*})(1-G^{S|B}(a^{S*}|a^{B*}))}{1-G^S(a^{S*})}\right] + \int_{a^{S*}}^{+\infty}\int_{a^{B*}}^{+\infty}[g^S(s)g^{B|S}(b|\,a^{S*}) + g^B(b)g^{S|B}(s|\,a^{B*}) - g(s,b)]dbds > 0. \tag{2}$$

Proof. See appendix. ■

Why would imposing a small bundled membership fee on everyone be profitable? The gains come from two sources - cost saving and revenue increase - each represented by a term in condition (2).

Cost saving The first term on the left-hand side of condition (2) represents the cost saving from hybrid bundling, which is always *non-negative* regardless of the distribution of agents' valuations. The two parts in brackets of this term correspond respectively to the decreases in demands of pure sellers and pure buyers due to increasing A.

As Lemma 3 shows, raising A generally decreases the number of single-service users and increases the proportion of seller-buyers, who are more cost-effective whenever there are scope economies. Therefore raising A will generally save costs for the platform.

Note this effect becomes zero when there exist no scope economies ($F = F^S + F^B$), no seller-buyers ($N^{SB} = 0$), or no single-service users ($N^{PS} = N^{PB} = 0$). It is positive in all other situations.

Revenue change Raising the bundled membership fee by one unit from zero has three *first-order* effects on revenues, represented by the second term of condition (2), whose sign depends on the underlying distribution of valuations.

i) it reduces the number of users paying memberships (the losses in demands of single-service users N^{PS} and N^{PB} are of the same order as the change in A);

ii) it reduces the total volume of transactions between sellers and buyers, leading to lower total transaction fees paid (the total volume of transactions $N^S \cdot N^B$ decrease, as $N^S = N^{PS} + N^{SB}$ and $N^B = N^{PB} + N^{SB}$, where both N^{PS} and N^{PB} decrease and N^{SB} does not change); and

iii) it increases membership revenue per user by one unit.[25]

We analyze their net effect in each market segment:

Seller-buyers: We know from the two-part-tariff effect that raising A does not change the demand of seller-buyers, thus there is no loss in the volume of transactions caused by them, either. The platform therefore gets a pure gain of one unit from each and every seller-buyer, with the total gain exactly equal to the number of them.

Pure sellers: Their demand N^{PS} decreases, leading to losses in transaction revenue, but there are gains in their membership revenue due to the higher fee charged. The net effect may be positive or negative, depending on the distribution of valuations.

Pure buyers: Similar to pure sellers, their demand N^{PB} also decreases, leading to losses in transaction revenue, but there are also gains in their membership revenue. The sign of the net effect depends on the distribution of valuations.

Thus, the overall impact of hybrid bundling on revenues may be negative or positive, depending on the distribution of valuations. We present results for two special cases below.

4.4.1 The case of independence

Corollary 1 *Hybrid bundling strictly dominates unbundled sales if v^S and v^B are independent.*

[25] Raising A has other (indirect) effects on demand that are not listed here, which are due to the network effects. These effects are however negligible when $A = 0$, since they are all of order greater than the change in A. See formulae (8) and (9) in appendix for these effects when $A > 0$.

Proof. See appendix. ■

When v^S and v^B are independent, raising A by one unit from zero causes no change in revenues from either pure sellers or pure buyers. The optimality of a^{S*} and a^{B*} under unbundled sales guarantees that, the gains from their membership payments exactly compensate the losses in their demand *plus* the losses in their volume of transactions. This leaves us with seller-buyers, but we already know there is no loss in demand or transactions due to them.

Therefore the net effect of raising A by one unit on revenues will be as if the platform is extracting one unit of surplus from each and every one of the existing seller-buyers. This means the second term of condition (2) reduces exactly to the number of seller-buyers at $\mathbf{P}_0^*$, N_U^{SB*}.

Note that independence between v^S and v^B also implies $N_U^{SB*} = N_U^{S*} \cdot N_U^{B*}$, and Assumption 5 guarantees that the numbers of sellers and buyers must both be positive, otherwise the platform gets zero optimal profit under unbundled sales. Thus, there must also be a positive number of seller-buyers at the optimal unbundled-sales strategy, $N_U^{SB*} > 0$, and the gains from hybrid bundling is therefore strictly positive.

4.4.2 The case of negative correlation

Corollary 2 *Suppose* $\mathbf{P}_0^* = (a^{S*}, a^{B*}, 0, 0)$ *is an optimal unbundled-sales strategy, then hybrid bundling strictly dominates unbundled sales if* ***any*** *of the following conditions holds:*

(i) $G^{S|B}(a^{S*} | b)$ *is strictly increasing in* b *for all* $b > a^{B*}$ *in* $\mathbb{V}$*; or*

(ii) $G^{B|S}(a^{B*} | s)$ *is strictly increasing in* s *for all* $s > a^{S*}$ *in* $\mathbb{V}$*; or*

(iii) *The two functions above are both constants for all* $b > a^{B*}$ *and* $s > a^{S*}$ *in* $\mathbb{V}$*, respectively,* $G^{S|B}(a^{S*} | a^{B*}) < 1$ *and* $G^{B|S}(a^{B*} | a^{S*}) < 1$*.*

Proof. See appendix. ■

Recall that the second term on the left-hand side of condition (2) depends on the conditional distributions $G^{B|S}$ and $G^{S|B}$, which determines the "correlation" between v^S and v^B. The conditions in Corollary 2 can therefore be roughly interpreted as a form of *negative correlation* between v^S and v^B.

For instance, condition (i) means that within the set of all buyers (users with $v^B > a^{B*}$), the *higher* a buyer's valuation for buying is (high v^B), the more likely she will have a valuation for selling that is *lower* than a^{S*} (low v^S). In Figure 4, this means that, within the set of *pure buyers* (PB_H, where users have $v^B > a^{B*}$ but $v^S < a^{S*}$), there is a *higher* density of agents when v^B is *higher*, and a *lower* density when v^B is *lower.*

As we discussed previously, raising the bundled membership fee A makes the platform lose pure buyers with *low* types (the area *acde* in Figure 4 where v^B is low). Since when condition (i) holds, there are *fewer* pure buyers close to the *lower* end of PB_H, the losses in demand and transactions are therefore also *lower* compared to the case when (i) does not hold.

The gains in memberships from pure buyers due to raising A, however, is *higher* when condition (i) holds, as there are *more* of them close to the *higher* end of PB_H, who stay on the platform. These gains more than compensate for the losses, and thus the platform makes positive profits from pure buyers.

The revenue changes due to pure sellers will then be at least covered by the pure gains from seller-buyers, no matter how the distribution of valuations behave within the set of sellers. This results in a strictly positive net effect on total revenues.

The case when condition (ii) holds is completely symmetric. Thus, either (i) or (ii) will ensure that the overall impact of raising A on revenues, the second term of condition (2), is positive. This is why negative correlation favors hybrid bundling.

Condition (iii) corresponds to independence between v^S and v^B in a "weak" sense, and actually implies dominance of hybrid bundling under independence in the normal sense.

4.4.3 The case of positive correlation

When v^S and v^B are positively correlated, however, we cannot conclude for all kinds of distributions what the sign of the net impact of hybrid bundling on revenues would be.

In this case, unbundled sales generally can already do quite well in capturing the people with high valuations for both services (since they are positively correlated!). Therefore hybrid bundling is less likely to bring more revenues.

However, the effect of cost saving under hybrid bundling (the first term of condition (2)) is always present regardless of the correlation between valuations, and it therefore expands the range of situations of dominance by hybrid bundling.

4.4.4 The importance of mixedness

Corollary 3 *Condition (2) does* ***not*** *hold if the market is* ***not mixed*** *at the optimal unbundled-sales strategy.*

Proof. See appendix. ■

As mentioned in the discussion of Proposition 2, both sources of gains under hybrid bundling rely critically on seller-buyers. Without them, the market would not be mixed, and hence there would be no cost saving nor revenue gain by raising the bundled membership fee from zero.

5 Robustness

In this section we discuss two extensions of the model. In each of them, we modify a previous assumption to make the model more suitable for some real-life examples. They therefore serve as "robustness checks" of the results we have found.

5.1 Technological impact

The examples we have discussed span various industries, and the platforms in them may operate on different business models and employ distinct technologies. Like many other models in the literature of bundling, our model so far allows for much more generality on the consumption side (by imposing minimal restrictions on the distribution of consumer valuations) than on the production side (by using possibly oversimplified cost assumptions). In this section, we take a step further on the production side of the bundling story.

After all, most if not all the conditions in the bundling literature for one strategy to dominate another boil down to restrictions on costs and distribution of valuations. We believe that generalizing the cost assumptions used in the literature will add power to the conditions derived. We also think that the assumption we use in this section is more suitable than Assumption C for many of the examples we have discussed.

For this section only, we change the cost assumption to the following:

Assumption NC ***(New cost structure)*** *With unbundled sales, the platform incurs fixed costs* F^S *per seller and* F^B *per buyer. With hybrid bundling, it incurs a fixed cost* F *for* ***any*** *user. The variable cost it incurs is the same with either strategy, which is* c *per transaction. We still assume* $\min(F^S, F^B, F, c) \geq 0$*, and* $F \in [\max(F^S, F^B), F^S + F^B]$.

Under Assumption NC, the fixed cost incurred for a seller-buyer is $F^S + F^B$ with unbundled sales, while it is F with hybrid bundling. Thus there is again possibility for scope economies. However, Assumption NC also implies that the single-service users are more costly under hybrid bundling than unbundled sales. Thus changing the cost assumption from C to NC does not necessarily favor either strategy *ex ante*. However we show in Proposition 3 that hybrid bundling again finds a better way to exploit the scope economies.

Real-life examples Assumption NC suits the production technologies of platforms whose membership include some hardware device or registration process which represents a large part of the fixed cost per member. In mobile telecommunications, for instance, it is reasonable to assume that the fixed cost of a new subscriber is mainly due to the mobile phone given out along with the subscription. Suppose some network provider used unbundled sales, then it would need to provide each caller with a device that makes calls, and each receiver with one that receives; a user who joins both sides would then obtain both devices. With hybrid bundling, instead, the platform needs to provide every user with one device combining the two functions. Thus Assumption NC describes this example better than Assumption C.

For financial intermediaries, for instance, the fixed cost involved in setting up a new account is presumably mainly the time and effort it takes to gather, verify and evaluate relevant information of the account holder. If the platform used unbundled sales, it might need to gather and process different sets of information from borrowers (on their trustworthiness, say) and lenders (on their available funds, say); for a same user who joins both sides, it is unlikely that the platform can save much between the two procedures, as two different accounts (one for each side) would need to be set up for that user. If it used hybrid bundling, however, it would need both sets of information of every account holder, since it would need to provide both services to each of them. Thus Assumption NC again seems a better fit than Assumption C.

Impact on results The main impact of Assumption NC on our earlier results is that the transition from the optimal degenerate unbundled sales strategy to its equivalent hybrid bundling strategy will not be "smooth" in profit any more. This is because there is a complete change in the structure of fixed costs when we switch from unbundled sales to hybrid bundling, including:

i) additional fixed cost of each pure seller (increased by $F - F^S \geq 0$) and each pure buyer (increased by $F - F^B \geq 0$); and

ii) reduction in fixed cost of each seller-buyer (decreased by $F^S + F^B - F \geq 0$).

Thus we may need additional conditions to balance these changes in fixed costs. These conditions are summarized in the following proposition.

Proposition 3 *Under Assumption NC, hybrid bundling strictly dominates unbundled sales if* ***any*** *of the following conditions holds:*

(i) $F^S \neq F^B$, $\min(F^S, F^B) > 0$, *and*

$$m(\mathbf{P}_0^*) \geq \frac{F - \min(F^S, F^B)}{\min(F^S, F^B)} \tag{3}$$

where $m(\mathbf{P}_0^*) = \frac{N_U^{SB*}}{N_U^*}$ *is the degree of mixedness of the market at the optimal unbundled sales price* $\mathbf{P}_0^*$*;*

(ii) $F^S = F^B > 0$, *and condition (3) holds with strict inequality;*

(iii) $F^S = F^B > 0$, *condition (3) is binding, and the following condition holds:*

$$(F - F^S)\frac{g^S(a^{S*})G^{B|S}(a^{B*}|a^{S*})}{1-G^B(a^{B*})} + (F - F^B)\frac{g^B(a^{B*})G^{S|B}(a^{S*}|a^{B*})}{1-G^S(a^{S*})} + \int_{a^{S*}}^{+\infty}\int_{a^{B*}}^{+\infty}[g^S(s)g^{B|S}(b|\, a^{S*}) + g^B(b)g^{S|B}(s|\, a^{B*}) - g(s,b)]dbds > 0; \quad (4)$$

(iv) $F^S = F^B = F = 0$, *and condition (4) holds.*

Proof. See appendix. ■

The intuition of Proposition 3 is two-fold.

First, we find a new condition under Assumption NC: condition (3), which controls the cost changes during the transition from unbundled sales to hybrid bundling. Its left-hand side is the proportion of seller-buyers in all users at the optimal unbundled sales prices. Since seller-buyers are still potentially more cost-effective under Assumption NC, a higher proportion of them will save more fixed cost during the transition from unbundled sales to hybrid bundling. The right-hand side of condition (3) is the *percentage increase in fixed cost per single-service user on the low-cost side.* Condition (3) therefore ensures that there are sufficiently many seller-buyers who bring savings in fixed cost that at least cover the losses due to single-service users, so that the transition is at least smooth in profit.

Second, condition (4) is the new version of condition (2), which represents the profit changes under hybrid bundling due to raising A. The first two terms of its left-hand side come from the cost savings during the transition between strategies. The two-part-tariff effect says that raising the bundled membership fee will only reduce the demand from single-service users on the first order, who under Assumption NC are more costly to the platform with hybrid bundling. Therefore reducing their proportion always saves fixed cost for the platform. The third term here is exactly the same as the second term of condition (2) we presented earlier, which represents the net profit implications that are not directly related to production technology.

Conditions (3) and (4) are useful in different situations of cost parameters.

New insights Proposition 3, especially condition (3), provides new insights for the observation that some mixed two-sided platforms bundle while others do not. In situations (i), (ii) and (iii), condition (3) implies that, given the same percentage change in costs, hybrid bundling is *more likely* to dominate unbundled sales in markets with *higher* degrees of mixedness.

As we discussed in section 1, this result is consistent with platform pricing behaviors observed in financial intermediation markets. Unlike stock exchanges, social lending platforms usually employ unbundled sales of their services to lenders and borrowers. We interpret this as the result of a much lower proportion of users who both lend and borrow in social lending than the proportion of users who both sell and buy in stock exchanges.

Even if platforms in both markets used unbundled sales, stock exchanges would most likely still end up with a higher degree of mixedness, and thus they are more likely to find hybrid bundling more attractive.

5.2 Bounded agent rationality

The assumption in section 4 that agents take into consideration all future benefits from trading in their membership decisions seems to work in favor of hybrid bundling vis-à-vis unbundled sales. Unlike unbundled sales, hybrid bundling has only *one* membership fee, which does not directly "prompt" an agent to associate it with a particular side of the market. While a *fully rational* agent will consider the expected surplus from trading on *both* sides, an agent with *bounded rationality* may ignore part of these surpluses, which will generally result in lower demand for the platform's services under hybrid bundling.

In this section, we "test" the results we have found so far under an assumption of bounded agent rationality. We show that they remain robust even if each agent in her membership decision considers only the benefit from trading on *one* side of the market.

We continue to use hybrid bundling price $\mathbf{P}_H = (a_H^S, a_H^B, A)$.

Motivating example Consider an agent who wants to buy a CD on eBay. Before signing up, suppose she has a high expected payoff from this purchase. She may also have some idea about the possibility that she could one day sell her old CDs there and expects a positive payoff from selling (which is lower than that from the current purchase). Nonetheless, before joining she is not sure how the system will work and how the payoff from future selling can be realized. Thus at the moment she may have to make the decision based only on her current need. After becoming a user, however, through time she gets familiarized with the services and can finally realize her expected payoff from selling as well. In this example, the complexity of the system, or the agent's lack of information (of how to realize expected payoff), makes it impractical to take account of the "less urgent" need when making the decision. The agent here is not fully rational - she cannot consider all potential needs *ex ante*. It is this particular kind of bounded rationality that we will focus on in this section, and it is summarized in the following assumption:

Assumption BR ***(Bounded Rationality)*** *Given hybrid bundling price $\mathbf{P}_H$, each agent makes her membership decision based only on her more urgent need - the higher of u^S and u^B.*

We believe for online trading and social lending markets, this is a more realistic assumption than the one of full rationality. Note that Assumption BR does not apply when the platform uses unbundled sales, because the membership fees of that strategy are separated for buying and selling, and consumption separability (Assumption 3) guarantees that an agent does not need to compare u^S and u^B in her membership decision regarding either side.[26] The agent is "forced" to make the rational membership decision under unbundled sales. Thus, compared to full rationality, Assumption BR simply works against the appeal of hybrid bundling vis-à-vis unbundled sales.

Under Assumption BR, the group of all users at hybrid bundling price $\mathbf{P}_H$ is the following set:[27]

$$Y_{BR} \equiv \{\omega \mid \max(u^S, u^B) > A\}$$

From an *ex post* view, the set of seller-buyers at $\mathbf{P}_H$ takes the following form:

$$\begin{aligned} SB_{BR} &\equiv \{\omega \in Y_{BR} \mid \min(u^S, u^B) > 0\} \\ &= \{\omega \mid \max(u^S, u^B) > A, \text{ and } \min(u^S, u^B) > 0\} \end{aligned}$$

Each seller-buyer, in the first instance of using the platform, is only interested in either buying or selling. The first requirement $\max(u^S, u^B) > A$ makes sure she is willing to sign up in the first instance. Nonetheless, after becoming a member she is able to use the service of the opposite side without any further fixed fees. The second requirement $\min(u^S, u^B) > 0$ is to guarantee that she will indeed do so later on. Thus it is exactly the agents in set SB_{BR} that will end up using the services of both sides.

The pure sellers and pure buyers are then the following sets:

$$\begin{aligned} PS_{BR} &= \{\omega \mid u^S > A, u^B \leq 0\} \\ PB_{BR} &= \{\omega \mid u^S \leq 0, u^B > A\} \end{aligned}$$

Because the same bundled membership fee A still applies to every user, it will affect the demand in all market segments, just like under full agent rationality. Therefore the bundled membership fee is again *not* redundant. The allocation of agents at $\mathbf{P}_H$ is illustrated in Figure 5.[28]

[26] See the demand functions of system (1) in section 3.

[27] We use subscript BR for all variables under Assumption BR, except for the price which is still denoted $\mathbf{P}_H$.

[28] Compared to Figure 3, the main change due to bounded agent rationality in Figure 5 happens in the shape and size of segment SB_{BR}, which appears to be missing an extra triangular area of seller-buyers from SB_H in Figure 3. Those are the agents whose valuation for neither service is high enough to compensate A alone, but their sum is.

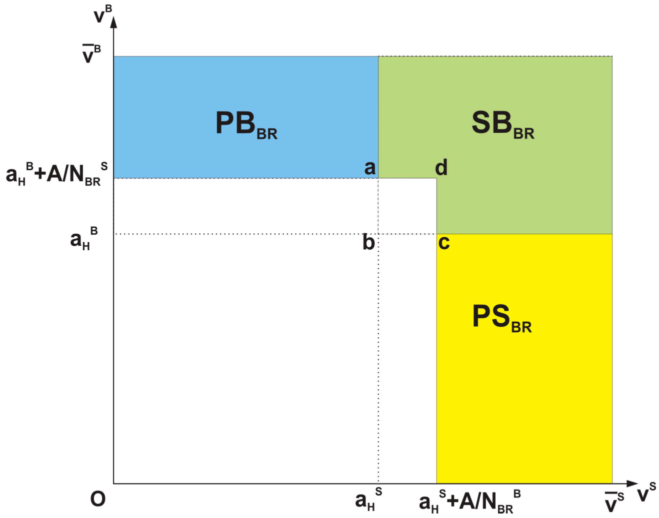


Figure 5: Agent Allocation under Hybrid Bundling (with Bounded Rationality)

The demands from two sides at $\mathbf{P}_H = (a_H^S, a_H^B, A)$ are determined simultaneously by the following two equations:[29]

$$
\begin{aligned}
N_{BR}^S &= 1 - G^S(a_H^S) - \int_{a_H^S}^{\frac{A}{N_{BR}^B}+a_H^S} \int_{-\infty}^{\frac{A}{N_{BR}^S}+a_H^B} g(s,b)dbds \\
N_{BR}^B &= 1 - G^B(a_H^B) - \int_{-\infty}^{\frac{A}{N_{BR}^B}+a_H^S} \int_{a^B}^{\frac{A}{N_{BR}^S}+a_H^B} g(s,b)dbds
\end{aligned} \tag{5}
$$

The platform's profit is then:

$$\Pi_{BR}(\mathbf{P}_H) \equiv (A - F^S)N_{BR}^{PS} + (A - F^B)N_{BR}^{PB} + (A - F)N_{BR}^{SB} + (a_H^S + a_H^B - c)N_{BR}^S N_{BR}^B$$

We have the following general conclusion.

Proposition 4 *Propositions 1 through 3, Corollaries 1 through 3, and Lemmas 1 through 3 all hold with bounded agent rationality.*

Proof. See appendix. ■

Although in general the demand under bounded rationality is lower than that under full rationality at the same price, when that price includes a zero membership fee, the differences in demands vanish. When an agent does not need to pay any membership fee, there is no real membership decision to make - she only needs to decide what transactions to make, which is not affected by our assumptions of her rationality. This means Lemma 2 holds.

The two-part-tariff effect also holds. The only difference from the full rationality case is that a rise in the bundled membership fee is now compensated by the higher of u^S and

[29] In case the solutions to the system are correspondences, we still take their suprema.

u^B instead of their sum, which leads to a "larger" decrease in demand of seller-buyers. The lost seller-buyers here is represented by a rectangular area *abcd* in Figure 5 instead of the triangular area *abc* in Figure 4. However, the total probability measure of them is still of the second order, again because the fee rise applies to both services and its "burden" is still shared in two dimensions.[30]

Therefore all the implications of the two-part-tariff effect still apply under bounded rationality.

Proposition 4 implies that the strength of hybrid bundling doesn't depend on agents' rationality of being able to foresee and take into account the service they value *less* when they decide whether to join the platform. It allows our earlier results to be applied to wider and more realistic contexts.

6 Discussion

In this section we provide detailed discussions of our model, concepts and results in relation to those in the existing literature.

6.1 Hybrid bundling and pure bundling

The hybrid bundling strategy we study is different from pure bundling strategies in the literature, because it involves two-part tariffs while the latter involves only "one-part" prices (that is, one purchase involves only one fee payment).

In pure bundling, there is only one price - the price for the pure bundle - which means all users would pay exactly the same fee for the service(s) they use on the platform.

Although the hybrid bundling strategy also involves pure bundling of memberships, the membership fee is just one part of the final price users pay. After becoming members, they still have the free choice of whether or not to pay the separate transaction fees, and they would not all pay the same fee(s) if they have different preferences. Therefore hybrid bundling generally induce different demand and profit than pure bundling.

6.2 Hybrid bundling and mixed bundling

In mixed bundling, users can also choose which service(s) to use and therefore there are ways to induce the same choices as a given hybrid bundling strategy. The main difference between them, however, lies in the different mechanisms induced by manipulations of prices in these strategies.

[30] As $SB_{BR} = \{\omega \mid \max(u^S, u^B) > A, \text{ and } \min(u^S, u^B) > 0\}$, where both functions still depend on benefits from both dimensions, the rise in A will still be shared by increases in seller-buyers' benefits from two dimensions.

6.2.1 Comparison between mechanisms

In a "one-sided" market setting, Long (1984) establishes the equivalence between mixed bundling and *simple* two-part tariffs, under a *unit*-consumption assumption. He shows that a two-part tariff consisting of prices of two separate products, q_1 and q_2, say, and one additional fixed fee for all consumers, q, is equivalent to a mixed bundling strategy with prices for separate products $p_1 = q_1 + q$, $p_2 = q_2 + q$ and price for the bundle $p_B = q_1 + q_2 + q$, where "*the bundle discount is like a fixed fee*".

Our discussion of hybrid bundling generalizes Long's discussion, because the strategy we study involves *general* two-part tariffs that allow for *multi-unit* consumption in the trading stage.

However, we further argue that the bundle discount in mixed bundling is *not* equivalent to the fixed/membership fee in either simple two-part tariffs or the hybrid bundling strategy, because it cannot achieve the two-part-tariff effect.

Unilateral price manipulation in mixed bundling As we mentioned in section 4.2, raising the fixed/membership fee in hybrid bundling leads to a simultaneous and commensurate rise in all users' final prices, which results in zero first-order decrease in seller-buyers' demand. However, manipulations of the bundle discount or any other price in mixed bundling always lead to demand changes of the same order as the price change itself, which are therefore never negligible.

This distinction is due to the different prices used. In mixed bundling, all three prices are final prices that consumers face. Unilateral adjustments of any final price will generally change the price difference between "neighboring" market segments involving the product (or bundle) under adjustment, which will in turn result in demand changes of the same order as the price change.

For example, consider the mixed bundling strategy (p_1, p_2, p_B) in McAfee, McMillan and Whinston (1989)[31] defined the same way as above, where the bundle discount is $\varepsilon \equiv p_1 + p_2 - p_B$. Raising the bundle discount whilst keeping p_1 and p_2 unchanged is equivalent to lowering p_B, which lowers the price difference between either separate product and the bundle, and therefore makes the bundle look more attractive. This change will generally cause consumers of both separate products "residing" along the *entire* "boundary" of the bundle-segment (those who "hardly" preferred these separate products to the bundle to begin with) to choose the bundle instead. The boundaries of the bundle-segment therefore shifts outwards in both dimensions, resulting in a first-order increase in the demand for the bundle.

Raising the bundle discount ε can also be achieved by raising p_1 whilst keeping p_2 and p_B constant, which will lower the price difference between "separate product 1" and the bundle, and makes the bundle look more attractive. This again will cause those consumers

[31] See Figures II and III in McAfee, McMillan and Whinston (1989).

of "separate product 1" who "hardly" preferred it to the bundle to begin with, to switch to the bundle. In other words, the boundary between these two segments will shift towards the segment of "separate product 1" by a distance proportional to the change in p_1. This again means a first-order increase in the demand of the bundle.

Multilateral price manipulation in mixed bundling Although there are ways to replicate the two-part-tariff effect through simultaneous manipulations of all the prices in mixed bundling, the required manipulations are further complicated by the multi-unit transactions that hybrid bundling permits users to make, which may make such replication impractical.

For instance, given a hybrid bundling strategy (a^S, a^B, A), a mixed bundling strategy that can replicate the demand and profit under it is $\mathbf{Q} \equiv (p^S, p^B, \varepsilon) = (a^S + \frac{A}{N^B}, a^B + \frac{A}{N^S}, A)$, where p^S is the fee any seller pays per sale, so that her total payment is $p^S N^B (= a^S N^B + A)$; p^B is the fee any buyer pays per purchase, so that her total payment is $p^B N^S (= a^B N^S + A)$; and ε is the lump-sum bundle discount offered to a seller-buyer, so that her total payment is $p^S N^B + p^B N^S - \varepsilon (= a^S N^B + a^B N^S + A)$.

The first problem with this replication is that offering a *lump-sum* bundle discount *only* to seller-buyers does not seem easy in practice. Nonetheless, there is no straightforward way to offer the same total discount on a *per-transaction* basis either, since a seller-buyer may well make *unequal* numbers of sales and purchases.

More problems arise when we use $\mathbf{Q}$ to replicate the two-part-tariff effect induced by raising A from zero to Δ, say. Just raising the bundle discount ε from zero to Δ does not work, since this will cause positive measures of single-service users to become seller-buyers. In addition, we need to raise p^S by $\frac{\Delta}{N^B}$ and p^B by $\frac{\Delta}{N^S}$ to make sure all users' final prices are raised by Δ (so that no single-service user would want to become seller-buyer). These intricate and impractical price manipulations are necessary exactly because of the *different* numbers of *multiple* transactions made by a buyer and a seller under *general* two-part tariffs.

This example shows the difficulty in replicating the hybrid bundling strategy and its two-part-tariff effect using mixed bundling.[32]

Analogous results of price discrimination Although hybrid bundling and mixed bundling have different mechanisms, they both can achieve multiproduct price discrimination, since they both have a price instrument that links the two dimensions of the type space (that is, the bundled membership fee and the bundle price, respectively). The striking contrast between our results and those in the mixed bundling literature is that

[32] Other difficulties arise in the case of $n > 2$ products, where a complete mixed bundle consists of $2^n - 1$ "sub-bundles" and therefore requires $2^n - 1$ prices to fully characterize. All these prices need to be changed equally in order to fully replicate the two-part-tariff effect in that case.

similar price discrimination can be achieved through these two mechanisms with price manipulation in *opposite* directions. McAfee, McMillan and Whinston (1989) achieve this by *reducing the bundle price* from the sum of optimal separate prices, while we achieve it by *raising the bundled membership fee* on top of the optimal separate transaction fees. The reason why both deviations can be profitable is of course that they are not exactly opposite - again, raising the bundled membership fee effectively changes everyone's final price, while reducing the bundle price does not reduce the final prices single-product consumers pay.[33]

6.2.2 Comparison between conditions

The general conditions for hybrid bundling to dominate unbundled sales, condition (2) in Proposition 2 and condition (4) in Proposition 3, are restrictions on the behavior of the valuation distribution and cost parameters. In this sense, they bear certain resemblance to the conditions for mixed bundling strategies to dominate unbundled sales in "one-sided" markets (for instance, equation (20) of Schmalensee (1984) and Proposition 1 of McAfee, McMillan and Whinston (1989)). All these conditions capture the marginal gains in profits from introducing two-dimensional price discrimination into an optimal strategy of one dimension.

However, the main difference between our results and theirs is that, conditions (2) and (4) incorporate two new elements: the cost savings from scope economies, and profit implications of transactions between sellers and buyers (who are all consumers of the monopolist). The former is due to our generalized cost assumptions, and the latter is due to our two-sided market context. These new elements generalize the existing results in the literature.

Some implications of our results in special situations, such as weakly negative correlation favoring hybrid bundling (Corollaries 1 and 2), confirm that similar intuition of correlation carries over from one-sided to two-sided markets.[34]

The new conditions we present (such as the conditions in Proposition 3) provide new insights specific to two-sided markets (such as the importance of mixedness to hybrid bundling).

[33] We show in Gao (2009b) that both deviations discussed here lead to *local* profit improvements which should not be used to infer properties of the *global* optimizers (e.g. whether the optimal bundling strategy would involve higher or lower "average price per product" compared to unbundled sales). There we provide an example where the optimal MMW-type mixed bundling strategy results in a *higher* average price per product, contrary to the impression one might derive from MMW's mechanism of offering a bundle discount that an optimal mixed bundling strategy would involve a *lower* average price.

[34] For example, in one-sided market settings, Adams and Yellen (1976) illustrates results of negative correlation with examples of discrete distributions; Long (1984) discusses a condition of negative correlation involving different conditional probabilities; Schmalensee (1984) illustrates both independence and negative correlation with bivariate normal distributions; McAfee, McMillan and Whinston (1989) provide a result of dominance by bundling in the case of independence.

6.3 Hybrid bundling and multiproduct nonlinear pricing

There is a connection between the two-part-tariff effect we have uncovered and the "optimality of exclusion" result by Armstrong (1996)[35], where he shows in a multiproduct pricing framework that, when agents have multidimensional types, it will generally be optimal to exclude a positive measure of agents with the lowest types from the market.

The two-part-tariff effect is actually the underlying mechanism that drives the "optimality of exclusion". If we apply the two-part-tariff effect to the whole type space, that is, if we start from transaction fees that induce all agents to be seller-buyers, then it implies that the platform will find it optimal to raise the bundled membership fee from zero until it excludes a positive measure of seller-buyers from the market, which is the same conclusion as Armstrong's.[36]

However, the two-part-tariff effect does not only apply to the whole type space - it applies to *all* possible sets of two-service users induced by *any* level of transaction fees, which are subsets of the whole type space.

Another (technical) difference between these results is that the two-part-tariff effect only requires *weak* convexity of the support of distribution; while "optimality of exclusion" requires *strict* convexity of the support in order to guarantee negligible first-order demand changes when the price manipulation Armstrong used starts from multidimensional-type strategies (such as mixed bundling). We are not concerned with such situations as the price manipulation we use start from one-dimensional unbundled sales strategy.

7 Conclusion

In real life, many two-sided markets are mixed - sellers may also buy and buyers can also sell. In this paper we have provided a model for the dynamics in such markets. As one same consumer may want to use the services that a monopoly platform provides to both sides, the platform becomes a multiproduct monopolist. Such platforms therefore can and will exploit their market power in multiple dimensions. Many real-life platforms achieve this through a hybrid bundling strategy involving two-part tariffs. We have found the key mechanism of such a strategy, the two-part-tariff effect, and presented general conditions for hybrid bundling to dominate the conventional unbundled strategy. This provides a new "multiproduct" explanation for the prevalence of two-part tariffs observed in platform pricing strategies.

This paper can be viewed from two different perspectives.

[35] See Proposition 1 of Armstrong (1996). We thank Dezsoe Szalay for pointing out this connection.

[36] Armstrong (1996) studies an n-product monopolist. We show in Gao (2009b) that a "generalized" two-part-tariff effect exists in the n-product case, where the decrease in demand for an "m-bundle" (that is a bundle consisting of m products, where $1 \leq m \leq n$) due to a rise in the bundled membership fee from zero is of order m. This confirms that the two-part-tariff effect also drives the "optimality of exclusion" in the n-product case.

First, it extends the theoretical literature on two-sided markets to the mixed case, where sellers may also buy and buyers can also sell; and

Second, it extends the bundling literature (and hence the multiproduct pricing literature) to the case where there exist network effects between the consumption of different products.

There is much more work to be done in either direction and this paper is simply the first step. In the first direction, we are starting to study the implications of mixedness for established results of standard two-sided markets, such as platform competition. Along the second line, we have found some properties of the optimal nonlinear pricing strategy, where the general intuition is that the network effects serve as a channel to transfer demand changes on one market side to the other and therefore amplifies the overall impact on profit. Difficulties in modelling the network effects arise when the number of services provided exceeds two and the platform becomes *multi-sided*. While there seem to be great real-life examples of such platforms (such as Google and iPhone), the general pattern of how network effects work across multiple sides is far from clear.

References

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

[2] Armstrong, M. (1996): “Multiproduct Nonlinear Pricing”, *Econometrica*, 64(1), pp. 51-75.

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

[4] Armstrong, M. and Wright J. “Two-Sided Markets, Competitive Bottlenecks and Exclusive Contracts.” Mimeo, 2004.

[5] Banal-Estanol, A. and Ottaviani, M. (2007): "Notes on Pure Bundling", working paper.

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

[7] Evans, D. (2003): "Some Empirical Aspects of Multi-sided Platform Industries." *Review of Network Economics*, Vol. 2.

[8] Evans, D. (2005): “The Industrial Organization of Markets with Two-Sided Platforms.” NBER Working Paper 11603.

[9] Fang, H. and Norman, P. (2006): "To Bundle or Not to Bundle", *Rand Journal of Economics*, 37(4), pp. 946-963.

[10] Gao, M. (2008): "Multiproduct Pricing with Network Effects", working paper.

[11] Gao, M. (2009b): "Multiproduct Price Discrimination with Two-Part Tariffs", working paper.

[12] Hagiu, A. (2006): "Pricing and Commitment by Two-Sided Platforms", *Rand Journal of Economics*, 37(3), pp. 720-737.

[13] Hausman, J. and Wright, J. (2006): "Two Sided Markets with Substitution: Mobile Termination Revisited." Working Paper.

[14] Hermalin B.E. and Katz, M. (2004): "Sender or receiver: who should pay to exchange an electronic message?" *Rand Journal of Economics*, Vol. 35, pp. 423-448.

[15] Jeon, D.S., Laffont, J.J., and Tirole, J. (2004): "On the Receiver Pays Principle." *Rand Journal of Economics*, Vol. 35, pp. 85-110.

[16] Littlechild, S.C. (1975): "Two-Part Tariffs and Consumption Externalities." *Bell Journal of Economics*, Vol. 6, pp. 661-670.

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

[18] 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.

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

[20] Rochet, J.C. and Tirole, J. (2004): "Defining Two-Sided Markets." Working Paper.

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

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

8 Appendix - Notes and Proofs

Assumption 5 - Discussion

This is mainly a restriction on the valuation distribution and the cost parameters so that costs are not too high or too low compared to agents' valuation, which allows us to focus on the more "interesting" situations. For example, it rules out the outcomes where costs are so high that there does not exist a profitable unbundled-sales strategy.

Compared to the literature, both our cost structure (which involves per-user fixed costs and transaction costs) and our type space (which is only weakly convex and can be the whole $\mathbb{R}^2$) are more general. This means the full characterization of the necessary and sufficient conditions for the existence of an optimal unbundled-sales strategy is more complicated.

For instance, one necessary condition would be:

There exists a real number $k \in [0, c]$ such that there exists a positive measure of agents who have $v^S > k$, and a positive measure of agents who have $v^B > c - k$.

This condition implies there is surplus at all for the platform to extract from transactions.

Other necessary conditions include that the per-user fixed costs (F, F^S, and F^B) are not too high so that the platform's profit from transactions plus the profit (or minus the loss) from user memberships does not go negative.

One simple sufficient condition that guarantees the latter is that $F = F^S = F^B = 0$, but it is not necessary and we are actually interested in the cases with positive fixed costs.

The part of Assumption 5 that $\mathbf{P}_0^*$ is on the interior of $\mathbb{V}$ only rules out situations where $\mathbb{V}$ is closed below in both dimensions (e.g. $\mathbb{V} = [a, +\infty) \times [b, +\infty)$) and the costs are so low that at least one optimal unbundled-sales fee is the lower bound of $\mathbb{V}$, in which case the first order conditions may not hold.

Since our main interest in this paper is not these existence-related conditions, we have chosen to use Assumption 5 as a single simplifying assumption instead.

Lemma 2

Demand under unbundled-sales price $\mathbf{P}_U = (a^S, a^B, A^S, A^B)$:

$$N_U^S = 1 - G^S\left(\frac{A^S}{N_U^B} + a^S\right)$$
$$N_U^B = 1 - G^B\left(\frac{A^B}{N_U^S} + a^B\right)$$
$$N_U^{SB} = \int_{\frac{A^S}{N_U^B}+a^S}^{+\infty} \int_{\frac{A^B}{N_U^S}+a^B}^{+\infty} g(s,b)\,db\,ds$$

Demand under hybrid bundling (with full agent rationality) at price $\mathbf{P}_H =$

(a^S, a^B, A), N_H^S and N_H^B, are determined by the following two simultaneous equations:

$$\begin{aligned} N_H^S &= N_H^{PS} + N_H^{SB} \\ N_H^B &= N_H^{PB} + N_H^{SB} \end{aligned} \tag{6}$$

where

$$N_H^{PS} = \int_{\frac{A}{N_H^B}+a^S}^{+\infty} \int_{-\infty}^{a^B} g(s,b)dbds \tag{7a}$$

$$N_H^{PB} = \int_{-\infty}^{a^S} \int_{\frac{A}{N_H^S}+a^B}^{+\infty} g(s,b)dbds \tag{7b}$$

$$N_H^{SB} = \int_{a^S}^{+\infty} \int_{a^B}^{+\infty} g(s,b)dbds - \int_{a^S}^{\frac{A}{N_H^B}+a^S} \int_{a^B}^{\frac{1}{N_H^S}[A-N_H^B(s-a^S)]+a^B} g(s,b)dbds \tag{7c}$$

At $\mathbf{P}_{U0} = (a^S, a^B, 0, 0)$ and $\mathbf{P}_{H0} = (a^S, a^B, 0)$ we have:

$$\begin{aligned} N_U^{SB} &= N_H^{SB} = \int_{a^S}^{+\infty} \int_{a^B}^{+\infty} g(s,b)dbds \\ N_U^S &= N_H^S = 1 - G^S(a^S) \\ N_U^B &= N_H^B = 1 - G^B(a^B). \blacksquare \end{aligned}$$

Proposition 1

Step 1: First consider the case where the point (a^S, a^B) lies in the interior of support $\mathbb{V}$ (which includes the case when $\mathbb{V}$ is a open set).

Step 1.1: Suppose $\mathbf{P}_{H0} = (a^S, a^B, 0)$ induces positive equilibrium demand N_H^S and N_H^B, which must solve system (1) because $(a^S, a^B, 0)$ is also a degenerate unbundled-sales strategy.

Then the demand at $\mathbf{P}_{H0}$ can be represented by (7) above.

The first-order derivative of N_H^{SB} with respect to A at general hybrid bundling price $\mathbf{P}_H$ is:

$$\frac{\partial N_H^{SB}}{\partial A}(\mathbf{P}_H) = -\frac{N_H^B - A \cdot \frac{\partial N_H^B}{\partial A}(\mathbf{P}_H)}{(N_H^B)^2} \cdot W(a^S, a^B, A, \frac{A}{N_H^B} + a^S) - \int_{a^S}^{\frac{A}{N_H^B}+a^S} \frac{\partial}{\partial A} W(a^S, a^B, A, s)ds \tag{8}$$

where

$W(a^S, a^B, A, t) \equiv \int_{a^B}^{\frac{1}{N_H^S}[A-N_H^B(t-a^S)]+a^B} g(t,b)db$, and

$$\begin{aligned} \frac{\partial}{\partial A} W(a^S, a^B, A, t) = &\frac{1}{(N_H^S)^2} \cdot \{[1 - \frac{\partial N_H^B}{\partial A}(\mathbf{P}_H)(t-a^S)]N_H^S - \frac{\partial N_H^S}{\partial A}(\mathbf{P}_H)[A - N_H^B(t-a^S)]\} \\ &\cdot g(t, \frac{1}{N_H^S}[A - N_H^B(t-a^S)] + a^B) \end{aligned}$$

And the first-order derivative of N_H^{PS} and N_H^{PB} with respect to A are:

$$\frac{\partial N_H^{PS}}{\partial A}(\mathbf{P}_H) = - \quad \frac{N_H^B - A \cdot \frac{\partial N_H^B}{\partial A}(\mathbf{P}_H)}{(N_H^B)^2} \cdot \int_{-\infty}^{a^B} g(\frac{A}{N_H^B}+a^S, b)db \tag{9a}$$

$$\frac{\partial N_H^{PB}}{\partial A}(\mathbf{P}_H) = - \quad \frac{N_H^S - A \cdot \frac{\partial N_H^S}{\partial A}(\mathbf{P}_H)}{(N_H^S)^2} \cdot \int_{-\infty}^{a^S} g(s, \frac{A}{N_H^S}+a^B)ds \tag{9b}$$

where

$$\frac{\partial N_H^S}{\partial A}(\mathbf{P}_H) \;=\; \frac{\partial N_H^{PS}}{\partial A}(\mathbf{P}_H)+\frac{\partial N_H^{SB}}{\partial A}(\mathbf{P}_H) \tag{10a}$$

$$\frac{\partial N_H^B}{\partial A}(\mathbf{P}_H) \;=\; \frac{\partial N_H^{PB}}{\partial A}(\mathbf{P}_H)+\frac{\partial N_H^{SB}}{\partial A}(\mathbf{P}_H) \tag{10b}$$

Let $A = 0$ and we immediately see that the domains of all the integrals in the expression of $\frac{\partial N_H^{SB}}{\partial A}(\mathbf{P}_H)$ become single points. Therefore at $\mathbf{P}_{H0} = (a^S, a^B, 0)$, for any $(a^S, a^B) \in \mathbb{R}^{+2}$, it must be

$$\frac{\partial N_H^{SB}}{\partial A}(\mathbf{P}_{H0}) = 0.$$

Step 1.2: Suppose there exists no positive equilibrium demand at $\mathbf{P}_{H0}$.

Then the only equilibrium demand at $\mathbf{P}_{H0}$ is $N_H^S = N_H^B = 0$ (recall that either one of them being zero implies both will be zero, by definition of u^S and u^B). We therefore have $N_H^{SB} = 0$ and raising A from 0 does not change N_H^{SB}. Actually in this case, $N_H^{SB} = 0$ for all $A \geq 0$.

Step 2: Now consider the case where the point (a^S, a^B) lies on the boundary of support $\mathbb{V}$ (which requires that $\mathbb{V}$ is at least partially closed).

This change may only reduce some of the domains of integration in the expression of $\frac{\partial N_H^{SB}}{\partial A}(\mathbf{P}_H)$, as density g will be zero if the domain of integration extends outside of $\mathbb{V}$ (in this case the decrease in SB will be only part of the area *abc* in Figure 4), but it will cause no other difference whatsoever compared to the situation considered in Step 1. However our proof in Step 1 is immune to reduction in domains as all we need is that they become single points when $A = 0$. Thus the result still holds in this case. The decrease in N_H^{SB} due to raising A from zero is still of order strictly greater than the change in A.

A caveat here is that Proposition 1 depends on Assumption 1, and it may not hold if the support $\mathbb{V}$ does not have full dimension on $\mathbb{R}^2$. For example, if $\mathbb{V}$ is a line segment, and if raising A causes changes in demand along it, then it may be that $\frac{\partial N_H^{SB}}{\partial A}(\mathbf{P}_{H0}) > 0$.■

Lemma 3

Step 1: First suppose $N_H > 0$ at $\mathbf{P}_{H0} = (a^S, a^B, 0)$.

By Proposition 1 we know $\frac{\partial N_H^{SB}}{\partial A} = 0$. Since $m = \frac{N_H^{SB}}{N}$, we have

$$\frac{\partial m}{\partial A} = \frac{\frac{\partial N_H^{SB}}{\partial A} \cdot N_H - N_H^{SB} \cdot \frac{\partial N_H}{\partial A}}{N_H^2} = N_H^{SB} \cdot \frac{-\frac{\partial N_H}{\partial A}}{N_H^2} \tag{11}$$

And since $N_H = N_H^{PS} + N_H^{PB} + N_H^{SB}$, we have $\frac{\partial N_H}{\partial A} = \frac{\partial N_H^{PS}}{\partial A} + \frac{\partial N_H^{PB}}{\partial A}$

Let $A = 0$ in (9), we have

$$\begin{aligned}
\frac{\partial N_H^{PS}}{\partial A}(\mathbf{P}_{H0}) &= -\frac{g^S(a^S) \cdot G^{B|S}(a^B \big| a^S)}{N_H^B} \\
\frac{\partial N_H^{PB}}{\partial A}(\mathbf{P}_{H0}) &= -\frac{g^B(a^B) \cdot G^{S|B}(a^S \big| a^B)}{N_H^S}
\end{aligned}$$

Step 1.1: Now we prove the "If" part:

Since $N_H = N_H^{PS} + N_H^{PB} + N_H^{SB} > 0$, and $m < 1$, it must be that at least one of N_H^{PS} and N_H^{PB} is strictly positive. Thus at least one of $\frac{\partial N_H^{PS}}{\partial A}(\mathbf{P}_{H0})$ and $\frac{\partial N_H^{PB}}{\partial A}(\mathbf{P}_{H0})$ must be strictly negative, while the other is non-positive. Thus we have $\frac{\partial N_H}{\partial A}(\mathbf{P}_{H0}) < 0$. And since $N_H^{SB} > 0$, we get $\frac{\partial m}{\partial A}(\mathbf{P}_{H0}) > 0$.

Step 1.2: "Only if":

Since $\frac{\partial m}{\partial A}(\mathbf{P}_{H0}) > 0$, we immediately know $m(\mathbf{P}_{H0}) < 1$, and by (11) we know $N_H^{SB} \neq 0$. Thus $N_H^{SB} > 0$.

Step 2: Suppose $N_H = 0$ at $\mathbf{P}_{H0} = (a^S, a^B, 0)$.

Since $N_H = 1 - G(a^S, a^B)$, we know $G(a^S, a^B) = \int_{-\infty}^{a^S}\int_{-\infty}^{a^B} g(s,b)dbds = 1$. Convexity of support $\mathbb{V}$ then implies that $g(s,b) = 0$ for all $(s,b) \in ((a^{S*}, +\infty) \times \mathbb{R}) \cup (\mathbb{R} \times (a^{B*}, +\infty))$. Then by (7) we must have $N_H(a^S, a^B, A) = N_H^{SB}(a^S, a^B, A) = 0$, for all $A \geq 0$, and therefore $m(\mathbf{P}_{H0})$ does not exist. In this case, neither of these conditions (i.e. $\frac{\partial m}{\partial A}(\mathbf{P}_{H0}) > 0$ and $[N_H^{SB} > 0, m > 1]$) holds. Thus we are done.■

Proposition 2

Step 1: Find first order conditions for optimality of $\mathbf{P}_0^* = (a^{S*}, a^{B*}, 0, 0)$ under unbundled sales

Under unbundled sales, we only need to consider degenerate strategy $\mathbf{P}_{U0} = (a_U^S, a_U^B, 0, 0) \in \mathbb{R}^{+4}$, where

$$\begin{aligned}
&N_U^S = 1 - G^S(a_U^S); \qquad N_U^B = 1 - G^B(a_U^B); \\
&N_U^{PS} = \int_{a_U^S}^{+\infty}\int_{-\infty}^{a_U^B} g(s,b)dbds; \quad N_U^{PB} = \int_{-\infty}^{a_U^S}\int_{a_U^B}^{+\infty} g(s,b)dbds; \\
&N_U^{SB} = \int_{a_U^S}^{+\infty}\int_{a_U^B}^{+\infty} g(s,b)dbds; \\
&\Pi_U(\mathbf{P}_{U0}) = (a_U^S + a_U^B - c)N_U^S N_U^B - F^S N_U^{PS} - F^B N_U^{PB} - F N_U^{SB}
\end{aligned}$$

At $\mathbf{P}_0^* = (a^{S*}, a^{B*}, 0, 0)$, denote $N_U^{S*} = 1 - G^S(a^{S*})$; $N_U^{B*} = 1 - G^B(a^{B*})$

Thus we have

$$\begin{array}{ll}
\frac{\partial N_U^S}{\partial a_U^S}(\mathbf{P}_0^*) = -g^S(a^{S*}); & \frac{\partial N_U^S}{\partial a_U^B}(\mathbf{P}_0^*) = 0; \\
\frac{\partial N_U^B}{\partial a_U^S}(\mathbf{P}_0^*) = 0; & \frac{\partial N_U^B}{\partial a_U^B}(\mathbf{P}_0^*) = -g^B(a^{B*}); \\
\frac{\partial N_U^{SB}}{\partial a_U^S}(\mathbf{P}_0^*) = -g^S(a^{S*}) \cdot [1 - G^{B|S}(a^{B*} | a^{S*})]; & \frac{\partial N_U^{SB}}{\partial a_U^B}(\mathbf{P}_0^*) = -g^B(a^{B*}) \cdot [1 - G^{S|B}(a^{S*} | a^{B*})]; \\
\frac{\partial N_U^{PS}}{\partial a_U^S}(\mathbf{P}_0^*) = -g^S(a^{S*}) \cdot G^{B|S}(a^{B*} | a^{S*}); & \frac{\partial N_U^{PS}}{\partial a_U^B}(\mathbf{P}_0^*) = g^B(a^{B*}) \cdot [1 - G^{S|B}(a^{S*} | a^{B*})]; \\
\frac{\partial N_U^{PB}}{\partial a_U^S}(\mathbf{P}_0^*) = g^S(a^{S*}) \cdot [1 - G^{B|S}(a^{B*} | a^{S*})]; & \frac{\partial N_U^{PB}}{\partial a_U^B}(\mathbf{P}_0^*) = -g^B(a^{B*}) \cdot G^{S|B}(a^{S*} | a^{B*}).
\end{array} \tag{12}$$

The first order conditions for the optimality of a^{S*} and a^{B*} are

$$\begin{aligned}
\frac{\partial \Pi_U}{\partial a^S}(\mathbf{P}_0^*) &= [(a^{S*} + a^{B*} - c)N_U^{B*} - F^S]\frac{\partial N_U^S}{\partial a_U^S}(\mathbf{P}_0^*) + N_U^{S*}N_U^{B*} - (F^S + F^B - F)\frac{\partial N_U^{PB}}{\partial a_U^S}(\mathbf{P}_0^*) = 0 \\
\frac{\partial \Pi_U}{\partial a^B}(\mathbf{P}_0^*) &= [(a^{S*} + a^{B*} - c)N_U^{S*} - F^B]\frac{\partial N_U^B}{\partial a_U^B}(\mathbf{P}_0^*) + N_U^{S*}N_U^{B*} - (F^S + F^B - F)\frac{\partial N_U^{PS}}{\partial a_U^B}(\mathbf{P}_0^*) = 0
\end{aligned}$$

which reduce to

$$(a^{S*} + a^{B*} - c)N_U^{B*} - F^S = (F^S + F^B - F)\frac{\frac{\partial N_U^{PB}}{\partial a_U^S}(\mathbf{P}_0^*)}{\frac{\partial N_U^S}{\partial a_U^S}(\mathbf{P}_0^*)} - \frac{N_U^{S*}N_U^{B*}}{\frac{\partial N_U^S}{\partial a_U^S}(\mathbf{P}_0^*)} \tag{13a}$$

$$(a^{S*} + a^{B*} - c)N_U^{S*} - F^B = (F^S + F^B - F)\frac{\frac{\partial N_U^{PS}}{\partial a_U^B}(\mathbf{P}_0^*)}{\frac{\partial N_U^B}{\partial a_U^B}(\mathbf{P}_0^*)} - \frac{N_U^{S*}N_U^{B*}}{\frac{\partial N_U^B}{\partial a_U^B}(\mathbf{P}_0^*)} \tag{13b}$$

Step 2: Find $\frac{\partial}{\partial A}\Pi_H(a^{S*}, a^{B*}, 0)$.

Denote hybrid bundling strategy $\mathbf{P}_{H0} = (a^{S*}, a^{B*}, 0)$. By Proposition 1, we have $\frac{\partial N_H^{SB}}{\partial A}(\mathbf{P}_{H0}) = 0$. Substituting $\mathbf{P}_{H0}$ in (9) and (10) (note $A = 0$ in $\mathbf{P}_{H0}$), we have

$$\frac{\partial N_H^{PS}}{\partial A}(\mathbf{P}_{H0}) = \frac{\partial N_H^S}{\partial A}(\mathbf{P}_{H0}) = -\frac{g^S(a^{S*}) \cdot G^{B|S}(a^{B*} | a^{S*})}{N_U^{B*}} \tag{14a}$$

$$\frac{\partial N_H^{PB}}{\partial A}(\mathbf{P}_{H0}) = \frac{\partial N_H^B}{\partial A}(\mathbf{P}_{H0}) = -\frac{g^B(a^{B*}) \cdot G^{S|B}(a^{S*} | a^{B*})}{N_U^{S*}} \tag{14b}$$

And by Lemma 2 we have

$N_H^S = N_U^{S*} = 1 - G^S(a^{S*})$

$N_H^B = N_U^{B*} = 1 - G^B(a^{B*})$

$N_H^{SB} = N_U^{SB*} = \int_{a^{S*}}^{+\infty}\int_{a^{B*}}^{+\infty} g(s,b)dbds$

Since $\Pi_H(P_H) = (A - F^S)N_H^{PS} + (A - F^B)N_H^{PB} + (A - F)N_H^{SB} + (a_H^S + a_H^B - c)N_H^S N_H^B$, we have

$$\frac{\partial}{\partial A}\Pi_H(\mathbf{P}_{H0}) = \frac{\partial N_H^{PS}}{\partial A}(\mathbf{P}_{H0})[(a^{S*}+a^{B*}-c)N_U^{B*}-F^S]+N_U^{PS*} \\ +\frac{\partial N_H^{PB}}{\partial A}(\mathbf{P}_{H0})[(a^{S*}+a^{B*}-c)N_U^{S*}-F^B]+N_U^{PB*} \\ +N_U^{SB*} \tag{15}$$

Substituting (13) into (15), we get

$$\frac{\partial}{\partial A}\Pi_H(\mathbf{P}_{H0}) = (F^S+F^B-F)\frac{\frac{\partial N_U^{PB}}{\partial a_U^S}(\mathbf{P}_0^*)\frac{\partial N_H^{PS}}{\partial A}(\mathbf{P}_{H0})}{\frac{\partial N_U^S}{\partial a_U^S}(\mathbf{P}_0^*)} - N_U^{S*}N_U^{B*}\frac{\frac{\partial N_H^{PS}}{\partial A}(\mathbf{P}_{H0})}{\frac{\partial N_U^S}{\partial a_U^S}(\mathbf{P}_0^*)}+N_U^{PS*} \tag{16}$$

$$+(F^S+F^B-F)\frac{\frac{\partial N_U^{PS}}{\partial a_U^B}(\mathbf{P}_0^*)\frac{\partial N_H^{PB}}{\partial A}(\mathbf{P}_{H0})}{\frac{\partial N_U^B}{\partial a_U^B}(\mathbf{P}_0^*)} - N_U^{S*}N_U^{B*}\frac{\frac{\partial N_H^{PB}}{\partial A}(\mathbf{P}_{H0})}{\frac{\partial N_U^B}{\partial a_U^B}(\mathbf{P}_0^*)}+N_U^{PB*}$$

$$+N_U^{SB*}$$

Then substituting (12) and (14) into (16), we have

$$\frac{\partial}{\partial A}\Pi_H(\mathbf{P}_{H0}) = (F^S+F^B-F)[\frac{g^S(a^{S*})G^{B|S}(a^{B*}|\,a^{S*})(1-G^{B|S}(a^{B*}|\,a^{S*}))}{1-G^B(a^{B*})} \\ +\frac{g^B(a^{B*})G^{S|B}(a^{S*}|\,a^{B*})(1-G^{S|B}(a^{S*}|\,a^{B*}))}{1-G^S(a^{S*})}] \\ +\int_{a^{S*}}^{+\infty}\int_{a^{B*}}^{+\infty}[g^S(s)g^{B|S}(b|\,a^{S*})+g^B(b)g^{S|B}(s|\,a^{B*})-g(s,b)]dbds \tag{17}$$

This is exactly the left-hand side of condition (2). Thus we are done.■

Corollary 1

This case is implied by part (iii) of Corollary 2. See proof of the latter below.■

Corollary 2

We first prove the following lemma:

Lemma 4 *At* $\mathbf{P}_0^*=(a^{S*},a^{B*},0,0)$, $N_U^{SB*}=0$ ***if and only if*** $G^{S|B}(a^{S*}|\,a^{B*})=G^{B|S}(a^{B*}|\,a^{S*})$ $=1$.

Proof. Step 1: "Only if"

Since $N_U^{SB*}=\int_{a^{B*}}^{+\infty}\int_{a^{S*}}^{+\infty}g(s,b)dsdb$, and g has a weakly convex support $\mathbb{V}$ with full dimension in $\mathbb{R}^2$, $N_U^{SB*}=0$ implies that the point (a^{S*},a^{B*}) is either on the boundary or outside of $\mathbb{V}$. Thus we must have $g(s, a^{B*}) = 0$, for all $s \in (a^{S*}, +\infty)$, which implies $1 - G^{S|B}(a^{S*}| a^{B*}) = \int_{a^{S*}}^{+\infty} \frac{g(s,a^{B*})}{g^B(a^{B*})} ds$ =0, and thus $G^{S|B}(a^{S*}| a^{B*}) = 1$.

Similarly, we must have $g(a^{S*}, b) = 0$, for all $b \in (a^{B*}, +\infty)$ which implies $G^{B|S}(a^{B*} | a^{S*})$ $= 1$.

Step 2: "If"

$G^{S|B}(a^{S*}| a^{B*}) = 1$ implies $1 - G^{S|B}(a^{S*}| a^{B*}) = \int_{a^{S*}}^{+\infty} \frac{g(s,a^{B*})}{g^B(a^{B*})} ds$ =0 which in turn implies $g(s, a^{B*}) = 0$ for all $s \in (a^{S*}, +\infty)$. Similarly $G^{B|S}(a^{B*}| a^{S*}) = 1$ implies $g(a^{S*}, b) = 0$ for all $b \in (a^{B*}, +\infty)$. Then by convexity of $\mathbb{V}$, we must have $g(s, b) = 0$ for all $(s, b) \in (a^{S*}, +\infty) \times (a^{B*}, +\infty)$, thus $N_U^{SB*} = \int_{a^{B*}}^{+\infty} \int_{a^{S*}}^{+\infty} g(s, b) ds db = 0$. ■

Now we use Lemma 4 to prove Corollary 2:

(i) For all $b > a^{B*}$, since $G^{S|B}(a^{S*}| b)$ is strictly increasing in b and bounded from above by 1, we must have $G^{S|B}(a^{S*}| a^{B*}) < 1$ and

$$1 - G^{S|B}(a^{S*}| b) < 1 - G^{S|B}(a^{S*}| a^{B*}) \tag{18}$$

Denote $I \equiv 1 - G^{S|B}(a^{S*}| a^{B*})$, then I is a positive constant.

$(18) \Rightarrow \int_{a^{S*}}^{+\infty} \frac{g(s,b)}{g^B(b)} ds < I \Rightarrow \int_{a^{S*}}^{+\infty} g(s, b) ds \leq g^B(b) \cdot I$

Integrate both sides with respect to b on $[a^{B*}, +\infty)$, we get

$\Rightarrow \int_{a^{B*}}^{+\infty} \int_{a^{S*}}^{+\infty} g(s, b) ds db < \int_{a^{B*}}^{+\infty} g^B(b) db \cdot I \Rightarrow$

$$N_U^{SB*} < N_U^{B*}[1 - G^{S|B}(a^{S*}| a^{B*})] \tag{19}$$

Thus by (16), $F^S + F^B \geq F$, $N_U^{S*}[1 - G^{B|S}(a^{B*}| a^{S*})] \geq 0$ and (19), we have:

$$\begin{aligned} & \frac{\partial}{\partial A} \Pi_H(\mathbf{P}_{H0}) \\ \geq \; & N_U^{S*}[1 - G^{B|S}(a^{B*}| a^{S*})] + N_U^{B*}[1 - G^{S|B}(a^{S*}| a^{B*})] - N_U^{SB*} \\ > \; & 0. \end{aligned}$$

(ii) Using symmetry of our model regarding the two sides of the market, in this part we only need to relabel all the notations used in (i) (i.e. swap "S" and "B") and we are done.

(iii) By Lemma 4, $G^{S|B}(a^{S*}| a^{B*}) < 1$ implies $N_U^{SB*} > 0$.

Since both $G^{S|B}(a^{S*}| b)$ and $G^{B|S}(a^{B*}| s)$ are continuous at $(s, b) = (a^{S*}, a^{B*})$, $G^{S|B}(a^{S*}| b)$ being a constant for all $b > a^{B*}$ implies $G^{S|B}(a^{S*}| b) = G^{S|B}(a^{S*}| a^{B*})$ for all $b \geq a^{B*}$, and $G^{B|S}(a^{B*}| s)$ being a constant for all $s > a^{S*}$ implies $G^{B|S}(a^{B*}| s) = G^{B|S}(a^{B*}| a^{S*})$ for all $s \geq a^{S*}$.

By the same argument as in (i), except that all the inequalities need to be changed to equations since now $G^{S|B}(a^{S*} | b)$ is a constant, we get the equality version of (19):

$$N_U^{SB*} = N_U^{B*}[1 - G^{S|B}(a^{S*} | a^{B*})]$$

Similarly, from $G^{B|S}(a^{B*} | s)$ being a constant and $G^{B|S}(a^{B*} | a^{S*}) < 1$ we have

$$N_U^{SB*} = N_U^{S*}[1 - G^{B|S}(a^{B*} | a^{S*})]$$

Thus we have

$$\begin{aligned} & \frac{\partial}{\partial A}\Pi_H(\mathbf{P}_{H0}) \\ \geq \quad & N_U^{S*}[1 - G^{B|S}(a^{B*} | a^{S*})] + N_U^{B*}[1 - G^{S|B}(a^{S*} | a^{B*})] - N_U^{SB*} \\ = \quad & N_U^{SB*} > 0. \blacksquare \end{aligned}$$

Corollary 3

When $N_U^{SB*} = 0$, by Lemma 4 we know it must be $G^{S|B}(a^{S*} | a^{B*}) = G^{B|S}(a^{B*} | a^{S*})$ $= 1$. Then from (16) we know that both terms on the left-hand side of condition (2) are equal to 0, thus condition (2) does not hold.■

Proposition 3

Assumption NC does not change the demand functions, but it does change the cost structure and hence the profit functions under both unbundled sales and hybrid bundling. The new profit functions are:

At unbundled-sales price $\mathbf{P}_U = (a_U^S, a_U^B, A^S, A^B)$,
$\Pi_U^{NC}(\mathbf{P}_U) = (a_U^S + a_U^B - c)N_U^S N_U^B + (A^S - F^S)N_U^S + (A^B - F^B)N_U^B$
And at hybrid bundling price $\mathbf{P}_H = (a_H^S, a_H^B, A)$,
$\Pi_H^{NC}(\mathbf{P}_H) = (a_H^S + a_H^B - c)N_U^S N_U^B + (A - F)N_H$
By definition, at $\mathbf{P}_0^* = (a^{S*}, a^{B*}, 0, 0)$ the platform achieves the highest unbundled-sales profit, which is

$$\Pi_U^{NC}(\mathbf{P}_0^*) = (a^{S*} + a^{B*} - c)N_U^{S*} N_U^{B*} - F^S N_U^{S*} - F^B N_U^{B*} \tag{20}$$

where $N_U^{S*} = 1 - G^S(a^{S*})$ and $N_U^{B*} = 1 - G^B(a^{B*})$.

By Lemma 2, the platform's profit at hybrid bundling price $\mathbf{P}_{H0} = (a^{S*}, a^{B*}, 0)$ is

$$\Pi_H^{NC}(\mathbf{P}_{H0}) = (a^{S*} + a^{B*} - c)N_U^{S*} N_U^{B*} - F(N_U^{S*} + N_U^{B*} - N_U^{SB*}) \tag{21}$$

Thus

$$\Pi_H^{NC}(\mathbf{P}_{H0}) - \Pi_U^{NC}(\mathbf{P}_0^*) = F N_U^{SB*} - (F - F^S)N_U^{S*} - (F - F^B)N_U^{B*} \tag{22}$$

(i) When $F^S \neq F^B$ and $\min(F^S, F^B) > 0$, without lost of generality, suppose $0 < F^S < F^B (\leq F)$. Then condition (3) implies

$m(\mathbf{P}_0^*) = \frac{N_U^{SB*}}{N_U^*} \geq \frac{F - \min(F^S, F^B)}{\min(F^S, F^B)} = \frac{F - F^S}{F^S}$

$\Rightarrow F^S N_U^{SB*} \geq (F - F^S) N_U^* \Rightarrow F N_U^{SB*} - (F - F^S) N_U^{SB*} \geq (F - F^S) N_U^*$

$\Rightarrow F N_U^{SB*} \geq (F - F^S)(N_U^* + N_U^{SB*}) = (F - F^S)(N_U^{S*} + N_U^{B*}) > (F - F^S) N_U^{S*} + (F - F^B) N_U^{B*}$

$\Rightarrow \Pi_H^{NC}(\mathbf{P}_{H0}) - \Pi_U^{NC}(\mathbf{P}_0^*) = F N_U^{SB*} - (F - F^S) N_U^{S*} - (F - F^B) N_U^{B*} > 0$. Done.

(ii) When $F^S = F^B > 0$, and condition (3) holds with strict inequality,

$m(\mathbf{P}_0^*) = \frac{N_U^{SB*}}{N_U^*} > \frac{F - \min(F^S, F^B)}{\min(F^S, F^B)} = \frac{F - F^S}{F^S}$

$\Rightarrow F^S N_U^{SB*} > (F - F^S) N_U^* \Rightarrow F N_U^{SB*} - (F - F^S) N_U^{SB*} > (F - F^S) N_U^*$

$\Rightarrow F N_U^{SB*} > (F - F^S)(N_U^* + N_U^{SB*}) = (F - F^S)(N_U^{S*} + N_U^{B*}) = (F - F^S) N_U^{S*} + (F - F^B) N_U^{B*}$

$\Rightarrow \Pi_H^{NC}(\mathbf{P}_{H0}) - \Pi_U^{NC}(\mathbf{P}_0^*) > 0$. Done.

(iii) When $F^S = F^B > 0$ and $m(\mathbf{P}_0^*) = \frac{F - \min(F^S, F^B)}{\min(F^S, F^B)}$ the same argument in (ii) above with equations yields $\Pi_H^{NC}(\mathbf{P}_{H0}) = \Pi_U^{NC}(\mathbf{P}_0^*)$.

Now we prove the left-hand side of condition (4) is exactly $\frac{\partial}{\partial A} \Pi_H^{NC}(\mathbf{P}_{H0})$.

First, the first order conditions for $\mathbf{P}_0^*$ to be the optimal unbundled-sales strategy under NC are:

$$(a^{S*} + a^{B*} - c) N_U^{B*} - F^S = \frac{N_U^{S*} N_U^{B*}}{g^S(a^{S*})} \tag{23a}$$

$$(a^{S*} + a^{B*} - c) N_U^{B*} - F^B = \frac{N_U^{S*} N_U^{B*}}{g^B(a^{B*})} \tag{23b}$$

Now take the derivative of $\Pi_H^{NC}(\mathbf{P}_H)$ with respect to A, and evaluate at $\mathbf{P}_{H0} = (a^{S*}, a^{B*}, 0)$, we have

$$\begin{aligned} \frac{\partial}{\partial A} \Pi_H^{NC}(\mathbf{P}_{H0}) = {} & N_H^{PS} + N_H^{PB} + N_H^{SB} \\ & + \frac{\partial N_H^{PS}}{\partial A}(\mathbf{P}_{H0})[(a^{S*} + a^{B*} - c) N_U^{B*} - F] \\ & + \frac{\partial N_H^{PB}}{\partial A}(\mathbf{P}_{H0})[(a^{S*} + a^{B*} - c) N_U^{B*} - F] \end{aligned} \tag{24}$$

Substituting (14) and (23) in (24), we get

$$\begin{aligned} \frac{\partial}{\partial A} \Pi_H^{NC}(\mathbf{P}_{H0}) = {} & (F - F^S) \frac{g^S(a^{S*}) G^{B|S}(a^{B*} | a^{S*})}{1 - G^B(a^{B*})} + (F - F^B) \frac{g^B(a^{B*}) G^{S|B}(a^{S*} | a^{B*})}{1 - G^S(a^{S*})} \\ & + \int_{a^{S*}}^{+\infty} \int_{a^{B*}}^{+\infty} [g^S(s) g^{B|S}(b | a^{S*}) + g^B(b) g^{S|B}(s | a^{B*}) - g(s, b)] db ds \end{aligned}$$

which is exactly the left-hand side of condition (4).

Thus $\frac{\partial}{\partial A}\Pi_H^{NC}(\mathbf{P}_{H0}) > 0$ when condition (4) holds. And since we have already shown $\Pi_H^{NC}(\mathbf{P}_{H0}) = \Pi_U^{NC}(\mathbf{P}_0^*)$, we are done.

(iv) When $F^S = F^B = F = 0$, from (22) we have $\Pi_H^{NC}(\mathbf{P}_{H0}) = \Pi_U^{NC}(\mathbf{P}_0^*)$. And in (iii) we have shown that $\frac{\partial}{\partial A}\Pi_H^{NC}(\mathbf{P}_{H0}) > 0$ when condition (4) holds. Thus we are done.■

Proposition 4

Under Assumption BR, demand N_{BR}^S and N_{BR}^B at hybrid bundling price $\mathbf{P}_H = (a_H^S, a_H^B, A)$ is determined by the simultaneous system (5), with which we can get

$$\begin{aligned} N_{BR} &= 1 - \int_{-\infty}^{\frac{A}{N_{BR}^B}+a_H^S} \int_{-\infty}^{\frac{A}{N_{BR}^S}+a_H^B} g(s,b)dbds \\ N_{BR}^{SB} &= N_{BR}^S + N_{BR}^B - N_{BR} \end{aligned} \tag{25}$$

We only need to prove the following two results with BR, and the rest will just follow. Note that Lemmas 1 and 4 are not affected by BR since they are properties about the unbundled sales strategy only.

Lemma 2 with BR:

Let $\mathbf{P}_{U0} = (a^S, a^B, 0, 0)$ and $\mathbf{P}_H = (a^S, a^B, 0)$, then demand in any market segment converges under unbundled sales and hybrid bundling with BR, and so does platform profit. Thus we are done.

Proposition 1 with BR:

Simply take the first order derivatives of demands in (5) and (25) with respect to A at $\mathbf{P}_{H0} = (a^S, a^B, A = 0)$. It is straightforward to see that $\frac{\partial N_{BR}}{\partial A}(\mathbf{P}_{H0}) = \frac{\partial N_{BR}^S}{\partial A}(\mathbf{P}_{H0}) + \frac{\partial N_{BR}^B}{\partial A}(\mathbf{P}_{H0})$, and therefore $\frac{\partial N_{BR}^{SB}}{\partial A}(\mathbf{P}_{H0}) = \frac{\partial N_{BR}^S}{\partial A}(\mathbf{P}_{H0}) + \frac{\partial N_{BR}^B}{\partial A}(\mathbf{P}_{H0}) - \frac{\partial N_{BR}}{\partial A}(\mathbf{P}_{H0}) = 0$. This holds for any $(a^S, a^B) \in \mathbb{R}^{+2}$.■

Part II
Multiproduct Price Discrimination with Two-Part Tariffs*

Ming Gao†

First version: March 2009. This version: July 2010

Abstract

This paper gives a new "multiproduct" explanation of the wide application of two-part tariffs, complementary to the classical "single-product" efficiency-related explanation. We consider a monopolist provider of n (> 1) products who uses two-part tariffs consisting of a membership fee that is common to all consumers, and separate prices for different product bundles. We show that the change in demand for any bundle of $k \in [1, n]$ products caused by imposing an extra membership fee on top of any separate pricing strategy is proportional to the membership fee to the power of k. Therefore a small extra membership fee has no first-order impact on the demand for any multi-product bundles, which enables the firm to extract more consumer surplus. When this positive effect dominates the loss of single-product demand, two-part tariff dominates separate pricing. We present conditions that guarantee such an outcome, which generalize McAfee, McMillan and Whinston (1989)'s result from two products to multiple products. Our results suggest that two-part tariffs can achieve multidimensional price discrimination and should be subject to similar regulatory scrutiny as other discriminatory pricing strategies.

Key Words: two-part tariff, multiproduct pricing, price discrimination, bundling

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

*I thank Jean-Pierre Benoît, Emeric Henry, Eric Maskin, Preston McAfee and Joao Montez for valuable discussions and comments.

†Economics Department, London Business School, Regent's Park, London NW1 4SA, UK. Email: mgao.phd2005@london.edu. Homepage: phd.london.edu/mgao/.

Contents

1 Introduction

A two-part tariff is one of the most prevalent pricing strategies observed in real life. Whenever the total price that a consumer pays for consumption of a product or service consists of two parts instead of a one-off payment, a (general) two-part tariff applies. Telecommunication network providers, energy companies, fitness clubs, etc, all use two-part tariffs.

A typical *single-product* two-part tariff involves a *fixed part* and a *variable part*, both pertaining to the same product. For instance, a pay-as-you-go mobile network tariff in the UK normally consists of a fixed fee for access to the network (e.g. the cost of a SIM card and/or a phone), and an additional fee for each minute spent making a call[1]. A typical UK electricity/gas tariff also has a fixed part irrespective of usage (e.g. the "standing charge"), and a variable part dependent on usage. The classical explanation of the desirability of this kind of *single-product* two-part tariff is that it can reduce deadweight loss and hence increase profit (e.g. with homogeneous consumers, a monopolist can set the unit price of its product equal to marginal cost and capture all social surplus by an entry fee, which also achieves efficiency).

In this paper we study a different kind of two-part tariff involving *multiple products*, which consists of an *individual part* that may be different for different products, and a *common part* that applies to all products. For instance, a fitness club tariff may involve different pay-as-you-go fees for different gym activities or classes, plus a membership fee for access no matter what classes are taken.

We seek to answer the following question: *Aside from the classical "efficiency" benefits, what makes (multiproduct) two-part tariffs desirable to a multiproduct monopolist?*

The reason why we focus on two-part tariffs instead of, say, the optimal pricing strategy, is three-fold: i) in theory, there is no known method to identify the general optimal pricing strategy; ii) in real life, we rarely see pricing strategies that are as complex as the potential optimal pricing strategy might be (e.g. mixed bundling consisting of hundreds or thousands of prices), even for large numbers of products; and iii) in real life, two-part tariffs are easily implementable and are actually widely used. (We further discuss our motivation in section 2 in relation to the literature.)

Armstrong (1999) shows an *asymptotic* result that a cost-based two-part tariff can be "almost" optimal when the number of products is large. This result is very useful for understanding the pricing behavior of, say, supermarkets and bookshops, which have thousands of products to sell (although it is not straightforward to tell how well a cost-based two-part tariff describes the actual tariffs used by these firms in real life). However, this result does not help us much to understand why firms with much fewer products, such

[1]Note that in some countries, e.g. US and China, the service of receiving calls on mobile phones may not be free. If this is the case, it is useful to consider the reception service as a different product from the calling service, in which case the tariff becomes a *multiproduct* two-part tariff.

as a fitness center, would also want to use a two-part tariff.

In this paper we show that two-part tariffs can achieve *multiproduct price discrimination* and we uncover the underlying mechanism through which it is achieved, i.e. the *two-part-tariff effect.* This effect exploits the multi-dimensional nature of a multiproduct firm's demand, and has nothing to do with cost or efficiency. It applies to any two or more products, and holds for general forms of multivariate distributions. Our results therefore complement Armstrong (1999) and provide a more complete explanation for the prevalence of two-part tariffs in real life.

It is the common part of a two-part tariff that enables it to achieve multiproduct price discrimination. Consider for instance a fitness club that provides three classes: 1, 2 and 3. Suppose the firm originally only sells them separately at respective prices p_1, p_2 and p_3. Then a consumer of classes 2 and 3, say, needs to pay $p_2 + p_3$. Suppose, in addition to p_1, p_2 and p_3, the firm now charges an extra membership fee m to everyone who wants to use its facilities at all. Now a consumer has to pay total $q_2 = p_2 + m$ for class 2 alone, and $q_3 = p_3 + m$ for class 3 alone. But classes 2 and 3 together now cost $p_2 + p_3 + m = q_2 + q_3 - m$, cheaper (by m) than separate purchases of classes 2 and 3. Moreover, attending all three classes together now costs $p_1+p_2+p_3+m = q_1+q_2+q_3-2m$ (where $q_1 = p_1+m$), even cheaper (by $2m$) compared to separate purchases of three classes. This is a form of multiproduct price discrimination, and it is achieved because consumers of any combinations of classes pay the same *membership fee m*. We show that when m is small, it leads to a decrease proportional to m^2 in the demand for any two-class bundles, and a decrease proportional to m^3 in the demand for all three classes together. Such demand decreases are higher-order effects than the increase in total membership fee collected from the consumers of these classes. Therefore the fitness club makes strictly more profits from these consumers.

We use a general model for a monopolist provider of n (> 1) products facing heterogeneous consumers, and study when he would find it profitable to use the kind of two-part tariff described in the previous example. In particular, we identify the demand implications of imposing an extra membership fee on top of separate-product pricing strategies, which we call the *two-part-tariff effect.* We show that the change in demand for any bundle of $k > 1$ products due to a small membership fee is proportional to the membership fee to the power of k. This is because, for a k-bundle consumer, the burden of the extra fee is shared by her valuations in k dimensions. Therefore, a small extra membership fee has no first-order impact on the demand for any multiproduct bundles, but surpluses extracted by the membership fee from consumers of such bundles are first-order gains. When these gains dominate the losses from single products, two-part tariff dominates separate pricing. We present conditions that guarantee such an outcome.

Our conditions generalize McAfee, McMillan and Whinston (1989)'s (henceforth MMW) results to the multiproduct case. Their paper addresses the case of two products and provides conditions for *mixed bundling* to strictly dominate separate pricing. The two-part tariff we study can be viewed as a particular way of mixed bundling, where the membership fee and its multiples serve as the "bundle discounts" in mixed bundling (as discussed in the example of three products previously).

Although both two-part tariffs and mixed bundling can achieve multiproduct price discrimination, they work through different mechanisms. MMW show that, offering a discount to the bundle of two products (down from the sum of their separate prices) will achieve the effect of increasing the demand for *both* products by just lowering *one* bundle price, thus increasing total profits. We show that imposing a small membership fee has *zero* first-order impact on the demand from *all* multiproduct consumers, thereby enabling more surplus extraction from them and increasing total profits.

There is a "surprising" contrast between these mechanisms - MMW's result on mixed bundling involves a profitable *decrease* of the final price charged to consumers, whereas our result on two-part tariff involves a profitable *increase* in the final price. This dichotomy implies that the *optimal* pricing strategy will *not necessarily* involve a "discount" (i.e. lower final price) compared to separate pricing, contrary to the impression that one might incorrectly derive from MMW's result. Indeed, we provide a two-product example in section 8.4.2 where the optimal MMW-type mixed bundling strategy results in higher final prices than separate pricing.

It is important to note that i) these two deviations are not exactly opposite to each other - the exact opposite mechanism of an additional membership fee on all consumers is the same amount of discount for all consumers, not just for multiproduct consumers; and ii) both deviations lead to *local* profit improvements, and one should be cautious when using them to infer properties of the *global* optimizers.

Since two-part tariffs can achieve multiproduct price discrimination, an implication of our results is that (multiproduct) two-part tariffs should be subject to similar antitrust scrutiny as other discriminatory pricing strategies, such as bundling.

The remainder of this paper is organized as follows: Section 2 discusses related literature and the position of this paper; Section 3 describes the model; Section 4 shows the effect of two-part tariffs on demand, i.e. the two-part-tariff effect; Section 5 compares profits under two-part tariffs and separate pricing, and provides conditions for the former to dominate the latter; Section 6 generalizes two-part tariffs to allow for negative membership fees; Section 7 discusses the usage of more than one membership fees; Section 8 compares two-part tariffs and general mixed bundling strategies, where we also provide results for specific distributions; Section 9 concludes.

2 Literature

This paper fits in the theoretical literature of multiproduct pricing, which is embedded in the larger literature of multi-dimensional mechanism design (especially multi-dimensional screening), and has evolved from the earlier literature on commodity bundling.

At its early stage, the bundling literature was first concerned with a two-product firm and two classic categories of pricing strategies: *pure bundling* (only selling two goods in bundles but not separately) and *mixed bundling* (providing two goods both separately and in bundles). Scholars have focused on understanding how these strategies can do better than separate pricing, and discussed the conditions with examples (Stigler (1963) and Adams and Yellen (1976)), with particular distributions (Schmalensee (1984)), and then with more general distributions (McAfee, McMillan and Whinston (1989)).

The general insights from these studies are: i) Pure bundling can yield higher profit than separate pricing because consumers' valuation of the bundle is generally *less dispersed* than their valuation of each separate good and hence pure bundling can increase the probability of trade at certain prices; ii) Mixed bundling can increase profit because by lowering the price of the bundle below the sum of the prices of separate goods (that is, lowering *one* price) a monopolist can increase demands for both goods (that is, increase *two* demands).

Afterwards the literature entered the era of multiple products. Along with the advancement of understanding in multi-dimensional screening (Rochet and Chone (1998)), scholars attempt to find the optimal multiproduct pricing strategy in a general setting. However, so far this goal has not been achieved. The closest results that have been found are characterizations of some properties of the optimal solution to the general screening problem (Rochet and Chone (1998)), those of the optimal non-linear pricing strategy (Armstrong (1996)) and those of the optimal mixed bundling strategy (Manelli and Vincent (2006)).

The first obstacle in identifying a solution is that the format of the optimal mechanism is still unknown. Although mixed bundling is generally the most intuitive candidate, Manelli and Vincent (2006) have shown that every multiproduct mixed bundling strategy may actually be dominated by a mechanism involving random assignments.

Even when the search is narrowed down to mixed bundling strategies, the general distributional assumption of consumers' valuations (of products) as well as the high dimensionality of the mathematical problem pose a second challenge - we only know some properties of the optimal strategy under some well-behaved distributions (see Manelli and Vincent (2006)), while the computational problem of finding the general solution is NP-complete[2] (Conitzer and Sandholm (2003)).

[2] NP-completeness (where NP stands for "nondeterministic polynomial time") implies that there is no known efficient way to locate a solution and the computational time required to solve the problem using any currently known algorithm increases very quickly as the size of the problem grows, and may easily reaches into the billions or trillions of years. See Madarász and Prat (2010) for a discussion of this problem.

Moreover, even if we can find the optimal mixed bundling strategy given the specific form of distribution, practicality becomes the third concern: For a supplier of n products, a mixed bundling strategy needs to specify $2^n - 1$ prices. This quickly makes mixed bundling impractical in real life as the number of products increases beyond even the lower tens (e.g. even if $n = 5$, there are still more than 30 prices involved).

In view of these difficulties and concerns, the literature has developed in two different directions:

The first direction pursues approximation of the optimal strategy through carefully designed computational algorithms (see Madarász and Prat (2010)).

The second direction returns to studying the mechanisms of particular kinds of simpler pricing strategies, and discusses the situations where these strategies may be desirable. Papers in this direction include Fang and Norman (2006) and Banal-Estanol and Ottaviani (2007) that focus on multiproduct pure bundling; and Chu, Leslie and Sorensen (2009) that focus on "bundle-size" pricing.

The current paper follows the second direction and focuses on multiproduct two-part tariffs. One nice feature of two-part tariff is that it is easily implementable: For a supplier of n products, a two-part tariff only uses $n + 1$ prices (which is comparable to the small number of prices used in "bundle-size" pricing, for instance). As discussed in section 1, our results are complementary to Armstrong (1999) and provide a more complete explanation for the prevalence of two-part tariffs in real life.

Long (1984) is the first in the bundling literature to make a connection between two-part tariffs and mixed bundling in the two product case. He shows that a two-part tariff consisting of prices of two separate products, q_1 and q_2, say, and one additional fixed fee for all consumers, m, is equivalent to a mixed bundling strategy with prices for separate products $p_1 = q_1 + m$, $p_2 = q_2 + m$ and price for the bundle $p_B = q_1 + q_2 + m$, where "*the bundle discount* ($p_1 + p_2 - p_B = m$) *is like a fixed fee*".

Our model and results generalize Long's discussion to the case of multiple products. Furthermore, we show that the bundle discount in mixed bundling is *not* equivalent to the membership fee in two-part tariffs, in either two-product case or multiproduct case, because the bundle discount does not apply to single-product consumers and thus cannot achieve the two-part-tariff effect (see section 8.3).

3 Model

There is only one firm, which produces $n \in \mathbb{N}$ different kinds of products. We use $j \in \{1, 2, ..., n\}$ to denote a **product**. There is no cost of production. The firm maximizes total profit.

There is a continuum of **consumers**, each of whom has a **valuation** for each of these products (i.e. the utility she derives from the product) and demands 0 or 1 unit of each product (i.e. 2 or more units of any one product will provide the exact same utility as 1 unit of that product). The total utility a consumer derives from consuming various products is simply the sum of her valuations for these products. No consumption results in zero utility. We denote a consumer's **type** by an n-dimensional real-valued parameter $\mathbf{x} \equiv (x_1, ..., x_n)$, where x_j is this consumer's valuation for product j.

The firm does not know each consumer's type. Rather, it has prior (density) $f(\mathbf{x})$ of the distribution of $\mathbf{x}$ among consumers. The support of f is denoted by $\mathbf{S} \equiv \times_{j=1}^{n} S_j \subset \mathbb{R}^n$, where S_j is the support in dimension j (i.e. for product j).

Assumption 1 *f is atomless, and $\mathbf{S}$ is weakly convex and has full dimension in $\mathbb{R}^n$.*

For expository simplicity and to allow for a direct comparison with Manelli and Vincent (2006), in this paper we focus on the case when $\mathbf{S} = [0,1]^n \equiv I^n$. All the results can easily be generalized for general $\mathbf{S}$ that satisfies Assumption 1.[3]

Assumption 2 *f is atomless, and $f(\mathbf{x}) > 0$ if and only if $\mathbf{x} \in I^n$.*

A **bundle** is a set of different products. We denote the full bundle of all n products as $N = \{1, 2, ..., n\}$. Any bundle, denoted by J, is therefore a subset of N, i.e. $J \subset N$. The empty bundle is $\varnothing \subset N$.

When it does not cause confusion, we also use j to represent the bundle $\{j\}$ (i.e. containing only product j). And j^c simply means $\{j\}^c$.

For bundle J, we denote the **bundle size** (i.e. the number of products in it) by $|J|$.

A general rule we use in the notation below is: superscript represents dimensionality; subscript represents bundle or product.

Definition 1 (Price Schedule) *A **price schedule** $\mathbf{P}$ specifies the price for each possible bundle, $\mathbf{P} \equiv \{p_J\}_{J \subset N}$, where $p_J \in \mathbb{R}^+$ for any $J \subset N$.*

Note that $\mathbf{P}$ consists of 2^n prices since there are 2^n possible bundles (including the full bundle and the empty bundle).

Definition 2 (IC) *Consider any K different subsets of N, denoted $\{J_k\}_{k=1,2,...,K}$ where $J_k \subset N$ $\forall k = 1, 2, ..., K$. A price schedule $\mathbf{P} = \{p_J\}_{J \subset N}$ is **incentive compatible** (IC) if the following condition holds for all $K = 1, 2, 3, ..., 2^n$*

$$p_{\bigcup_{k=1}^{K} J_k} \leq \sum\nolimits_{k=1}^{K} p_{J_k}$$

[3] Our notation also mostly follows that of Manelli and Vincent (2006).

Interpretation In an IC price schedule, the price of a bundle would not exceed the sum of the prices of any "profile" of its sub-bundles that forms a full "cover" of this bundle, otherwise no consumer would ever demand this bundle. IC is a necessary condition for each bundle to attract some demand.

Since a *partition* of a bundle is a profile of sub-bundles that form a full "cover" of this bundle, IC therefore implies that *the price of a bundle in an IC price schedule would not exceed the sum of the sub-bundle prices in any partition of this bundle.*

Our discussion from now on focuses only on IC price schedules.

Definition 3 (Additivity/Separate Pricing) *A price schedule* $\mathbf{P} = \{p_J\}_{J \subset N}$ *is* ***additive*** *(or* ***separate pricing****) if* $p_\varnothing = 0$ *and* $p_J = \sum_{j \in J} p_j$ *for any non-empty* $J \subset N$*.*

$p_\varnothing = 0$ is actually a constraint on all price schedules that satisfy consumers' individual rationality (IR). Note that this definition implies the following result.

Lemma 1 *If* $\mathbf{P} = \{p_J\}_{J \subset N}$ *is additive, then* $p_J + p_K = p_{J \cup K} + p_{J \cap K}$ *for any two bundles* J *and* K*.*

Proof. By additivity of $\mathbf{P}$, we have

$$\begin{aligned} p_{J \cup K} + p_{J \cap K} &= \sum_{i \in J \cup K} p_i + \sum_{h \in J \cap K} p_h \\ &= \sum_{j \in J} p_j + \sum_{l \in K \backslash J} p_i + \sum_{h \in J \cap K} p_h \\ &= \sum_{j \in J} p_j + \sum_{k \in K} p_k \\ &= p_J + p_K. \end{aligned}$$

■

Lemma 2 *Additivity implies IC.*

Proof. Apply Lemma 1 on any two subsets of N, say J_1 and J_2, and we have

$p_{J_1 \cup J_2} + p_{J_1 \cap J_2} = p_{J_1} + p_{J_2}$

Since $p_{J_1 \cap J_2} \geq 0$, we know

$p_{J_1 \cup J_2} \leq p_{J_1} + p_{J_2}$

Now apply Lemma 1 to a third subset J_3 and $J_1 \cup J_2$, and in turn we have

$p_{J_1 \cup J_2 \cup J_3} \leq p_{J_1 \cup J_2 \cup J_3} + p_{(J_1 \cup J_2) \cap J_3} = p_{J_1 \cup J_2} + p_{J_3} \leq p_{J_1} + p_{J_2} + p_{J_3}$

We therefore can continue to add additional subsets one by one in the same way as above and show that $p_{\bigcup_{k=1}^{K} J_k} \leq \sum_{k=1}^{K} p_{J_k}$ holds for all $K = 1, 2, 3, ..., 2^n$. ■

Definition 4 (Demand Segment) *Given IC price schedule* $\mathbf{P} = \{p_J\}_{J \subset N}$*, the* ***demand segment*** *for any bundle* $J \subset N$*, denoted* A_J*, is the set of all the consumers that buy bundle* J*:*

$$A_J \equiv \{\mathbf{x} \in I^n | \sum_{j \in J} x_j - p_J \geq \sum_{k \in K} x_k - p_K, \forall K \subset N\}$$

Note: From this definition, all demand segments are *closed*, and their intersections define their "boundaries".

Definition 5 (Allocation) *A (consumer)* ***allocation*** *given price schedule* $\mathbf{P}$ *is the profile of demand segments of all bundles induced by* $\mathbf{P}$*, denoted* $\{A_J\}_{J \subset N}$*.*[4]

Lemma 3 (Additive Allocation) *If* $\mathbf{P} = \{p_J\}_{J \subset N}$ *is additive, the allocation it induces* $\{A_J\}_{J \subset N}$ *must satisfy for any* $J \subset N$

$$A_J = \{\mathbf{x} \in I^n | x_j \geq p_j, \forall j \in J; x_k \leq p_k, \forall k \in J^c\} \tag{1}$$

Intuition An additive price schedule allocates all consumers into "cubes" delineated by orthogonal hyperplanes.

Proof.

What we need to show is that, when $\mathbf{P} = \{p_J\}_{J \subset N}$ is additive, the following two conditions are equivalent for any $J \subset N$:

(i) $\sum_{j \in J} x_j - p_J \geq \sum_{k \in K} x_k - p_K, \forall K \subset N$;

(ii) $x_j \geq p_j, \forall j \in J; x_k \leq p_k, \forall k \in J^c$.

The property we need to prove the equivalence is exactly additivity of $\mathbf{P}$: that is, $p_\varnothing = 0$ and $p_J = \sum_{j \in J} p_j$ for any $J \neq \varnothing$.

First, it is easy to see that for $J = \varnothing$, both (i) and (ii) reduce to: $x_k \leq p_k, \forall k \in N$ and are therefore equivalent.

Now consider $J \neq \varnothing$:

Step 1: (i)$\Rightarrow$(ii)

Since $J \neq \varnothing$, we must be able to partition J into some $\{i\}$ and K where $i \in J$, $K \subset N$, $i \notin K$, and $J = \{i\} \cup K$ (note that K may be $\varnothing$).

Since (i) holds for any $K \subset N$, it must hold for this K. By additivity of $\mathbf{P}$ we know $p_J = p_i + p_K$. Therefore we have

$\sum_{j \in J} x_j - \sum_{k \in K} x_k \geq p_J - p_K \Leftrightarrow x_i \geq p_i$.

As this i can be any element in J, we have proved the first part of (ii).

To prove the second part of (ii), first notice that it only matters for $J \neq N$ (and does not exist for $J = N$). Therefore we can consider $L \equiv J \cup \{h\}$ where $h \in J^c$. By additivity of $\mathbf{P}$ we know $p_L = p_J + p_h$. Since (i) must hold for this L as well (as $L \subset N$), we have

[4] To lighten notation, we do not carry $\mathbf{P}$ in A_J or $\{A_J\}_{J \subset N}$, but it is always implied that a demand segment or allocation is induced by some price schedule.

$p_L - p_J \geq \sum_{l \in L} x_l - \sum_{j \in J} x_j \Leftrightarrow p_h \geq x_h$

And this holds for any $h \in J^c$. Done.

Step 2: (ii)⇒(i)

For any $K \subset N$, by additivity of $\mathbf{P}$ and (ii) we have:

$$
\begin{aligned}
& (\sum_{j \in J} x_j - p_J) - (\sum_{k \in K} x_k - p_K) \\
= \; & (\sum_{j \in J} x_j - \sum_{k \in K} x_k) - (p_J - p_K) \\
= \; & (\sum_{i \in J \backslash K} x_i - \sum_{h \in K \backslash J} x_h) - (p_{J \backslash K} - p_{K \backslash J}) \\
\geq \; & (\sum_{i \in J \backslash K} p_i - \sum_{h \in K \backslash J} p_h) - (p_{J \backslash K} - p_{K \backslash J}) \\
= \; & 0
\end{aligned}
$$

which implies (i). Done. ■

Definition 6 (Truncated Type) *Given any bundle $J \subset N$, a $J-$**truncated type parameter** is denoted $\mathbf{x}^J = \{x_j\}_{j \in J} \in I^J$, where $I^J \equiv \times_{j \in J} I_j$.*

We sometimes use x_j^J (where $j \in J$) to denote the element of $\mathbf{x}^J$ pertaining to product j.

Notice $\mathbf{x}^J$ is a $|J|$-dimensional vector (or a point) in I^J. We use J instead of $|J|$ as the superscript of $\mathbf{x}^J$ to emphasize that $\mathbf{x}^J$ keeps the dimensions in I^n *according to* bundle J, rather than any $|J|$ dimensions of I^n. This distinction is important for the following definitions.

Definition 7 (Projection) *For any (consumer set) $A \subset I^n$ and any bundle $K \neq \varnothing$, define*

$$A^K \equiv \{\mathbf{x}^K \in I^K | (\mathbf{x}^K, \mathbf{y}) \in A, \text{ for some } \mathbf{y} \in I^{K^c}\}$$

*which is the **projection of set** A on the $|K|$-dimensional hyperplane defined by the following $|K^c|$ equations:*

$$\{x_j^{K^c} = 0\}_{j \in K^c} \tag{2}$$

Note that when $K = N$, because $|N^c| = 0$, the projection operation above is the identity mapping, i.e. $A^N = A$.

Definition 8 (Projection of A_J) *For any two bundles $K, J \neq \varnothing$, the set*

$$A_J^K \equiv \{\mathbf{x}^K \in I^K | (\mathbf{x}^K, \mathbf{y}) \in A_J, \text{ for some } \mathbf{y} \in I^{K^c}\}$$

*is the **projection** of set A_J on the $|K|$-dimensional hyperplane defined by the $|K^c|$ equations of (2).*

Definition 9 (Probability Measure) *For any $A \subset I^n$, we define the probability measure of A as $M(A)$ which satisfies*

$$M(A) = \int_A f(\mathbf{x}) d\mathbf{x}$$

We use $M^J(A) = \int_A f(\mathbf{x}^J) d\mathbf{x}^J$ to denote the **marginal measure** in I^J of set A, for any $J \subset N$, which is particularly useful when A does not have full dimension in I^n but has full dimension in I^J.

4 Two-Part Tariffs and the Two-Part-Tariff Effect

Definition 10 (Two-Part Tariff) *A two-part tariff is a price schedule $\mathbf{Q} = \{q_J\}_{J \subset N}$ consisting of two parts $(m, \mathbf{P})$, where*

$$\begin{aligned} & m > 0\text{; and} \\ & \mathbf{P} = \{p_J\}_{J \subset N} \text{ is additive; and} \\ & q_J = \begin{cases} p_J + m & \text{, if } J \neq \varnothing \\ 0 & \text{, if } J = \varnothing \end{cases} \end{aligned} \tag{3}$$

And the allocations induced by $\mathbf{P}$ and $\mathbf{Q}$ are denoted $\{A_J\}_{J \subset N}$ and $\{C_J\}_{J \subset N}$, respectively.

Comment A two-part tariff $\mathbf{Q}$ consists of a *common part*, m, which is the *membership fee* that applies to all customers, and an *individual part*, $\mathbf{P}$, which is an additive price schedule of the prices of separate products. Compared to $\mathbf{P}$, $\mathbf{Q}$ simply *increases* the prices of all *non-empty* bundles by the *same amount* m. Since $\mathbf{P}$ is additive, $\mathbf{Q}$ will not be additive.

As we discussed in section 1, the two-part tariff in (3) achieves multiproduct price discrimination. To see this, suppose under $\mathbf{Q}$ consumer a demands bundle $\{1\}$ (by paying $q_1 = p_1 + m$), consumer b demands bundle $\{2\}$ (by paying $q_2 = p_2 + m$), and consumer c demands bundle $\{1, 2\}$ (by paying $q_{\{1,2\}} = p_{\{1,2\}} + m = p_1 + p_2 + m = q_1 + q_2 - m$). Compared to a and b, it is as if c gets a "discount" of m by buying two products together, since c only needs to pay membership fee m *once*. Actually, it is easy to see that under $\mathbf{Q}$ a consumer of any bundle J gets a "discount" of $(|J| - 1) \cdot m$ compared to the consumers of the $|J|$ individual products. This is a special feature of the two-part tariff we defined in (3).

4.1 Uniform Distribution

Theorem 1 *Consider the price schedules and allocations defined in (3). If* $\mathbf{x}$ *is* ***uniformly distributed****, then for* ***any*** $J \neq \varnothing$*, we have*

$$M(C_J) - M(A_J) = c(\mathbf{P}) \cdot m^{|J|}$$

where $c(\mathbf{P})$ *is a function of* $\mathbf{P}$ *(but not of* m*).*

That is, the demand change of any bundle J *due to an additional membership fee* m *on top of separate pricing is proportional to* m *to the power of the bundle size* $|J|$*.*

Intuition Starting from an additive price schedule, for a consumer of bundle J, the burden of an extra membership fee is shared by her valuations of all $|J|$ products in the bundle.

Note that the demand change, $M(C_J) - M(A_J)$, is non-positive because $C_J \subset A_J$ as we will show below.

Proof. Since $\mathbf{P}$ is additive, by Lemma 3 we know the allocation induced by $\mathbf{P}$ is defined in (1), that is:

$$A_J = \{\mathbf{x} \in I^n | x_j \geq p_j, \forall j \in J; x_k < p_k, \forall k \in J^c\}$$

And by Definition 4 and the definition of $\mathbf{Q}$ in (3), we know the allocation $\mathbf{Q}$ induces is

$$\begin{aligned} C_\varnothing &= \{\mathbf{x} \in I^n | \sum_{k \in K} x_k < p_K + m, \forall K \neq \varnothing, K \subset N\} \\ &= \{\mathbf{x} \in I^n | \sum_{k \in K} x_k < \sum_{k \in K} p_k + m, \forall K \neq \varnothing, K \subset N\}; \text{ and} \end{aligned}$$

$$\begin{aligned} C_{J(\neq\varnothing)} &= \{\mathbf{x} \in I^n | \sum_{j \in J} x_j \geq p_J + m; \sum_{j \in J} x_j - p_J \geq \sum_{k \in K} x_k - p_K, \forall K \subset N\} \\ &= \{\mathbf{x} \in I^n | \sum_{j \in J} x_j \geq p_J + m; x_j \geq p_j, \forall j \in J; x_k \leq p_k, \forall k \in J^c\} \end{aligned} \tag{4}$$

(For the last equation, see proof of Lemma 3 where we show (i)$\Leftrightarrow$(ii).)

Definition 11

$$A_J(m) \equiv \{\mathbf{x} \in A_J | 0 \leq \sum_{j \in J} x_j - p_J < m\}, \text{ for } J \subset N. \tag{5}$$

Therefore we have

$$\begin{aligned} A_J(m) &= \{\mathbf{x} \in A_J | 0 \leq \sum_{j\in J} x_j - p_J < m\} \quad (6) \\ &= \{\mathbf{x} \in I^n | x_j \geq p_j, \forall j \in J; \sum_{j\in J} x_j < \sum_{j\in J} p_j + m; x_k < p_k, \forall k \in J^c\} \end{aligned}$$

which implies:

Lemma 4 *$\forall J \subset N$, we have $A_J = C_J \cup A_J(m)$. That is, $A_J(m)$ is exactly the lost demand for bundle J when the price schedule changes from $\mathbf{P}$ to $\mathbf{Q}$.*

Rewrite this relationship in terms of measures and we have:

Lemma 5 *$\forall J \subset N$, we have $M(A_J(m)) = M(A_J \backslash C_J) = M(A_J) - M(C_J)$.*

Denote the projection of $A_J(m)$ in dimensions J (see Definition 8) as $A_J^J(m) \equiv (A_J(m))^J$, and that in dimensions J^c as $A_J^{J^c}(m) \equiv (A_J(m))^{J^c}$. Then by Definition 8, (1) and (6) we know:

$$\begin{aligned} A_J^J(m) &= \{\mathbf{x}^J \in I^J | (\mathbf{x}^J, \mathbf{y}) \in A_J(m), \text{for some } \mathbf{y} \in I^{J^c}\} \quad (7) \\ &= \{\mathbf{x}^J \in I^J | x_j \geq p_j, \forall j \in J; \sum_{j\in J} x_j < \sum_{j\in J} p_j + m\} \end{aligned}$$

From this expression we know $A_J^J(m)$ has full dimension in I^J, and each of its "$|J|$ sides" has "length" exactly equal to m.

By (6) we also know:

$$\begin{aligned} A_J^{J^c}(m) &= \{\mathbf{x}^{J^c} \in I^{J^c} | (\mathbf{x}^{J^c}, \mathbf{y}) \in A_J(m), \text{for some } \mathbf{y} \in I^J\} \quad (8) \\ &= \{\mathbf{x}^{J^c} \in I^{J^c} | x_k < p_k, \forall k \in J^c\} \\ &= \{\mathbf{x}^{J^c} \in I^{J^c} | 0 \leq x_k < p_k, \forall k \in J^c\} \end{aligned}$$

which implies:

Lemma 6

$$A_J^{J^c}(m) = A_J^{J^c}, \forall J \subset N. \quad (9)$$

That is, $A_J^{J^c}(m)$ does not depend on m for any $J \subset N$.

Lemma 7

$$A_J(m) = A_J^J(m) \times A_J^{J^c}, \forall J \subset N. \quad (10)$$

That is, when the price schedule changes from $\mathbf{P}$ to $\mathbf{Q}$, the demand change of any bundle J, $A_J(m)$, is the Cartesian product of its projections in own dimensions J and in complement dimensions J^c.

(Notice that our proof up until this point applies to any general distribution f satisfying Assumption 2.)

Therefore by mutual independence among all x_j's ($j \in N$) implied by uniform distribution, we have

$$M(A_J(m)) = M^J(A_J^J(m)) \cdot M^{J^c}(A_J^{J^c}) \tag{11}$$

Note: $M^J(\cdot)$ and $M^{J^c}(\cdot)$ are the marginal measures of M in dimensions J and dimensions J^c, respectively.

Now we need to find $M^J(A_J^J(m))$ and $M^{J^c}(A_J^{J^c})$.

By (7) and $f(\mathbf{x}) = 1, \forall \mathbf{x} \in I^n$, we have

$$M^J(A_J^J(m)) = \int_{A_J^J(m)} d\mathbf{x}^J = \frac{m^{|J|}}{|J|!} \tag{12}$$

By (8) we have

$$M^{J^c}(A_J^{J^c}(m)) = \int_{A_J^{J^c}(m)} d\mathbf{x}^{J^c} = \prod_{k \in J^c} p_k \tag{13}$$

Finally, putting (12) and (13) together, we have

$$M(C_J) - M(A_J) = -M(A_J(m)) = \frac{-\prod_{k \in J^c} p_k}{|J|!} \cdot m^{|J|} \tag{14}$$

where the first part $\frac{-\prod_{k \in J^c} p_k}{|J|!} \equiv c(\mathbf{P})$ is a function of $\mathbf{P}$ only (more precisely it is a function of $\mathbf{P}^{J^c}$ only), and does not depend on m. ■

Comment By (8) we know $M^{J^c}(A_J^{J^c}(m))$ only depends on $\mathbf{P}$ and distribution f, but does not depend on m. By (7) we know each of $A_J^J(m)$'s "$|J|$ sides" has "length" m, which means $M^J(A_J^J(m))$ will be proportional to $m^{|J|}$. Therefore $M(A_J(m))$ is also proportional to $m^{|J|}$.

Under uniform distribution, Theorem 1 tells us that when we impose an extra membership fee on top of an additive pricing strategy, the impact on the demand for any bundle is of order equal to the bundle size. This implies that, for multiproduct bundles, there is no first-order demand impact. This implication turns out to hold for any general distribution, as summarized in the following result.

4.2 General Distribution

Theorem 2 (Two-Part-Tariff Effect) *Consider the price schedules and allocations defined in (3). For* ***any general*** *f satisfying Assumption 2, we have*

(i) $\frac{\partial}{\partial m}[M(C_J) - M(A_J)]|_{m=0} \leq 0$, *for all J such that $|J| = 1$;*
(ii) $\frac{\partial}{\partial m}[M(C_J) - M(A_J)]|_{m=0} = 0$, *for all J such that $|J| > 1$.*

That is, imposing an additional membership fee on top of separate pricing has ***no*** *first-order impact on the demand for any multiproduct bundle.*

Intuition Similar to the results shown in Theorem 1, even when distribution f does not satisfy independence, when m is very small, we can still think of $M(A_J(m))$ as "proportional" to $m^{|J|}$, and thus its first order derivative with respect to m would be "proportional" to $m^{|J|-1}$, which goes to 0 as $m \to 0$, unless $|J| = 1$.

Proof. The expressions (6), (7) and (8) derived above hold for any general distribution f that satisfies Assumption 2. Although property (11) requires mutual independence among x_j's, we do not really need it here as all we care about is the first order derivative of $M(A_J(m))$, not $M(A_J(m))$ itself. Therefore Lemma 7 will suffice.

Part (i): when $|J| = 1$, i.e. $J = \{j\}$, $\forall j \in N$.

By (6) we know

$$\begin{aligned} A_j(m) &= \{\mathbf{x} \in I^n | p_j \le x_j < p_j + m; x_k < p_k, \forall k \in j^c\} \\ &= \times_{k \in j^c} \{x_k \in I^k | 0 \le x_k < p_k\} \times \{x_j \in I^j | p_j \le x_j < p_j + m\} \end{aligned}$$

where by (8) and (9) we know that $A_j^{j^c}(m) = A_j^{j^c} = \times_{k \in j^c} \{x_k \in I^k | 0 \le x_k < p_k\}$

Then by (10) we have

$$M(A_j(m)) = \int_{p_j}^{p_j+m} [\int_{A_j^{j^c}} f(\mathbf{x}^{j^c}, x_j) d\mathbf{x}^{j^c}] dx_j$$

Notice that the integral in the brackets above is a function of x_j and $\mathbf{P}^{j^c}$ only (it depends on $\mathbf{P}^{j^c}$ because of $A_j^{j^c}$), and does not depend on m. Therefore we can define it as

$$W_j(x_j, \mathbf{P}^{j^c}) \equiv \int_{A_j^{j^c}} f(\mathbf{x}^{j^c}, x_j) d\mathbf{x}^{j^c}$$

And rewrite

$$M(A_j(m)) = \int_{p_j}^{p_j+m} W_j(x_j, \mathbf{P}^{j^c}) dx_j$$

Therefore

$$\begin{aligned} \frac{\partial}{\partial m}(M(C_j) - M(A_j))|_{m=0} &= -\frac{\partial M(A_j(m))}{\partial m}|_{m=0} \\ &= -W_j(p_j + m, \mathbf{P}^{j^c})|_{m=0} \\ &= -W_j(p_j, \mathbf{P}^{j^c}) \\ &= -\int_{A_j^{j^c}} f(\mathbf{x}^{j^c}, p_j) d\mathbf{x}^{j^c} \\ &\le 0 \end{aligned}$$

Notice that the last inequality above will be strict if $\mathbf{P}$ in (3) satisfies $p_j > 0\ \forall j \in N$.

Part (ii): when $|J| > 1$. Since all n dimensions are "symmetric" in our setting, without loss of generality, we consider $J = \{1, 2, ..., |J|\}$.

By (9) and (10) we have

$$M(A_J(m)) = \int_{A_J(m)} f(\mathbf{x})d\mathbf{x} = \int_{A_J^{J^c}} [\int_{A_J^J(m)} f(\mathbf{x}^{J^c}, \mathbf{x}^J)d\mathbf{x}^J]d\mathbf{x}^{J^c}$$

Since $A_J^{J^c}$ does not depend on m (by (8) and (9)), we have

$$\frac{\partial M(A_J(m))}{\partial m} = \int_{A_J^{J^c}} [\frac{\partial}{\partial m} \int_{A_J^J(m)} f(\mathbf{x}^{J^c}, \mathbf{x}^J)d\mathbf{x}^J]d\mathbf{x}^{J^c} \tag{15}$$

Now we focus on the part in brackets, $\frac{\partial}{\partial m}\int_{A_J^J(m)} f(\mathbf{x}^{J^c}, \mathbf{x}^J)d\mathbf{x}^J$.

First notice that by (7) we know

$$A_J^J(m) = \{\mathbf{x}^J \in I^J | x_j \geq p_j, \forall j \in J; \sum_{j\in J} x_j < \sum_{j \in J} p_j + m\}$$

In the following expression we write out $\int_{A_J^J(m)} f(\mathbf{x}^{J^c}, \mathbf{x}^J)d\mathbf{x}^J$ in all $|J|$ dimensions, in ascending order of product indices from inside outwards.

$$\int_{A_J^J(m)} f(\mathbf{x}^{J^c}, \mathbf{x}^J)d\mathbf{x}^J$$

$$= \int_{p_{|J|}}^{m+p_{|J|}} \int_{p_{|J|-1}}^{m+p_{|J|-1}+p_{|J|}-x_{|J|}} ... \int_{p_k}^{m+\sum\limits_{j\geq k} p_j - \sum\limits_{j>k} x_j} ... \int_{p_1}^{m+\sum\limits_{j\geq 1} p_j - \sum\limits_{j>1} x_j} f(\mathbf{x})dx_1...dx_k...dx_{|J|-1}dx_{|J|}$$

Since $|J| > 1$, this expression will have at least two "layers". We focus on the first (outmost) layer, and denote all the parts inside the first layer of integration as

$$V(|J|-1, \mathbf{P}, m, x_{|J|}) \equiv \int_{p_{|J|-1}}^{m+p_{|J|-1}+p_{|J|}-x_{|J|}} ... \int_{p_k}^{m+\sum\limits_{j\geq k} p_j - \sum\limits_{j>k} x_j} ... \int_{p_1}^{m+\sum\limits_{j\geq 1} p_j - \sum\limits_{j>1} x_j} f(\mathbf{x})dx_1...dx_k...dx_{|J|-1} \tag{16}$$

where $|J|-1$ represents the number of layers of integration in V; $\mathbf{P}$ and m enter V because they define the limits of integration; and, finally, only $x_{|J|}$ instead of $\mathbf{x}^J$ enters V because in the $|J|-1$ layers of integration from inside out, $dx_1, ..., dx_{|J|-1}$ each acts as the variable of integration and therefore gets integrated away, leaving only $x_{|J|}$ (from the limits of integration) in V.

Therefore we can rewrite $\int_{A_J^J(m)} f(\mathbf{x}^{J^c}, \mathbf{x}^J)d\mathbf{x}^J$ as

$$\int_{A_J^J(m)} f(\mathbf{x}^{J^c}, \mathbf{x}^J)d\mathbf{x}^J = \int_{p_{|J|}}^{m+p_{|J|}} V(|J|-1, \mathbf{P}, m, x_{|J|})dx_{|J|}$$

Therefore we have

$$\begin{aligned}
&\frac{\partial}{\partial m}\int_{A^J_J(m)} f(\mathbf{x}^{J^c},\mathbf{x}^J)d\mathbf{x}^J \\
= \quad & V(|J|-1,\mathbf{P},m,x_{|J|}=m+p_{|J|})+\int_{p_{|J|}}^{m+p_{|J|}}\frac{\partial}{\partial m}V(|J|-1,\mathbf{P},m,x_{|J|})dx_{|J|}
\end{aligned} \tag{17}$$

Now examine the first part of (17), $V(|J|-1,\mathbf{P},m,x_{|J|}=m+p_{|J|})$, which is found by letting $x_{|J|}=m+p_{|J|}$ in (16). We only need to focus on the upper limit of integration in (16), which is $m+p_{|J|-1}+p_{|J|}-x_{|J|}$. We immediately see that it is equal to $p_{|J|-1}$ when $x_{|J|}=m+p_{|J|}$. But this means the upper and lower limits of integration of (16) are the same. Therefore

$$V(|J|-1,\mathbf{P},m,x_{|J|}=m+p_{|J|})=0$$

Now consider the second part of (17). When $m=0$, its upper and lower limits of integration also coincide, and therefore it also equals 0 when $m=0$.

Hence we have

$$\frac{\partial}{\partial m}\int_{A^J_J(m)} f(\mathbf{x}^{J^c},\mathbf{x}^J)d\mathbf{x}^J|_{m=0}=0$$

Substitute back to (15) and we are done. ■

Comment Part (i) of Theorem 2 says that, imposing a small extra membership fee on top of an additive price schedule will cause a first-order decrease in the demand for single-product bundles. Part (ii) says that such a price manipulation has no first-order impact on the demand for all multi-product bundles (consisting of two or more products).

Theorem 2 summarizes the demand implications of two-part tariffs, so we name it the two-part-tariff effect.

The two-part-tariff effect is crucial to the profitability of two-part tariffs. When the firm charges everyone an extra membership fee, part (ii) of Theorem 2 implies that this will lead to a pure first-order gain in profit from all multi-product consumers, as their demand does not decrease (on the first order) as a result. This gives rise to the possibility of higher overall profit for the firm. In section 5 we discuss when this gain will dominate the loss from the decreased demand for single products.

4.3 Deviation Starting from Non-Additive Price Schedules

The two-part-tariff effect exploits the multi-dimensional nature of the demand of a multiproduct firm. This argument is made in comparison to separate pricing (i.e. additive) strategies.

Ignoring implementability concerns, would the two-part-tariff effect in Theorem 2 still hold had the deviation started from a non-additive price schedule?

The answer is *not necessarily.*

Actually, starting from a non-additive price schedule, the first-order impact on demand for multiproduct bundles due to an extra membership fee may be strictly negative. This is because the original non-additive pricing strategy may already be exploiting the multiple dimensionality, since it is not additive.

For instance, in the case when $n = 2$ with uniform distribution[5], consider a two-part tariff $(m_1, \mathbf{P})$ with $m_1 > 0$ as the original pricing strategy. Suppose we impose a second membership fee applicable to all consumers on top of m_1. It is straightforward to see that this deviation is equivalent to increasing the first membership fee m_1 of the original strategy. Since $M(A_{\{1,2\}}) = (1 - p_1) \cdot (1 - p_2) - \frac{1}{2}m^2$ under uniform distribution, we have $\frac{\partial M(A_{\{1,2\}})}{\partial m}|_{m=m_1} = -m_1 < 0$. That is, imposing an additional membership fee will strictly reduce the demand for multiproduct bundles. Therefore there is no two-part-tariff effect starting from any two-part tariff $(m_1, \mathbf{P})$ with $m_1 > 0$.

5 Two-Part Tariffs vs. Separate Pricing

5.1 The Profit Maximization Problem

If the firm uses general price schedule $\mathbf{P} = \{p_J\}_{J \subset N}$ to maximize its profit, its maximization problem is

$$\max_{\{p_J\}_{J \subset N}} \sum_{J \subset N} p_J M(A_J)$$

where $\{A_J\}_{J \subset N}$ is the allocation induced by $\mathbf{P}$.

As discussed previously, so far the literature has not succeeded in finding the general solution of this problem. The difficulties in solving this problem are discussed in section 2.

Given the purpose of this paper, we focus on a comparison between two-part tariffs and separate pricing.

5.2 Separate Pricing

In Definition 3, we have defined additivity to be a synonym of separate pricing to reflect the fact that a separate pricing strategy does not involve any manipulation of the prices of different combinations of products.

Note that, due to additivity, a separate pricing strategy $\mathbf{P} = \{p_J\}_{J \subset N}$ can also be written as $\mathbf{P} = \{p_j\}_{j \in N}$ which only lists the prices of single products, as they *uniquely and completely* determine all the other prices in schedule $\mathbf{P}$.

If the firm only uses separate pricing strategies, its profit maximization problem is simple and solvable. Actually, because additive price schedules allocate all consumers into "cubes" delineated by orthogonal hyperplanes (see (1)), the firm's maximization problem

[5] See section 8.4.1 for a detailed discussion of this special case.

reduces to n separate maximization problems, each regarding one single product. Therefore in the optimal separate pricing strategy the monopolist can set the optimal price for each single product *irrespective* of all the other products and their prices (i.e. acting as a single-product monopolist). We call the separate pricing strategy comprised of these "individually optimal" prices the monopoly separate pricing strategy.

Definition 12 (Monopoly Separate Pricing/MSP) *A price schedule* $\mathbf{P} = \{p_J\}_{J \subset N}$ *is a monopoly separate pricing strategy if it is additive and* $\forall j \in N$,

$$p_j = \arg\max_{p'_j \geq 0} p'_j \cdot \Pr[x_j \geq p'_j] \tag{18}$$

Lemma 8 *The optimal separate pricing strategy is an MSP. There exists MSP* $\mathbf{P} = \{p_J\}_{J \subset N}$ *such that* $p_j \in (0,1)$ $\forall j \in N$ *and* $\mathbf{P}$ *yields strictly positive profit.*

Proof. With separate pricing strategy $\mathbf{P}$, the allocation is given by (1).

We consider the set of all consumers who buy some bundle that contains product j, i.e. the total demand for product $j \in N$:

$$\mathcal{B}_j \equiv \bigcup\nolimits_{J \ni j} A_J$$

Since $\{A_J\}_{J \subset N}$ satisfies (1), we have

$$\mathcal{B}_j = \{\mathbf{x} \in I^n | x_j \geq p_j\}, \forall j \in N$$

That is, under separate pricing, all consumers with valuation for product j higher than p_j will buy j.

Note that instead of writing the firm's total profit according to bundles, $\sum\limits_{J \subset N} p_J M(A_J)$, there is an alternative way of writing it according to products. When the firm uses an additive pricing strategy, the profit margin (equal to the price, as production cost is 0) is the same for the same product however it is sold (either separately or as part of a bundle). Therefore, we have

$$\sum_{J \subset N} p_J M(A_J) = \sum_{j \in N} p_j M(\mathcal{B}_j) = \sum_{j \in N} p_j \Pr[x_j \geq p_j]$$

Thus the maximization problem reduces to

$$\max_{\{p_j\}_{j \in N}} \sum_{j \in N} p_j \Pr[x_j \geq p_j]$$

and therefore the optimal strategy must be an MSP.

The existence of an MSP is implied by Assumption 2, under which the maximization problem of each individual product, $\max_{p'_j \geq 0} p'_j \cdot \Pr[x_j \geq p'_j]$, is well-behaved. ■

5.3 Comparison

Theorem 3 *When all x_j's ($j \in N$) are mutually **independent**, the optimal separate pricing strategy is strictly dominated by a two-part tariff.*

Proof. Suppose $\mathbf{P}$ is the optimal separate pricing strategy, then Lemma 8 tells us $\mathbf{P}$ is an MSP and satisfies (18).

Now we use this $\mathbf{P}$ to define a two-part tariff $\mathbf{Q}$ according to (3). Our strategy in this proof is to find first the difference in profits from $\mathbf{P}$ and $\mathbf{Q}$, and then show that when $m \to 0$, $\mathbf{Q}$ yields strictly higher profit than $\mathbf{P}$.

First define profit functions:

$$\begin{aligned}\pi(\mathbf{P}) &\equiv \sum_{J \subset N} p_J \cdot M(A_J) \\ \pi(\mathbf{Q}) &\equiv \sum_{J \subset N} q_J \cdot M(C_J)\end{aligned}$$

Notice by (3) and (5), we have

$$\begin{aligned}\Delta\pi &\equiv \pi(\mathbf{Q}) - \pi(\mathbf{P}) = \sum_{J \subset N} q_J \cdot M(C_J) - \sum_{J \subset N} p_J \cdot M(A_J) \\ &= \sum_{J \neq \varnothing} \{(p_J + m) \cdot [M(A_J) - M(A_J(m))] - p_J \cdot M(A_J)\} \\ &= m \cdot \sum_{J \neq \varnothing} M(A_J) - \sum_{J \neq \varnothing} (p_J + m) \cdot M(A_J(m))\end{aligned}$$

By definition, $M(A_J)$ does not depend on m, $\forall J \subset N$. Therefore

$$\begin{aligned}\frac{\partial \Delta\pi}{\partial m} &= \sum_{J \neq \varnothing} M(A_J) - \sum_{J \neq \varnothing} [M(A_J(m)) + (p_J + m) \cdot \frac{\partial M(A_J(m))}{\partial m}] \\ &= \sum_{J \neq \varnothing} M(A_J) - \sum_{J \neq \varnothing} M(A_J(m)) - \sum_{J \neq \varnothing} (p_J + m) \cdot \frac{\partial M(A_J(m))}{\partial m}\end{aligned}$$

Thus

$$\begin{aligned}
\frac{\partial \Delta\pi}{\partial m}|_{m=0} &= \sum_{J\neq\varnothing} M(A_J) - \sum_{J\neq\varnothing}(p_J)\cdot\frac{\partial M(A_J(m))}{\partial m}|_{m=0} \\
&= \sum_{j\in N} M(A_j) - \sum_{j\in N}(p_j)\cdot\frac{\partial M(A_j(m))}{\partial m}|_{m=0} \\
&\quad + \sum_{J\subset N,|J|>1} M(A_J) - \sum_{J\subset N,|J|>1}(p_J)\cdot\frac{\partial M(A_J(m))}{\partial m}|_{m=0}
\end{aligned}$$

where in the last equality we have used separate expressions for single-product bundles and multi-product bundles.

By Theorem 2 we know that for any $J \subset N, |J| > 1$, $\frac{\partial M(A_J(m))}{\partial m}|_{m=0} = -\frac{\partial M(A_J\backslash C_J)}{\partial m}|_{m=0} = 0$. Therefore we are left with

$$\frac{\partial \Delta\pi}{\partial m}|_{m=0} = \sum_{j\in N} M(A_j) - \sum_{j\in N}(p_j)\cdot\frac{\partial M(A_j(m))}{\partial m}|_{m=0} + \sum_{J\subset N,|J|>1} M(A_J) \qquad (19)$$

Now we study $\frac{\partial M(A_j(m))}{\partial m}|_{m=0}$.

We first focus on $M(A_j(m))$. By (10) we know

$$\begin{aligned}
A_j(m) &= \{\mathbf{x}\in I^n | p_j \le x_j < p_j + m; x_k < p_k, \forall k \in j^c\} \\
&= \times_{k\in j^c}\{x_k \in I^k | 0 \le x_k < p_k\} \times \{x_j \in I^j | p_j \le x_j < p_j\}
\end{aligned}$$

where by (9) we know that $A_j^{j^c}(m) = A_j^{j^c} = \times_{k\in j^c}\{x_k \in I^k | 0 \le x_k < p_k\}$

Therefore

$$M(A_j(m)) = \int_{p_j}^{p_j+m}[\int_{A_j^{j^c}} f(\mathbf{x}^{j^c}, x_j) d\mathbf{x}^{j^c}] dx_j$$

The integral in the brackets only depends on x_j and $\mathbf{P}^{j^c}$, which we define as

$$W_j(x_j, \mathbf{P}^{j^c}) \equiv \int_{A_j^{j^c}} f(\mathbf{x}^{j^c}, x_j) d\mathbf{x}^{j^c}$$

And rewrite

$$M(A_j(m)) = \int_{p_j}^{p_j+m} W_j(x_j, \mathbf{P}^{j^c}) dx_j$$

Therefore

$$\frac{\partial M(A_j(m))}{\partial m}|_{m=0} = W_j(p_j + m, \mathbf{P}^{j^c})|_{m=0} = W_j(p_j, \mathbf{P}^{j^c})$$

Substituting in (19) we get

$$\frac{\partial \Delta\pi}{\partial m}|_{m=0} = \sum_{j=1}^{n}[M(A_j) - p_j \cdot W_j(p_j, \mathbf{P}^{j^c})] + \sum_{J\subset N, |J|>1} M(A_J) \tag{20}$$

Notice that our proof up until this point applies to all general price schedules $\mathbf{P}$ satisfying (3) and any general distribution f satisfying Assumption 2.

Now we use the fact that $\mathbf{P}$ is MSP and all x_j's ($j \in N$) are mutually independent to show that $M(A_j) = p_j \cdot W_j(p_j, \mathbf{P}^{j^c})$, $\forall j \in N$.

First denote by F_j and f_j the marginal distribution and density of x_j, respectively.

Since $\mathbf{P}$ is additive, by (1) of Lemma 3, we have $\forall j \in N$

$$A_j = \{\mathbf{x} \in I^n | x_j \geq p_j; x_k < p_k, \forall k \in j^c\}$$

and therefore

$$M(A_j) = \int_{p_j}^{1}[\int_{A_j^{j^c}} f(\mathbf{x}^{j^c}, x_j) d\mathbf{x}^{j^c}] dx_j$$

Since all x_j's ($j \in N$) are mutually independent, we know $f(\mathbf{x}) = \prod_{j\in N} f_j(x_j)$. And since $A_j^{j^c} = \times_{k\in j^c}\{x_k \in I^k | 0 \leq x_k < p_k\}$ (see part (i) proof of Theorem 2), we have

$$M(A_j) = \prod_{k\neq j} F_k(p_k) \cdot [1 - F_j(p_j)]$$

and

$$W_j(p_j, \mathbf{P}^{j^c}) = \int_{A_j^{j^c}} f(\mathbf{x}^{j^c}, p_j) d\mathbf{x}^{j^c} = \prod_{k\neq j} F_k(p_k) \cdot f_j(p_j)$$

Therefore

$$M(A_j) - p_j \cdot W_j(p_j, \mathbf{P}^{j^c}) = \prod_{k\neq j} F_k(p_k) \cdot [1 - F_j(p_j) - p_j \cdot f_j(p_j)]$$

In the last step, we use the fact that $\mathbf{P}$ is MSP, which must satisfy (18), whose first order condition is, for all $j \in N$,

$$1 - F_j(p_j) = p_j f_j(p_j)$$

and therefore $M(A_j) - p_j \cdot W_j(p_j, \mathbf{P}^{j^c}) = 0$, $\forall j \in N$.

Therefore we have

$$\frac{\partial \Delta\pi}{\partial m}|_{m=0} = \sum_{J\subset N, |J|>1} M(A_J) \geq 0$$

Now we only need to show that the last inequality above must be strict when $\mathbf{P}$ is MSP.

To see this, suppose instead $\sum_{J,|J|>1} M(A_J) = 0$. Since each $M(A_J)$ is non-negative, we must have $M(A_J) = 0$, for all $J \subset N, |J| > 1$.

Given that the support in each dimension is $I = [0, 1]$, $M(A_J) = 0$ for all $|J| > 1$ implies that, in price schedule $\mathbf{P}$, there must exists at least one product k such that $p_k \geq 1$. The reason is that if there were no such product, i.e. $p_j < 1$ for all $j \in N$, then at least $M(A_N)$ would be strictly positive.

But $p_k \geq 1$ implies that $\Pr[x_k \geq p_k] = 0$, which in turn implies p_k yields zero profit from product k and therefore contradicts the fact that p_k is the optimal MSP price for product k.

Thus we have proven

$$\frac{\partial \Delta\pi}{\partial m}|_{m=0} > 0$$

which implies that $\mathbf{Q}$ yields strictly higher profit than $\mathbf{P}$. ■

Interpretation From Theorem 3 we know that, with independence, no separate pricing strategy is optimal. One simple way to increase profit from the optimal separate pricing strategy $\mathbf{P}$ is to impose a small membership fee (that is, raising all the prices in $\mathbf{P}$ by the same small amount $m > 0$ except for the price of the empty bundle).

Theorem 4 *With* ***any general*** f *satisfying Assumption 2,* ***any*** *separate pricing strategy* $\mathbf{P}$ *is strictly dominated by the two-part tariff* $\mathbf{Q}$ *defined in (3) using* $\mathbf{P}$ *if the following condition holds at* $\mathbf{P}$

$$\sum_{j=1}^{n}[M(A_j) - p_j \cdot \int_{A_j^{j^c}} f(\mathbf{x}^{j^c}, p_j) d\mathbf{x}^{j^c}] + \sum_{J \subset N, |J|>1} M(A_J) > 0 \tag{21}$$

where $\{A_J\}_{J \subset N}$ *is the allocation induced by* $\mathbf{P}$ *as defined in (1).*

Proof. Consider the price schedules defined in (3). We need to show that when condition (21) holds, $\mathbf{Q}$ yields strictly higher profit than $\mathbf{P}$.

In exactly the same way as in the proof of Theorem 3 we can get result (20), which holds for all general price schedules that satisfies (3) and any general distribution f satisfying Assumption 2.

Substitute $W_j(p_j, \mathbf{P}^{j^c}) = \int_{A_j^{j^c}} f(\mathbf{x}^{j^c}, p_j) d\mathbf{x}^{j^c}$ in (20) and we see that condition (21) of Theorem 4 is exactly $\frac{\partial \Delta\pi}{\partial m}|_{m=0} > 0$. Therefore $\mathbf{P}$ is strictly dominated by $\mathbf{Q}$ when condition (21) holds. ■

Intuition As we have discussed in the comment of Theorem 2, imposing a small extra membership fee on top of a separate pricing strategy will only decrease the demand for single-product bundles, but has no first-order impact on the demand for all multi-product bundles. Since the firm charges everyone an extra membership fee, it gains from each and every one of multi-product consumers. This gain is represented by the term $\sum_{J\subset N,|J|>1} M(A_J)$ in condition (21) (which is exactly the "number" of all multi-product consumers). From single-product consumers, the firm charges a higher price, but also loses some demand. The net effect from single-products is represented by the term $\sum_{j=1}^{n}[M(A_j) - p_j \cdot \int_{A_j^{j^c}} f(\mathbf{x}^{j^c}, p_j)d\mathbf{x}^{j^c}]$, which may be positive or negative, depending on the pricing strategy $\mathbf{P}$. The overall impact on profit from the whole market is therefore captured by the left-hand side of condition (21).

Comment Theorem 4 generalizes Proposition 1 of McAfee, McMillan and Whinston (1989) to the multiproduct case. The latter addresses the case of two products and provides a condition for *mixed bundling* to strictly dominate separate pricing. It can be shown that when $n = 2$, our condition (21) reduces to their condition (1).

We postpone the discussion of the relationship between our results on two-part tariffs and MMW's on mixed bundling to section 8.

6 Generalized Two-Part Tariffs

Thus far our discussion has focused on the situations where condition (21) holds. Now suppose it does not hold at MSP $\mathbf{P}$. Can the firm make a profitable deviation in this case?

The answer is *yes*. Indeed, since the left-hand side of condition (21) is merely $\frac{\partial \Delta\pi}{\partial m}|_{m=0}$, in the case that it is negative, we expect that using a negative m should be profitable. In this section we generalize two-part tariffs to allow for negative membership fees.

6.1 Subsidizing Memberships

A negative m is actually a **membership subsidy**. In real life, membership subsidies often take the form of free gifts offered to consumers, such as complimentary appetizers or desserts at restaurants, and free "air-time" offered by mobile phone networks. We have not used membership subsidy in our general model because it may not always be feasible, as we have restricted the feasible price schedules to be non-negative. However, when we consider condition (21) at the optimal separate pricing strategy, Lemma 8 tells us that the MSP is strictly positive, and therefore a small membership subsidy is indeed feasible. Now we generalize our definition of two-part tariffs to incorporate subsidies.

Definition 13 (Generalized Two-Part Tariff) *A* ***generalized two-part tariff*** *is a* *price schedule* $\mathbf{Q} = \{q_J\}_{J \subset N}$ *consisting of two parts* $(m, \mathbf{P})$*, where*

$$\begin{aligned} & m \in \mathbb{R}; \textit{ and} \\ & \mathbf{P} = \{p_J\}_{J \subset N} \textit{ is additive; and} \\ & q_J = \begin{cases} p_J + m & , \textit{ if } J \neq \varnothing \\ 0 & , \textit{ if } J = \varnothing \end{cases} ; \textit{ and} \\ & q_J \geq 0 \textit{ for all } J \subset N. \end{aligned} \tag{22}$$

And the allocations induced by $\mathbf{P}$ *and* $\mathbf{Q}$ *are denoted* $\{A_J\}_{J \subset N}$ *and* $\{C_J\}_{J \subset N}$*, respectively. When* $m > 0$*, we call it a* ***membership fee****; when* $m < 0$*, we call its absolute value a* ***membership subsidy****.*

Once we allow the membership fee to be negative, we have the following result.

Theorem 5 *Suppose* $\mathbf{P}$ *is a monopoly separate pricing strategy, and the reverse of condition (21) holds with strict inequality at* $\mathbf{P}$*, then offering every consumer (who buys at least one product) a small* ***membership subsidy*** *(i.e. a negative membership fee) strictly increases profit.*

Intuition A membership subsidy leads to a first-order increase in the demand for single products, and does not affect the demand for any other bundles. Under the condition expressed in Theorem 5, this turns out to be profitable overall.

Proof. In order to distinguish membership subsidies from membership fees, we denote the former by $s \equiv -m$. In this proof, we require that $s \geq 0$. Therefore, by Definition 13, we have

$$q_J = \begin{cases} p_J - s & , \text{ if } J \neq \varnothing \\ 0 & , \text{ if } J = \varnothing \end{cases}$$
$$\text{where } s \geq 0.$$

Step 1: Finding the allocation induced by $\mathbf{Q}$.

Since $\mathbf{P}$ is additive, by Lemma 3 we know the allocation induced by $\mathbf{P}$ is defined in (1), that is:

$$A_J = \{\mathbf{x} \in I^n | x_j \geq p_j, \forall j \in J; x_k \leq p_k, \forall k \in J^c\}$$

Now we look at the allocation induced by $\mathbf{Q} = \{q_J\}_{J \subset N}$ defined above, $\{C_J\}_{J \subset N}$. By Definition 4, we know for any $J \subset N$

$$C_J = \{\mathbf{x} \in I^n | \sum_{j \in J} x_j - q_J \geq \sum_{k \in K} x_k - q_K, \forall K \subset N\}$$

Now we study J of different sizes.

i) When $|J| > 1$, for any such $J \subset N$, we have $q_J = p_J - s$, which implies

$$\begin{aligned}\sum_{j\in J} x_j - q_J \quad &\geq \quad \sum_{k\in K} x_k - q_K, \forall K \subset N \\ &\Leftrightarrow \begin{cases} \text{(a)} \ \sum_{j\in J} x_j - p_J \geq \sum_{k\in K} x_k - p_K & , \forall K \subset N, K \neq \varnothing \\ \text{(b)} \ \sum_{j\in J} x_j - p_J \geq -s & \text{(when } K = \varnothing) \end{cases}\end{aligned}$$

In the same way as we show that (i)⇔(ii) in the proof of Lemma 3, we know that for $|J| > 1$

$$\text{part (a) above} \Leftrightarrow \begin{cases} \text{(a1)} \ x_j \geq p_j & , \forall j \in J; \\ \text{(a2)} \ x_k \leq p_k & , \forall k \in J^c. \end{cases}$$

And it is straightforward to see that (a1) implies (b) above, as **P** is additive. (This means that when we use a membership subsidy $s > 0$, the IR condition of any multiproduct bundle is not binding.)

From (a1) and (a2) we know for any J such that $|J| > 1$, **Q** induces the *same* demand segment as **P**:

$$C_{J,|J|>1} = \{\mathbf{x} \in I^n | x_j \geq p_j, \forall j \in J; x_k \leq p_k, \forall k \in J^c\} = A_{J,|J|>1}$$

ii) When $|J| = 1$, i.e. $J = \{j\}$. For any $j \in N$, we have $q_j = p_j - s$, which implies

$$\begin{aligned}\sum_{j\in J} x_j - q_J \quad &\geq \quad \sum_{k\in K} x_k - q_K, \forall K \subset N \\ &\Leftrightarrow \begin{cases} \text{(c)} \ x_j - p_j \geq \sum_{k\in K} x_k - p_K & , \forall K \ni j, K \subset N; \\ \text{(d)} \ x_j - p_j \geq \sum_{k\in K} x_k - p_K & , \forall K \not\ni j, |K| > 1, K \subset N; \\ \text{(e)} \ x_j - p_j \geq x_k - p_k & , \forall k \in j^c; \\ \text{(f)} \ x_j - p_j \geq -s. & \text{(when } K = \varnothing) \end{cases}\end{aligned}$$

If we write out the right-hand side of (c) above for K of different sizes, we have

$$\text{part (c) above} \Leftrightarrow \text{(c1)} \ x_k \leq p_k, \forall k \in j^c.$$

Also, it is clear that (c1) and (e) together imply (d) as **P** is additive. Therefore we only need (c1), (e) and (f) to fully characterize the demand segment of any single product $j \in N$ that **Q** induces, which is

$$C_j = \{\mathbf{x} \in I^n | x_j - p_j \geq \max[-s, x_k - p_k], \text{ and } x_k \leq p_k, \forall k \in j^c\}$$

Note that this is different from the single-product segments induced by **P**.

iii) When $|J| = 0$, i.e. $J = \varnothing$, we have $q_\varnothing = p_\varnothing = 0$, which implies

$$
\begin{aligned}
\sum_{j\in J} x_j - q_J \quad &\geq \quad \sum_{k\in K} x_k - q_K, \forall K \subset N \\
&\Leftrightarrow \quad \begin{cases} \text{(g) } 0 \geq \sum_{k\in K} x_k - p_K + s & , \forall K \subset N, |K| > 1; \\ \text{(h) } 0 \geq x_k - p_k + s & , \forall k \in N. \end{cases}
\end{aligned}
$$

Clearly, (h) implies (g) as $s \geq 0$ and $\mathbf{P}$ is additive. Therefore we have

$$
C_\varnothing = \{\mathbf{x} \in I^n | x_k \leq p_k - s, \forall k \in N\}
$$

In summary, the allocation $\{C_J\}_{J\subset N}$ induced by $\mathbf{Q}$ with a membership subsidy s is for any $J \subset N$

$$
C_J = \begin{cases} \{\mathbf{x} \in I^n | x_j \geq p_j, \forall j \in J; x_k < p_k, \forall k \in J^c\} = A_J & , \text{ if } |J| > 1; \\ \{\mathbf{x} \in I^n | x_j - p_j \geq \max[-s, x_k - p_k], x_k \leq p_k, \forall k \in j^c\} & , \text{ if } J = \{j\}; \\ \{\mathbf{x} \in I^n | x_k \leq p_k - s, \forall k \in N\} & , \text{ if } J = \varnothing. \end{cases}
$$

Step 2: Finding the profit implication of raising s from 0, $\frac{\partial \Delta\pi}{\partial s}|_{s=0}$.

Compared to the allocation induced by $\mathbf{P}$ as shown in (1), the only difference in $\{C_J\}_{J\subset N}$ happens on the single-product bundles and the empty bundle. In particular, for any $j \in N$ we have

$$
\begin{aligned}
A_j &= \{\mathbf{x} \in I^n | x_j \geq p_j; x_k \leq p_k, \forall k \in j^c\} \\
C_j &= \{\mathbf{x} \in I^n | x_j - p_j \geq \max[-s, x_k - p_k], x_k \leq p_k, \forall k \in j^c\}
\end{aligned}
$$

Since $\max[-s, x_k - p_k] \leq 0$, it must be

$$
A_j \subset C_j, \forall j \in N
$$

Now define for any $j \in N$

$$
C_j(s) \equiv C_j \backslash A_j = \{\mathbf{x} \in I^n | \max[-s, x_k - p_k] \leq x_j - p_j \leq 0, x_k \leq p_k, \forall k \in j^c\} \tag{23}
$$

Then we have

$$
\begin{aligned}
M(C_j) &= M(A_j) + M(C_j(s)), \forall j \in N; \\
M(C_J) &= M(A_J), \forall J \subset N, |J| > 1.
\end{aligned}
$$

Note that $M(C_j(s = 0)) = 0$.

Therefore,

$$\begin{aligned}
\Delta\pi &\equiv \pi(\mathbf{Q})-\pi(\mathbf{P}) = \sum_{J\subset N} q_J \cdot M(C_J) - \sum_{J\subset N} p_J \cdot M(A_J) \\
&= \sum_{J\subset N,|J|>1} (q_J - p_J)\cdot M(A_J) + \sum_{j\in N}[q_j \cdot M(C_j) - p_j \cdot M(A_j)] \\
&= -s\cdot \sum_{J\subset N,|J|>1} M(A_J) + \sum_{j\in N}\{(p_j - s)\cdot[M(A_j)+M(C_j(s))] - p_j\cdot M(A_j)\} \\
&= -s\cdot \sum_{J\subset N,|J|>1} M(A_J) + \sum_{j\in N}\{p_j\cdot M(C_j(s)) - s\cdot[M(A_j)+M(C_j(s))]\}
\end{aligned}$$

By definition, $M(A_J)$ does not depend on s, $\forall J \subset N$. Therefore

$$\begin{aligned}
\frac{\partial\Delta\pi}{\partial s} &= -\sum_{J\subset N,|J|>1} M(A_J) + \sum_{j\in N}\{p_j\cdot\frac{\partial M(C_j(s))}{\partial s} - [M(A_j)+M(C_j(s))] - s\cdot\frac{\partial M(C_j(s))}{\partial s}\} \\
&= -\sum_{J\subset N,|J|>1} M(A_J) - \sum_{j\in N}[M(A_j)+M(C_j(s))] + \sum_{j\in N}[(p_j - s)\cdot\frac{\partial M(C_j(s))}{\partial s}]
\end{aligned}$$

Thus

$$\frac{\partial\Delta\pi}{\partial s}|_{s=0} = -\sum_{J\subset N} M(A_J) + \sum_{j\in N} p_j\cdot\frac{\partial M(C_j(s))}{\partial s}|_{s=0} \tag{24}$$

The only thing remains to be shown now is $\frac{\partial M(C_j(s))}{\partial s}|_{s=0}$.
By definition (in (23)), we have

$$C_j(s) = \{\mathbf{x}\in I^n|\max[-s, x_k - p_k] \le x_j - p_j \le 0, x_k \le p_k, \forall k\in j^c\}$$

Now we define the *section* of $C_j(s)$ on hyperplane $x_j = x'_j$ in dimensions j^c (that is orthogonal to dimension j), which is[6]

$$C_j^{j^c}(x'_j) \equiv \{\mathbf{x}^{j^c}\in I^{j^c}|0\le x_k < x'_j - p_j + p_k, \forall k\in j^c\} \tag{25}$$

Then we have

$$M(C_j(s)) = \int_{p_j - s}^{p_j}[\int_{C_j^{j^c}(x_j)} f(\mathbf{x}^{j^c}, x_j)d\mathbf{x}^{j^c}]dx_j$$

where if we relabel the products in $\{j\}^c$ as $j_1^c, j_2^c, ..., j_{n-1}^c$, we have

$$\int_{C_j^{j^c}(x_j)} f(\mathbf{x}^{j^c}, x_j)d\mathbf{x}^{j^c} = \int_0^{x_j - p_j + p_{j_1^c}}\int_0^{x_j - p_j + p_{j_2^c}}...\int_0^{x_j - p_j + p_{j_{n-1}^c}} f(\mathbf{x}^{j^c}, x_j)dx_{j_{n-1}^c}...dx_{j_2^c}dx_{j_1^c}$$

[6]Note the section $C_j^{j^c}(x'_j)$ defined here does not include its $(n-1)$ "edge-hyperplanes", i.e. it is missing the set $\{\mathbf{x}^{j^c}\in I^{j^c}|x_k = x'_j - p_j + p_k, \forall k\in j^c\}$. Given that f is atomless and $\mathbf{S}$ has full dimension in $\mathbb{R}^n$, this set has zero measure in $\mathbb{R}^n$. Therefore our expression of $M(C_j(s))$ using the current definition of $C_j^{j^c}(x'_j)$ is correct.

which is fully defined by x_j and $\mathbf{P}$, and does not depend on s.

Therefore, we have

$$\begin{aligned}\frac{\partial M(C_j(s))}{\partial s}|_{s=0} &= \int_{C_j^{j^c}(p_j-s)} f(\mathbf{x}^{j^c}, p_j - s) d\mathbf{x}^{j^c}|_{s=0} \\ &= \int_{C_j^{j^c}(p_j)} f(\mathbf{x}^{j^c}, p_j) d\mathbf{x}^{j^c}\end{aligned}$$

Now by definition (in (25)), we know

$$C_j^{j^c}(p_j) = \{\mathbf{x}^{j^c} \in I^{j^c} | 0 \leq x_k < p_k, \forall k \in j^c\}$$

By (8) and Lemma 6, we see that the $C_j^{j^c}(p_j)$ here is exactly the same as $A_j^{j^c}$ (the projection of the demand segment of j induced by additive $\mathbf{P}$, A_j, on a hyperplane in dimensions j^c). That is

$$C_j^{j^c}(p_j) = A_j^{j^c}$$

Therefore we have

$$\frac{\partial M(C_j(s))}{\partial s}|_{s=0} = \int_{A_j^{j^c}} f(\mathbf{x}^{j^c}, p_j) d\mathbf{x}^{j^c}$$

Substituting back in (24), we have

$$\frac{\partial \Delta\pi}{\partial s}|_{s=0} = -\sum_{J \subset N} M(A_J) + \sum_{j \in N} p_j \cdot \int_{A_j^{j^c}} f(\mathbf{x}^{j^c}, p_j) d\mathbf{x}^{j^c}$$

Therefore, $\frac{\partial \Delta\pi}{\partial s}|_{s=0} > 0$ if and only if

$$\sum_{J \subset N} M(A_J) - \sum_{j \in N} p_j \cdot \int_{A_j^{j^c}} f(\mathbf{x}^{j^c}, p_j) d\mathbf{x}^{j^c} < 0$$

which is exactly the reverse of condition (21) in Theorem 4. ■

Comment In the proof above, we have shown that although the membership subsidy s affects different demand segments (i.e. only single-product bundles) than the membership fee (which affects all bundles), its marginal impact on profits starting from an additive price schedule is actually exactly opposite to that of the membership fee. Therefore when the reverse of condition (21) holds at the optimal separate pricing strategy, a membership subsidy strictly increases profit.

Note that the two-part-tariff effect of Theorem 2 is no longer present when membership subsidies are used. The demand changes induced by a subsidy are all first-order impacts as $\frac{\partial M(C_j(s))}{\partial s}|_{s=0} = \int_{A_j^{j^c}} f(\mathbf{x}^{j^c}, p_j) d\mathbf{x}^{j^c} > 0$.

6.2 The "Power" of Generalized Two-Part Tariffs

When we allow membership fees to be either positive or negative, a (generalized) two-part tariff becomes much more powerful. The following result summarizes our discussion of both cases.

Theorem 6 (Dominance by Generalized Two-Part Tariffs) *Under **any general** f satisfying Assumption 2, suppose* $\mathbf{P}$ *is a monopoly separate pricing strategy. Then generalized two-part tariffs strictly dominate separate pricing if the following condition holds at* $\mathbf{P}$

$$\sum_{J\subset N} M(A_J) - \sum_{j\in N} p_j \cdot \int_{A_j^{j^c}} f(\mathbf{x}^{j^c}, p_j) d\mathbf{x}^{j^c} \neq 0 \tag{26}$$

where $\{A_J\}_{J\subset N}$ *is the allocation induced by* $\mathbf{P}$ *as defined in (1).*

Proof. This is implied by Theorems 4 and 5. In particular, when condition (26) holds, the optimal separate pricing strategy $\mathbf{P}$ is strictly dominated by the generalized two-part tariff $\mathbf{Q}$ defined in Definition 13 using $\mathbf{P}$. ■

7 Variable Membership Fees

7.1 Two Membership Fees

In real life, multiproduct firms sometimes use more than one level of membership fees. For instance, if we think of parking charges at a shopping mall as a form of membership fee charged by the mall, in practice such charges may vary according to how much shoppers purchase from the mall. A seemingly popular practice is that a shopper gets a *discount* on parking fee once her expenditure in the mall exceeds some threshold. This in effect creates *two* different levels of membership fees: Consumers with purchases lower than the threshold are charged a higher membership fee, say m, whereas consumers with purchases no lower than the threshold pay a lower membership fee, say $(m - \epsilon)$, where $\epsilon \geq 0$.

Now we use our model to study this situation. Suppose a shopping mall sells $n \geq 2$ products, and it is located in a remote area such that all consumers have to drive to shop there. Suppose the mall initially uses a two-part tariff $\mathbf{Q} = (m, \mathbf{P})$, where m is the charge for parking that is common to all consumers, and thus acts as a membership fee; and $\mathbf{P} = \{p_J\}_{J\subset N}$ is a separate pricing strategy. That is, $\mathbf{Q} = (m, \mathbf{P}) = \{q_J\}_{J\subset N}$ is exactly as in Definition 10.

Now suppose the mall wants to introduce a discount of $\epsilon \geq 0$ on parking for consumers who purchase at least $\underline{n}$ products from the mall. We denote by $\mathbf{R}$ the new price schedule after the introduction of discount ϵ in $\mathbf{Q}$. Our goal here is to see whether using such a discount is profitable or not.

Definition 14 (Two-Part Tariff with Two Membership Fees) *A two-part tariff with two membership fees is a price schedule* $\mathbf{R} = \{r_J\}_{J\subset N}$ *consisting of three parts* $(m, \mathbf{P}, \epsilon)$, *where*

$$(m, \mathbf{P}) = \mathbf{Q} = \{q_J\}_{J\subset N} \text{ is a two-part tariff; and}$$
$$\epsilon \in (0, m]\text{; and}$$
$$r_J = \begin{cases} q_J - \epsilon = p_J + (m - \epsilon) & , \text{ if } |J| \geq \underline{n} \\ q_J = p_J + m & , \text{ if } 0 < |J| < \underline{n} \\ 0 & , \text{ if } J = \varnothing \end{cases} \tag{27}$$
$$\text{where } \underline{n} \geq 2.$$

And the allocations induced by $\mathbf{P}$, $\mathbf{Q}$, *and* $\mathbf{R}$ *are denoted* $\{A_J\}_{J\subset N}$, $\{C_J\}_{J\subset N}$, *and* $\{D_J\}_{J\subset N}$ *respectively.*

If $\underline{n} = 1$, it is clear that $\mathbf{R} = (m - \epsilon, \mathbf{P})$, i.e. $\mathbf{R}$ reduce to a two-part tariff with membership fee $(m - \epsilon)$. In this case, the discussion of $\mathbf{R}$ is exactly the same as that of a normal two-part tariff $\mathbf{Q}$ in Definition 10, as the discount ϵ in is case has no different role than the original membership fee m.

Therefore, in this section we focus on the case when the threshold $\underline{n} \geq 2$.

7.1.1 Two Membership Fees vs. One

Theorem 7 (Two Membership Fees) *Under* ***any general*** f *satisfying Assumption 2, suppose* $\mathbf{Q} = (m, \mathbf{P}) = \{q_J\}_{J\subset N}$ *is the* ***optimal*** *two-part tariff with* $m > 0$. *Then offering all consumers who buy at least* $\underline{n}(\geq 2)$ *products a small discount* ϵ *on their membership fee* m *strictly increases profit over* $\mathbf{Q}$ *if the following condition holds at* $\mathbf{Q}$

$$\sum_{J\subset N, |J|\geq \underline{n}} q_J \cdot \frac{\partial M(D_J \backslash C_J)}{\partial \epsilon}|_{\epsilon=0} - \sum_{J\subset N, |J|\geq \underline{n}} M(C_J) - \sum_{J\subset N, |J|=\underline{n}-1} q_J \cdot \frac{\partial M(C_J \backslash D_J)}{\partial \epsilon}|_{\epsilon=0} > 0 \tag{28}$$

where $\{C_J\}_{J\subset N}$ *and* $\{D_J\}_{J\subset N}$ *are defined in (27).*

Intuition Offering all consumers who buy at least $\underline{n}$ products a small discount ϵ results in *decreases* in demand for all bundles of sizes smaller than $\underline{n}$, and *increases* in demand for all bundles of sizes at least equal to $\underline{n}$. However, on the *first order*, the additional discount ϵ only results in demand changes for bundles of sizes equal to or larger than $\underline{n} - 1$, and the demand changes for all other bundles are of orders higher than 1. The condition (28) in Theorem 7 controls these first-order demand changes at the optimal two-part tariff.

Proof. We denote by $\mathbf{R}$ the price schedule resulting from offering all consumers who buy at least $\underline{n}(\geq 2)$ products a small discount ϵ in $\mathbf{Q}$. Our plan in this proof is to first identify the allocation induced by $\mathbf{R}$, and then study the profit implication of raising ϵ from 0.

Step 1: Finding the allocation induced by $\mathbf{R} = \{r_J\}_{J\subset N}$. It is clear that this $\mathbf{R}$ is exactly the price schedule defined in (27) using $\mathbf{Q} = (m, \mathbf{P}) = \{q_J\}_{J\subset N}$. Therefore

$$r_J = \begin{cases} q_J - \epsilon = p_J + (m - \epsilon) & \text{, if } |J| \geq \underline{n} \\ q_J = p_J + m & \text{, if } 0 < |J| < \underline{n} \\ 0 & \text{, if } J = \varnothing \end{cases}$$

where $\mathbf{P} = \{p_J\}_{J\subset N}$ is additive.

By Definition 4 and the two-part tariff allocation in (4), we know the allocation induced by $\mathbf{Q}$ is

$$C_{J(\neq\varnothing)} = \{\mathbf{x} \in I^n | x_j \geq p_j, \forall j \in J; x_k \leq p_k, \forall k \in J^c; \sum_{j\in J} x_j \geq p_J + m\}$$

By Definition 4 and the definition of $\mathbf{R} = \{r_J\}_{J\subset N}$ above, we know the allocation induced by $\mathbf{R}$ is

$$\begin{aligned} D_{J\subset N, 0<|J|<\underline{n}-1} = \quad & \{\mathbf{x} \in I^n | x_j \geq p_j, \forall j \in J; x_k \leq p_k, \forall k \in J^c; \textstyle\sum_{j\in J} x_j \geq p_J + m; \\ & \textit{and} \textstyle\sum_{k\in K} x_k \geq p_K - \epsilon, \forall K \subset N, |K| \geq \underline{n} - |J|, K \cap J = \varnothing\} \\ D_{J\subset N, |J|=\underline{n}-1} = \quad & \{\mathbf{x} \in I^n | x_j \geq p_j, \forall j \in J; x_k \leq p_k - \epsilon, \forall k \in J^c; \textstyle\sum_{j\in J} x_j \geq p_J + m\} \\ D_{J\subset N, |J|=\underline{n}} = \quad & \{\mathbf{x} \in I^n | x_j \geq p_j - \epsilon, \forall j \in J; x_k \leq p_k, \forall k \in J^c; \textstyle\sum_{j\in J} x_j \geq p_J + m - \epsilon; \\ & \textit{and} \textstyle\sum_{h\in H} x_h - p_H \geq \sum_{k\in K} x_k \geq p_K, \forall H \subset J, K \subset N, K \cap J = \varnothing\} \\ D_{J\subset N, |J|>\underline{n}} = \quad & \{\mathbf{x} \in I^n | x_j \geq p_j, \forall j \in J; x_k \leq p_k, \forall k \in J^c; \textstyle\sum_{j\in J} x_j \geq p_J + m - \epsilon\} \end{aligned}$$

Compared to $\{C_J\}_{J\subset N}$ above, we have

$$\begin{aligned} D_{J\subset N, 0<|J|<\underline{n}-1} &\subset C_{J\subset N, 0<|J|<\underline{n}-1} \\ D_{J\subset N, |J|=\underline{n}-1} &\subset C_{J\subset N, |J|=\underline{n}-1} \\ D_{J\subset N, |J|=\underline{n}} &\supset C_{J\subset N, |J|=\underline{n}} \\ D_{J\subset N, |J|>\underline{n}} &\supset C_{J\subset N, |J|>\underline{n}} \end{aligned}$$

Therefore, offering all consumers who buy at least $\underline{n}(\geq 2)$ products a small discount ϵ results in decreases in demand for all bundles of sizes smaller than $\underline{n}$ and increases in demand for all bundles of sizes at least equal to $\underline{n}$.

Therefore, we have

$$M(D_J) - M(C_J) = \begin{cases} -M(C_J \backslash D_J) & \text{, if } 0 < |J| \leq \underline{n} - 1 \\ M(D_J \backslash C_J) & \text{, if } |J| \geq \underline{n} \end{cases}$$

Now we take a closer look at $\{D_J\}_{J\subset N}$ above, focusing on where ϵ causes a change from $\{C_J\}_{J\subset N}$. We want to find the lowest "polynomial order" of $M(D_J)$ in terms of ϵ.

As an intuitive illustration, we *suppose for now only* that consumers' types follow the *uniform distribution*. In this case, we can see from $\{D_J\}_{J \subset N}$ above that:

i) For $0 < |J| < \underline{n} - 1$, $M(D_J) \propto \epsilon^{|K|}(+\epsilon^{|H|(>|K|)})$, where $|K| \geq \underline{n} - |J| \geq 2$, (from condition $\sum_{k \in K} x_k \geq p_K - \epsilon, \forall K \subset N, |K| \geq \underline{n} - |J|, K \cap J = \varnothing$);

ii) For $|J| = \underline{n} - 1$, $M(D_J) \propto \epsilon$, (from condition $x_k \leq p_k - \epsilon, \forall k \in J^c$);

iii) For $|J| = \underline{n}$, $M(D_J) \propto \epsilon$, (from conditions $x_j \geq p_j - \epsilon, \forall j \in J$ and $\sum_{j \in J} x_j \geq p_J + m - \epsilon$);

iv) For $|J| > \underline{n}$, $M(D_J) \propto \epsilon$, (from condition $\sum_{j \in J} x_j \geq p_J + m - \epsilon$).

from which we see that the lowest polynomial orders of $M(D_J)$ in terms of ϵ are *all equal to* 1 *except only for* $0 < |J| < \underline{n} - 1$.

For general distribution f, we actually have

$$\frac{\partial M(D_J)}{\partial \epsilon}|_{\epsilon=0} = \begin{cases} -\frac{\partial M(C_J \backslash D_J)}{\partial \epsilon}|_{\epsilon=0} < 0 & \text{, if } |J| = \underline{n} - 1 \\ \frac{\partial M(D_J \backslash C_J)}{\partial \epsilon}|_{\epsilon=0} > 0 & \text{, if } |J| \geq \underline{n} \\ 0 & \text{, otherwise.} \end{cases} \tag{29}$$

That is, the additional discount ϵ only results in first-order demand changes for bundles of sizes equal to or larger than $\underline{n} - 1$.

Step 2: Finding the profit implication of raising ϵ from 0, $\frac{\partial \pi(\mathbf{R})}{\partial \epsilon}|_{\epsilon=0}$.

$$\begin{aligned} \pi(\mathbf{R}) &= \sum_{J \subset N} r_J \cdot M(D_J) \\ &= \sum_{J \subset N, |J| \geq \underline{n}} (q_J - \epsilon) \cdot M(D_J) + \sum_{J \subset N, |J| < \underline{n}} q_J \cdot M(D_J) \end{aligned}$$

Therefore

$$\begin{aligned} \frac{\partial \pi(\mathbf{R})}{\partial \epsilon}|_{\epsilon=0} &= \sum_{J \subset N, |J| \geq \underline{n}} q_J \cdot \frac{\partial M(D_J)}{\partial \epsilon}|_{\epsilon=0} - \sum_{J \subset N, |J| \geq \underline{n}} M(D_J)|_{\epsilon=0} + \sum_{J \subset N, |J| < \underline{n}} q_J \cdot \frac{\partial M(D_J)}{\partial \epsilon}|_{\epsilon=0} \\ &= \sum_{J \subset N, |J| \geq \underline{n}} q_J \cdot \frac{\partial M(D_J \backslash C_J)}{\partial \epsilon}|_{\epsilon=0} - \sum_{J \subset N, |J| \geq \underline{n}} M(C_J) - \sum_{J \subset N, |J| = \underline{n}-1} q_J \cdot \frac{\partial M(C_J \backslash D_J)}{\partial \epsilon}|_{\epsilon=0} \end{aligned}$$

where the last equation is due to (29) and that $D_J|_{\epsilon=0} = C_J, \forall J \subset N$.

Therefore $\frac{\partial \pi(\mathbf{R})}{\partial \epsilon}|_{\epsilon=0} > 0$ if condition (28) holds in Theorem 7. ■

7.1.2 An Example with Three Products and Uniform Distribution

Corollary 1 *Condition (28) holds under uniform distribution with three products when* $\underline{n} = 2$.

Proof. Under uniform distribution, when $n = 3$, the optimal two-part tariff is $\mathbf{Q} = (m, \mathbf{P})$, where $m \approx 0.643$, and $p_1 = p_2 = p_3 \approx 0.198$. When $\underline{n} = 2$, we have $\frac{\partial \pi(\mathbf{R})}{\partial \epsilon}|_{\epsilon=0} \approx 0.749 > 0$. ■

Corollary 1 tells us that under uniform distribution and with three products, offering all consumers who buy at least 2 products a small discount on their membership fee indeed strictly increases profit over the optimal two-part tariff.

7.2 General Variable Membership Fees

The fundamental reason why offering a discount that in effect creates a second (lower) level of membership fee can be profitable is that it may achieve better price discrimination, as there is one more price instrument (the discount) that the firm can use. Therefore, the firm should do at least as well with two membership fees as it does with one, and given "favorable" valuation distributions (i.e. those satisfying condition (28)), it can do strictly better.

When $n \geq 3$, the same argument applies to using a third membership fee. Generally, with more and more price instruments available to the firm, we expect the resulting price schedule to be able to mimic the full mixed bundling strategy (with $2^n - 1$ prices) more and more closely. However, the conditions also become (increasingly) more complicated and less intuitive. For this reason, we do not pursue further in this direction.

8 Two-Part Tariffs vs. Mixed Bundling

We first focus on the two-product case and then discuss the general case in the end.

When $n = 2$, we know that a mixed bundling strategy has only $2^2 - 1 = 3$ prices, and we have illustrated in the first example of section 1 that such a strategy can be replicated with a two-part tariff, which also has 3 prices when $n = 2$.

8.1 MMW's Result on Mixed Bundling

McAfee, McMillan and Whinston (1989) show a very strong result on the desirability of mixed bundling over separate pricing in the two-product case. They show a condition (condition (1) in their Proposition 1) that holds under a wide range of distributions (including all independent joint distributions), that guarantees mixed bundling strictly dominates separate pricing.

They identified this condition by using a third price instrument called *bundle discount* in addition to the two prices of the optimal separate pricing strategy. That is, they study a general IC price schedule, where the price for the bundle of two products cannot exceed the sum of the prices of two products separately - the difference being the bundle discount. They show that by offering a small bundle discount starting from the optimal separate pricing strategy, the firm increases the demand for *both* products (sold either separately or in bundles) although it lowers only *one* price (i.e. that of the bundle).

This effect is illustrated in the following figure which is directly adapted from MMW. For consumers of product 1 only, a bundle discount will lure some of them to buy product 2 as well, as will it lure some consumers of product 2 only to buy product 1 as well. This is why the bundle discount increases the demand for both products.

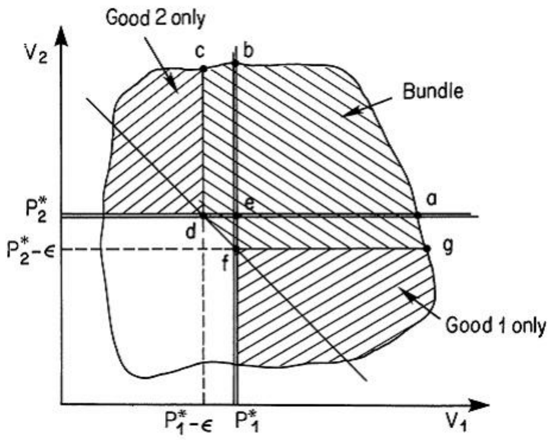


Figure II from McAfee, McMillan and Whinston (1989)

8.2 Outcome Equivalence

In the two-product case, Long (1984) shows that a two-part tariff consisting of prices of two separate products, q_1 and q_2, say, and one additional fixed fee for all consumers, m, is equivalent to a mixed bundling strategy with prices for separate products $p_1 = q_1 + m$, $p_2 = q_2 + m$ and price for the bundle $p_B = q_1 + q_2 + m$, where "*the bundle discount* $(p_1 + p_2 - p_B)$ *is like a fixed fee*".

It is true that when $n = 2$, a mixed bundling strategy and a two-part tariff can replicate each other's demand and profit - the profit maximization problems using these two strategies are therefore equivalent.

However, we want to point out that this "equivalence" exists only in the *outcome*, but *not* in the *mechanism.* To be specific, the bundle discount does not work in the same way as the membership fee, nor can it replicate the two-part tariff effect (in either two-product or multiproduct case). *In order to replicate the latter, manipulation of the bundle discount alone does not work - the prices for both single products also have to be manipulated at the same time.*

8.3 Mechanism Dichotomy

8.3.1 Difference in Demand Changes

When we consider the demand changes induced by an additional membership fee identified in Theorem 2, if we use a bundle discount ϵ instead of a membership fee m, we have the following result.

Theorem 8 *When $n = 2$, suppose $\mathbf{P}^* = \{p_1^*, p_2^*\}$ is the optimal separate pricing strategy (MSP) and induces allocation $\{A_J\}_{J \subset N}$. Consider (mixed bundling) price schedule $\mathbf{T} = \{p_1^*, p_2^*, t_{\{1,2\}} = p_1^* + p_2^* - \epsilon\}$ which induces allocation $\{C_J\}_{J \subset N}$, where ϵ represents the bundle discount. Then we have*

$$\text{(i) } \frac{\partial}{\partial \epsilon}[M(C_j) - M(A_j)]|_{\epsilon=0} < 0, \text{ for } j = 1, 2;$$
$$\text{(ii) } \frac{\partial}{\partial \epsilon}[M(C_{\{1,2\}}) - M(A_{\{1,2\}})]|_{\epsilon=0} > 0.$$

That is, the demand changes induced by a bundle discount are all first-order impacts.

Proof. Part (i):

Under $\mathbf{P}^*$ and $\mathbf{T}$, the demand segments of single product $j \in \{1, 2\}$ are respectively

$$\begin{aligned} A_j &= \{\mathbf{x} \in I^2 | x_j \geq p_j^*; x_k \leq p_k^*, k \neq j\}, \text{ and} \\ C_j &= \{\mathbf{x} \in I^2 | x_j \geq p_j^*; x_k \leq p_k^* - \epsilon, k \neq j\} \end{aligned}$$

As IC of $\mathbf{T}$ implies that $\epsilon \geq 0$, we have $C_j \subset A_j, \forall j = 1, 2$. Therefore

$$\begin{aligned} A_j \backslash C_j &= \{\mathbf{x} \in I^2 | x_j \geq p_j^*; p_k^* - \epsilon \leq x_k \leq p_k^*, k \neq j\}, \text{ and} \\ M(C_j) - M(A_j) &= -M(A_j \backslash C_j) = -\int_{p_k^* - \epsilon}^{p_k^*} \int_{p_j^*}^{1} [f(x_j, x_k) dx_j] dx_k \end{aligned}$$

And hence

$$\begin{aligned} \frac{\partial}{\partial \epsilon}[M(C_j) - M(A_j)]|_{\epsilon=0} &= -\int_{p_j^*}^{1} f(x_j, p_k^* - \epsilon) dx_j |_{\epsilon=0} \\ &= -\int_{p_j^*}^{1} f(x_j, p_k^*) dx_j \\ &< 0 \end{aligned}$$

where the last inequality is due to Lemma 8. Part (i) done.

Part (ii):

Under $\mathbf{P}^*$ and $\mathbf{T}$, the demand segments of bundle $\{1, 2\}$ are respectively

$$\begin{aligned} A_{\{1,2\}} &= \{\mathbf{x} \in I^2 | x_1 \geq p_1^*; x_2 \geq p_2^*\}, \text{ and} \\ C_{\{1,2\}} &= \{\mathbf{x} \in I^2 | x_1 \geq p_1^* - \epsilon; x_2 \geq p_2^* - \epsilon; x_1 + x_2 \geq p_1^* + p_2^* - \epsilon\} \end{aligned}$$

which implies $C_{\{1,2\}} \supset A_{\{1,2\}}$. Therefore

$$\begin{aligned} C_{\{1,2\}} \backslash A_{\{1,2\}} &= \{\mathbf{x} \in I^2 | p_1^* - \epsilon \leq x_1 \leq p_1^*; p_2^* - \epsilon \leq x_2 \leq p_2^*; x_1 + x_2 \geq p_1^* + p_2^* - \epsilon\} \\ &= (A_1 \backslash C_1) \cup (A_2 \backslash C_2) \cup D(\epsilon) \end{aligned}$$

where

$$D(\epsilon) \equiv \{\mathbf{x} \in I^2 | x_1 \leq p_1^*; x_2 \leq p_2^*; x_1 + x_2 \geq p_1^* + p_2^* - \epsilon\}$$

And hence

$$M(C_{\{1,2\}}) - M(A_{\{1,2\}}) = M(A_1 \backslash C_1) + M(A_2 \backslash C_2) + M(D(\epsilon))$$

Now we focus on $M(D(\epsilon))$, and we have

$$\begin{aligned} M(D(\epsilon)) &= \int_{p_1^*-\epsilon}^{p_1^*} \int_{p_1^*+p_2^*-\epsilon-x_1}^{p_2^*} [f(x_1, x_2) dx_2] dx_1, \text{ and} \\ \frac{\partial}{\partial \epsilon}[M(D(\epsilon))] &= \int_{p_1^*+p_2^*-\epsilon-(p_1^*-\epsilon)}^{p_2^*} f(p_1^* - \epsilon, x_2) dx_2 + \int_{p_1^*-\epsilon}^{p_1^*} f(x_1, p_1^* + p_2^* - \epsilon - x_1) dx_1 \end{aligned}$$

And therefore

$$\frac{\partial}{\partial \epsilon}[M(D(\epsilon))]|_{\epsilon=0} = 0$$

And since we have shown in the proof of part (i) that

$$\frac{\partial}{\partial \epsilon}[M(A_j \backslash C_j)]|_{\epsilon=0} = -\frac{\partial}{\partial \epsilon}[M(C_j) - M(A_j)]|_{\epsilon=0} > 0, \text{ for } j = 1, 2$$

we have

$$\begin{aligned} & \frac{\partial}{\partial \epsilon}[M(C_{\{1,2\}}) - M(A_{\{1,2\}})]|_{\epsilon=0} \\ = & \frac{\partial}{\partial \epsilon}[M(A_1 \backslash C_1)]|_{\epsilon=0} + \frac{\partial}{\partial \epsilon}[M(A_2 \backslash C_2)]|_{\epsilon=0} + \frac{\partial}{\partial \epsilon}[M(D(\epsilon))]|_{\epsilon=0} \\ > & 0. \end{aligned}$$

■

From Theorem 8 we know that the demand changes induced by a bundle discount are all *first-order* changes. This can be seen from the figure above, where a positive measure of consumers switch from buying either single product to buying the bundle (to be specific, these consumers are represented by the areas $bcde$ and $aefg$ in the figure).

This is in stark contrast to the two-part-tariff effect we have shown in Theorem 2, where the demand changes for all multiproduct bundles (including any two-product bundle) induced by an additional membership fee are all of order higher than one.

This is because the membership fee and the bundle discount are two different price instruments: The membership fee raises the final prices of both single products and the two-product bundle (i.e. it raises the final price that *all* consumers face by the same amount), which implies that there will be no consumer switching from buying either single product to buying the bundle or vice versa (as their price difference does not change). Therefore the only change in the demand for the bundle will be a loss of consumers whose valuations for two products were higher than the sum of the separate prices under separate pricing, but turn out to be insufficient to cover the additional membership fee. Their measure has to be confined in the second order, as the membership fee is shared by their valuations in two dimensions.

On the other hand, the bundle discount in mixed bundling only reduces the price for the two-product bundle, and does not change the final price of either single product. Therefore there is a change in the price difference between either single product and the bundle, which in turn leads to first-order changes in their demand.

8.3.2 Deviations in "Opposite" Directions

Having identified the different demand implications of mixed bundling and two-part tariffs, there is a second (and perhaps more striking) dichotomy between the mechanisms of mixed bundling and two-part tariffs.

Starting from the optimal separate pricing strategy, MMW's deviation involves a *decrease* in the bundle price (by offering a bundle discount), while our deviation involves an *increase* in all prices (by imposing a membership fee). Under the same condition, both deviations turn out to be profitable (when $n = 2$, our condition (21) reduces to condition (1) in MMW's Proposition 1).

The reason why these deviations in seemingly opposite directions can both be profitable is, of course, that the membership fee and the bundle discount are not exactly opposite to each other. The membership fee changes the final price that all consumers face, while the bundle discount does not change the final prices of single products. The exact opposite instrument of a membership fee is a membership "subsidy", i.e. the same amount of "discount" on both single products and the bundle (which we discuss in detail in section 6).

8.4 Average Price at Optimality - Lower or Higher?

The deviations in "opposite" directions mentioned previously may have confusing implications. For instance, since MMW's deviation involves a profitable decrease in the bundle price whilst keeping the prices for separate products unchanged, one might get an impression that the optimal mixed bundling strategy would result in a somewhat *lower* "average" price for each product compared to separate pricing. On the contrary, our result appears to render the opposite impression, i.e. a somewhat *higher* "average" price under the optimal two-part tariff compared to separate pricing, since imposing an additional membership fee on top of the optimal separate pricing strategy is profitable.

Given the outcome equivalence discussed previously, obviously only one of these two impressions can be correct. So which is it?

Actually, the answer is: *Neither.* In this section, we show several specific distributions which illustrate that, *at optimality, the "average" price a consumer pays for a product can go either way.*

First of all, we need to give "average" price a precise and sensible definition.

Definition 15 *Given price schedule* $\mathbf{P} = \{p_J\}_{J \subset N}$ *and the induced allocation* $\{A_J\}_{J \subset N}$*, the* ***average price per unit of product*** *is*

$$\bar{p} \equiv \frac{\sum_{J \subset N} p_J \cdot M(A_J)}{\sum_{J \subset N} |J| \cdot M(A_J)} = \frac{\pi(\mathbf{P})}{\sum_{J \subset N} |J| \cdot M(A_J)}.$$

That is, given a price schedule, we calculate the price for an "average" unit of product offered by the firm, which is equal to the firm's total revenue divided by the total number of units of products sold at the current price schedule. The total revenue is in turn equal to the total profit, as there is no cost of production.

This price precisely represents the expected price that an average consumer in the population pays for one unit of product sold by the firm.

Note that when $n = 2$, the optimal mixed bundling strategy and the optimal two-part tariff will imply the same $\bar{p}$, since they have outcome equivalence.

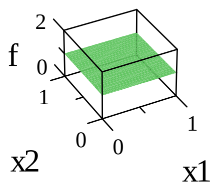
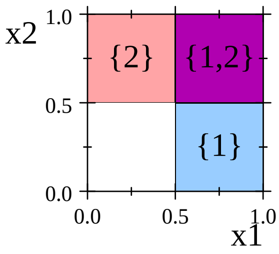
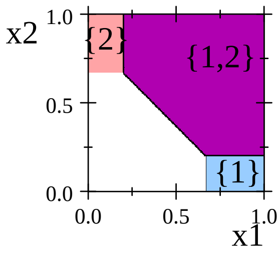
8.4.1 Uniform Distribution (Independence)

$$f(x_1, x_2) = 1, \text{ for } x_1, x_2 \in [0, 1]$$

Note that uniform distribution implies independence between x_1 and x_2.

Comparison between optimal strategies:

Table 1: Optimality Results under Uniform Distribution

$f(x_1, x_2)$ as shown	Separate Pricing		Two-Part Tariff
			
p_1^*, p_2^*	0.5	>	0.20
m^*	-		0.47
$p_j^* + m^*$	0.5	<	0.67
$p_1^* + p_2^*$ $+m^*$	1	>	0.87
$M(A_1), M(A_2)$	0.25	>	0.07
$M(A_{\{1,2\}})$	0.25	<	0.54
π $(= \sum_{J\subset N}(p_J^* + m^*)M(A_J))$	0.5	<	0.55
d $(= \sum_{J\subset N} \lvert J\rvert \cdot M(A_J))$	1	<	1.20
D $(= \sum_{J\subset N} M(A_J))$	0.75	>	0.67
$\bar{u}$ $(= \frac{d}{D})$	1.33	<	1.80
$\bar{e}$ $(= \frac{\pi}{D})$	0.67	<	0.82
$\bar{p}$ $(= \frac{\pi}{d})$	$0.5(= p_1^*)$	>	0.46

In the table above, we have shown graphs of the density function and the optimal allocations in the first row.

p_1^* and p_2^* are the optimal prices for single products in the relevant optimal price schedule (separate pricing or two-part tariff); and m^* is the optimal membership fee (only applicable to two-part tariff). When $m^* > 0$, $p_1^* + m^*$ and $p_2^* + m^*$ are the final prices for single products, as implied by the optimal two-part tariff, and $p_1^* + p_2^*$ $+m^*$ is the implied final price for the two-product bundle. These three "final prices" constitute the *optimal mixed bundling strategy* that is equivalent to the optimal two-part tariff (p_1^*, p_2^*, m^*).

$M(A_1)$ and $M(A_2)$ are the demands for single products, and $M(A_{\{1,2\}})$ is the demand for the bundle. π represents the firm's total profit (or revenue).

We have presented several new indicators in the table. In particular, $d \equiv \sum_{J\subset N} |J| \cdot M(A_J)$ represents the total units of products sold given the relevant price schedule; $D \equiv \sum_{J\subset N} M(A_J)$ represents the total measure of consumers who buy any product/bundle at all; $\bar{u} \equiv \frac{d}{D}$ represents the average number of units per consumer; $\bar{e} \equiv \frac{\pi}{D}$ represents the average expense per consumer.

In the last row of the table, we show the "average price per unit of product" $\bar{p}$, as defined in Definition 15. We see that with uniform distribution, $\bar{p}$ is **lower** under two-part tariff/mixed bundling compared to separate pricing $(0.46 < 0.5)$.

8.4.2 Highly Positively Correlated Distribution

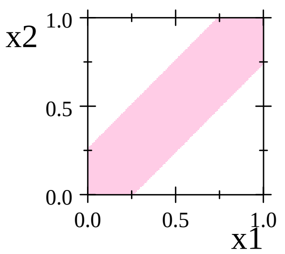
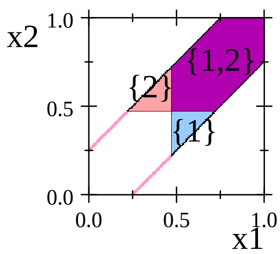
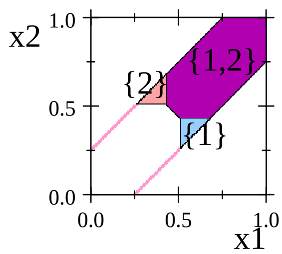
$$f(x_1, x_2) = \begin{cases} \frac{16}{7} - 4.375 \times 10^{-101} & \text{, if } \; x_1 - \frac{1}{4} \leq x_2 \leq x_1 + \frac{1}{4} \text{ and } 0 \leq x_1, x_2 \leq 1; \\ 10^{-100} & , \quad \text{otherwise.} \end{cases}$$

This distribution has correlation coefficient $corr[x_1, x_2] \approx 0.8617 > 0$. Under this distribution, virtually all consumers are distributed evenly within a $\frac{\sqrt{2}}{4}$-wide range along the 45-degree line of the support. We have made the distribution outside this range so "thin" that it is negligible in all the calculations.

We show below that even with this highly positively correlated joint distribution, two-part tariff still strictly dominates separate pricing. Moreover, the average price that the firm charges per unit of product, $\bar{p}$, is **higher** under the optimal two-part tariff/mixed bundling than under separate pricing $(0.471 > 0.469)$.

Comparison between optimal strategies:

Table 2: Optimality Results under Highly Positively Correlated Distribution

$f(x_1, x_2)$ as shown	Separate Pricing		Two-Part Tariff
			
p_1^*, p_2^*	$\frac{15}{32} \approx 0.47$	$>$	$\frac{41}{96} \approx 0.43$
m^*	-		$\frac{1}{12} \approx 0.08$
$p_j^* + m^*$	$\frac{15}{32} \approx 0.47$	$<$	$\frac{49}{96} \approx 0.51$
$p_1^* + p_2^* + m^*$	$\frac{15}{16} \approx 0.94$	$=$	$\frac{15}{16} \approx 0.94$
$M(A_1), M(A_2)$	$\frac{1}{14} \approx 0.07$	$>$	$\frac{2}{63} \approx 0.03$
$M(A_{\{1,2\}})$	$\frac{13}{28} \approx 0.47$	$<$	$\frac{127}{252} \approx 0.50$
$\pi \; (= \sum_{J \subset N} (p_J^* + m^*) M(A_J))$	$\frac{225}{448} \approx 0.502$	$<$	$\frac{6107}{12\,096} \approx 0.505$
$d \; (= \sum_{J \subset N} \lvert J \rvert \cdot M(A_J))$	$\frac{15}{14} \approx 1.07$	$=$	$\frac{15}{14} \approx 1.07$
$D \; (= \sum_{J \subset N} M(A_J))$	$\frac{17}{28} \approx 0.61$	$>$	$\frac{143}{252} \approx 0.57$
$\bar{u} \; (= \frac{d}{D})$	$\frac{30}{17} \approx 1.76$	$<$	$\frac{270}{143} \approx 1.89$
$\bar{e} \; (= \frac{\pi}{D})$	$\frac{225}{272} \approx 0.83$	$<$	$\frac{6107}{6864} \approx 0.89$
$\bar{p} \; (= \frac{\pi}{d})$	$\frac{15}{32} (= p_1^*) \approx 0.469$	$<$	$\frac{6107}{12\,960} \approx 0.471$

A rather interesting observation of the outcome shown above is that the total units of products sold, d, remains *exactly the same* under separate pricing and two-part tariff. Since the profit is increased under two-part tariff, the average price per unit of product goes up.

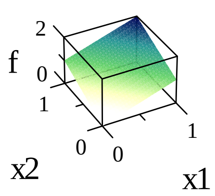
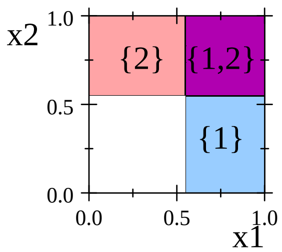
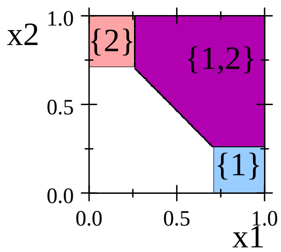
8.4.3 Negatively Correlated Distribution

$$f(x_1, x_2) = x_1 + x_2, \text{ for } x_1, x_2 \in [0, 1]$$

This distribution has correlation coefficient $corr[x_1, x_2] = -\frac{1}{11} \approx -0.09 < 0$.

Comparison between optimal strategies:

Table 3: Optimality Results under Negatively Correlated Distribution

$f(x_1, x_2)$ as shown	Separate Pricing		Two-Part Tariff
			
p_1^*, p_2^*	0.55	$>$	0.26
m^*	-		0.45
$p_j^* + m^*$	0.26	$>$	0.07
$p_1^* + p_2^* + m^*$	0.32	$<$	0.60
$M(A_1), M(A_2)$	0.55	$<$	0.71
$M(A_{\{1,2\}})$	1.1	$>$	0.97
$\pi\ (= \sum_{J \subset N}(p_J^* + m^*)M(A_J))$	0.63	$<$	0.69
$d\ (= \sum_{J \subset N} \lvert J \rvert \cdot M(A_J))$	1.15	$<$	1.35
$D\ (= \sum_{J \subset N} M(A_J))$	0.83	$>$	0.75
$\bar{u}\ (= \frac{d}{D})$	1.38	$<$	1.80
$\bar{e}\ (= \frac{\pi}{D})$	0.76	$<$	0.93
$\bar{p}\ (= \frac{\pi}{d})$	0.55	$>$	0.51

Under this distribution, the average price per unit of product is **lower** under two-part tariff/mixed bundling than under separate pricing ($0.51 < 0.55$).

8.4.4 Discussion

The examples above show that the "impressions" about the average price that one might derive from either MMW's result or ours are not reliable. This is because both results only reveal *local* deviations that lead to profit improvements. Caution should be taken when we use these results to infer properties of *globally optimal* strategies.

8.5 Two-Part Tariff vs. Mixed Bundling: General Case when $n > 2$

In our setting, a general price schedule can be thought of as a mixed bundling strategy, in the sense that it specifies all the prices of all possible bundles. Therefore the two-part tariff in Definition 10 is generally a special case of mixed bundling.

When $n > 2$, the number of possible multiproduct bundles is larger than 1. For instance, when $n = 3$, there are 3 two-product bundles and 1 three-product bundle. These bundles may all have different prices, each of which may in turn also differ from the sum of the prices of the component single products and/or component bundles. Unlike the case when $n = 2$, there is no longer a standard way to define the "bundle discount" of a mixed bundling strategy when $n > 2$. In general, for each bundle and each of its partitions (into component bundles/products), there is a relevant "discount" that is the difference between the bundle price and the sum of the prices of its component bundles/products. Since the total number of bundles involved in a mixed bundling strategy increases exponentially with the number of products, n, the total number of "bundle discounts" increases even more quickly. Therefore it becomes impossible to fully characterize the impact of "bundle discounts" on profits (i.e. MMW's condition (1)) in the general multiproduct case.

That is, we do not know generally whether offering bundle discount(s) would be profitable or not when $n > 2$.

However, our condition (21) of Theorem 4 and condition (28) of Theorem 7 bridge this gap partly.

As we have illustrated in the three-product example in section 1 and again in the early part of section 4, the two-part tariff we study creates "bundle discounts" at constant steps equal to the membership fee m, according to the size of the bundle. That is, under two-part tariff, a consumer of any bundle J gets a "bundle discount" of $(|J| - 1) \cdot m$ compared to the consumers of the $|J|$ individual products.

Our Theorem 4 says that bundle discounts offered in this way will indeed be profitable given condition (21).

Our Theorem 7 says that offering a further discount on membership fees for consumers of bundles of certain sizes or larger will further increase profit given condition (28).

These are generalizations of the two-product result by MMW to the multiproduct case.

8.6 Correlation and Theorem 4

In this section we examine how the correlation between consumers' valuations for two products affects the profitability of two-part tariffs.

The bundling literature has provided insights through particular examples and distributions that negative correlation generally works in favor of bundling strategies rather than separate pricing. By outcome equivalence in the two-product case discussed in section 8.2, these insights naturally apply to two-part tariffs as well.

Actually, we have the following conclusion for a family of negatively correlated distributions that is more general than the existing results in the literature.

Corollary 2 *Suppose $n = 2$ and $\mathbf{P} = (p_1^*, p_2^*)$ is a monopoly separate pricing strategy, then two-part tariff strictly dominates separate pricing if* ***any*** *of the following conditions holds:*

(i) $F_{1|2}(p_1^*|\, x_2)$ is strictly increasing in x_2 for all $x_2 > p_2^*$ in I^2; or

(ii) $F_{2|1}(p_2^*|\, x_1)$ is strictly increasing in x_1 for all $x_1 > p_1^*$ in I^2; or

(iii) The two functions above are both constants for all $x_2 > p_2^*$ and $x_1 > p_1^*$ in I^2, respectively.

Proof. When $n = 2$, since $\mathbf{P} = (p_1^*, p_2^*)$ is MSP, we have for $j = 1, 2$,

$$\begin{aligned} p_j^* &= \frac{1 - F_j(p_j^*)}{f_j(p_j^*)} \in (0, 1) \\ M(A_{\{j\}}) &= 1 - F_j(p_j^*) - M(A_{\{1,2\}}) \end{aligned}$$

Therefore the left-hand side of condition (21) of Theorem 4 becomes

$$\begin{aligned} \frac{\partial \Delta\pi}{\partial m}|_{m=0} &= M(A_{\{1\}}) - p_1^* \cdot \int_0^{p_2^*} f(p_1^*, x_2)dx_2 + M(A_{\{2\}}) - p_2^* \cdot \int_0^{p_1^*} f(x_1, p_2^*)dx_1 + M(A_{\{1,2\}}) \\ &= [1 - F_1(p_1^*)][1 - \int_0^{p_2^*} \frac{f(p_1^*, x_2)}{f_1(p_1^*)}dx_2] + [1 - F_2(p_2^*)][1 - \int_0^{p_1^*} \frac{f(x_1, p_2^*)}{f_2(p_2^*)}dx_1] - M(A_{\{1,2\}}) \\ &= [1 - F_1(p_1^*)][1 - F_{2|1}(p_2^*|\, p_1^*)] + [1 - F_2(p_2^*)][F_{1|2}(p_1^*|\, p_2^*)] - M(A_{\{1,2\}}) \end{aligned}$$

(i) For all $x_2 > p_2^*$ in I^2, since $F_{1|2}(p_1^*|\, x_2)$ is strictly increasing in x_2 and bounded from above by 1, we must have $F_{1|2}(p_1^*|\, p_2^*) < 1$ and

$$1 - F_{1|2}(p_1^*|\, x_2) < 1 - F_{1|2}(p_1^*|\, p_2^*) \tag{30}$$

Denote $I \equiv 1 - F_{1|2}(p_1^*|\, p_2^*)$, then I is a positive constant.
(30) $\Rightarrow \int_{p_1^*}^1 \frac{f(x_1,x_2)}{f_2(x_2)}dx_1 < I \Rightarrow \int_{p_1^*}^1 f(x_1, x_2)dx_1 \leq f_2(x_2) \cdot I$
Integrate both sides with respect to x_2 on $[p_2^*, 1]$, we get
$\Rightarrow \int_{p_2^*}^1 \int_{p_1^*}^1 f(x_1, x_2)dx_1 dx_2 < \int_{p_2^*}^1 f_2(x_2)dx_2 \cdot I \Rightarrow$

$$M(A_{\{1,2\}}) < [1 - F_2(p_2^*)][1 - F_{1|2}(p_1^*|\, p_2^*)] \tag{31}$$

Thus by (31) and $[1 - F_1(p_1^*)][1 - F_{2|1}(p_2^*|\, p_1^*)] \geq 0$, we have:

$$\frac{\partial \Delta\pi}{\partial m}|_{m=0} > 0.$$

(ii) Using symmetry of our model regarding the two products $j = 1, 2$, we only need to swap the product labels and we are done.

(iii) Since f is atomless (Assumption 2), when $F_{1|2}(p_1^*|\, x_2)$ is constant for all $x_2 > p_2^*$ in I^2, we have

$$1 - F_{1|2}(p_1^*|\, x_2) = 1 - F_{1|2}(p_1^*|\, p_2^*)$$

By the same argument as in (i), except that all the inequalities need to be changed to equations since now $F_{1|2}(p_1^*|\, x_2)$ is a constant, we get the equality version of (31):

$$M(A_{\{1,2\}}) = [1 - F_2(p_2^*)][1 - F_{1|2}(p_1^*|\, p_2^*)]$$

Similarly, from $F_{2|1}(p_2^*|\, x_1)$ being a constant for all $x_1 > p_1^*$ in I^2, we have

$$M(A_{\{1,2\}}) = [1 - F_1(p_1^*)][1 - F_{2|1}(p_2^*|\, p_1^*)]$$

Thus we have

$$\frac{\partial \Delta\pi}{\partial m}|_{m=0} = M(A_{\{1,2\}}) > 0.$$

■

It can be shown that conditions (i) and (ii) above are both satisfied by the negatively correlated distribution we discussed in section 8.4.3 (i.e. $f(x_1, x_2) = x_1 + x_2$, for $x_1, x_2 \in [0, 1]$), and condition (iii) is satisfied by the uniform distribution in section 8.4.1.

With positive correlation, we cannot draw a general conclusion as in the case of negative correlation. However, condition (21) of Theorem 4 can still hold with positive correlations. Actually, it does hold under the highly positively correlated distribution discussed in section 8.4.2 (with correlation coefficient≈ 0.9), where we can show that $\frac{\partial \Delta\pi}{\partial m}|_{m=0} = \frac{1}{14} > 0$.

9 Conclusion

Two-part tariffs are prevalent in life. In many cases, they involve more than one product provided by the same firm, such as the landline telephone and broadband services from a telephone company, or a bank account through which many services such as debit card, credit card, mortgage and travel insurance are provided.

As a multiproduct pricing strategy, a two-part tariff has the appeal of simplicity (in that it consists only $n + 1$ prices, for totally n products) and implementability (in that it simply adds on top of a separate strategy an additional membership fee). It is an easy way for a multiproduct firm to achieve price discrimination. We have examined the underlying general mechanism, the two-part-tariff effect, which distinguishes two-part tariffs from other kinds of pricing strategies, including mixed bundling.

The two-part-tariff effect also helps us understand real-life pricing strategies that may not look like two-part tariffs, such as free gift (as "membership subsidy") and discount on parking charge (as "variable membership fee"), which are actually general two-part tariffs in disguise.

Our results on the desirability of two-part tariffs have nothing to do with production cost or efficiency, and apply to general firms providing two or more products. They therefore provide a new perspective for the prevalence of two-part tariffs, as well as a practical way for multiproduct firms that use separate pricing for some reason, to increase profits - imposing a small membership fee or subsidy would generally be profit-improving. Another practical implication of this new perspective is that two-part tariffs should be subject to the same regulatory scrutiny as other discriminatory pricing strategies.

References

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

[2] Armstrong, M. (1996): “Multiproduct Nonlinear Pricing”, *Econometrica*, 64(1), pp. 51-75.

[3] Armstrong, M. (1999): "Price Discrimination by a Many-Product Firm", *Review of Economic Studies*, 66(1), pp. 151-168.

[4] Banal-Estanol, A. and Ottaviani, M. (2007): "Notes on Pure Bundling", working paper.

[5] Chu, C. S., Leslie, P. and Sorensen, A. (2009): "Bundle-Size Pricing as an Approximation to Mixed Bundling", *American Economic Review*, forthcoming.

[6] Conitzer, V. and Sandholm, T. (2003): "Automated Mechanism Design: Complexity Results Stemming from the Single-Agent Setting". *Proceedings of the 5th International Conference on Electronic Commerce (ICEC-03)*, Pittsburgh, PA, USA, pp. 17-24.

[7] Fang, H. and Norman, P. (2006): "To Bundle or Not to Bundle", *Rand Journal of Economics*, 37(4), pp. 946-963.

[8] Gao, M. (2009): "When to Allow Buyers to Sell? Bundling in Mixed Two-Sided Markets", working paper.

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

[10] Madarász, K. and Prat, A. (2010): "Approximate Nonlinear Pricing", working paper.

[11] 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.

[12] McAfee, R. P. and McMillan, J. (1988): "Multidimensional Incentive Compatibility and Mechanism Design", *Journal of Economic Theory*, 46, pp. 335-354.

[13] 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.

[14] Rochet, J. and Chone, P. (1998): "Ironing, Sweeping, and Multidimensional Screening", *Econometrica*, 66(4), pp. 783-826.

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

Part III
A Quasi-Cumulative Weighting Function for Prospect Theory: The (β, c) Model*

Ming Gao†

First version: December 2008. This version: July 2010

Abstract

We propose a simplified weighting function for cumulative prospect theory (CPT), which plays a similar role in models with risky choice as that of the quasi-hyperbolic discounting function in models with intertemporal choice. The inverse S-shaped weighting of cumulative probabilities posited in CPT causes difficulties in representation which hinders its application in wider situations of risky choice. The (β, c) model we propose has a weighting function that is linear with slope smaller than 1 on the open interval $(0, 1)$, jumps down to 0 at end point 0, and jumps up to 1 at end point 1. It achieves highly tractable utility representation for CPT whilst preserving the basic tenets of CPT. It by construction can explain all four major phenomena of risky choice violating the standard model that CPT was developed to reconcile, including reference dependence and certainty effect. It also allows Bayesian updating (with distortions at certainty) which CPT cannot accommodate. We systematically examine the (β, c) representation of discrete and continuous lotteries, and provide four applications which illustrate that the (β, c) model is a useful work horse to analyze implications of preferences exhibiting certainty effect and reference dependence in standard models.

Key Words: cumulative prospect theory, weighting function, risky choice, certainty effect.

JEL **Classification**: D03, D81, G11.

---

*I thank Roland Bénabou, Jean-Pierre Benoît, Francesca Cornelli, Emeric Henry, Vijay Krishna, Eric Maskin, Emre Ozdenoren, Kareen Rozen, Peter Wakker, John Wooders and Horst Zank for valuable discussions and comments. I also thank participants at FUR XIV International Conference 2010.

†Economics Department, London Business School, Regent's Park, London NW1 4SA, UK. Email: mgao.phd2005@london.edu. Homepage: phd.london.edu/mgao/.

Contents

1 Introduction

Tversky and Kahneman (1992) propose a cumulative version of their prospect theory (1979), called the cumulative prospect theory (henceforth, CPT). Among other things, they posit that, in risky choice situations, people do not use the true probabilities of outcomes to evaluate "lotteries" but use decision weights that are functions of these probabilities. Moreover, the weight people attach to each outcome does not only depend on the true probability of that outcome, but also depends on the probabilities of other outcomes - a.k.a. the *cumulative probability.*

They show with experiments that the so-called weighting functions are *inverse S-shaped* - they transform true probabilities non-linearly in such a way that they are more sensitive to probability changes (and hence are steeper) closer to impossibility and certainty, and less sensitive in the interim (and hence are flatter). The functional forms they used to fit experimental data are continuous functions with changing curvature throughout the unit interval.

The preference representation that Tversky and Kahneman (1992) propose is then a *cumulative functional* consisting of a complex system of functions that invokes inverse S-shaped weighting of cumulative probabilities. With complexity in the functional forms, they gain accuracy in approaching the "true" decision weights. However, what is lost is tractability in utility representation. This trade-off is not always worthwhile.

As a descriptive model, the CPT cumulative functional is quite powerful in reconciling, within one unified model, the "anomalous" phenomena that violate standard utility models. However, it is unlikely that there exists a "precise" cumulative functional, as Tversky and Kahneman (1992) said: *"..., the cumulative functional is unlikely to be accurate in detail. We suspect that decision weights may be sensitive to the formulation of the prospects, as well as to the number, the spacing and the level of outcomes."*

As a preference representation, the CPT cumulative functional is quite cryptic. Without calculation, it is difficult to comprehend what the representation of a given lottery is, and in most cases it is impossible to compare two lotteries that do not involve some kind of "dominance" (e.g. first-order stochastic dominance). We feel that this lack of tractability hinders the application of CPT in a wider range of risky choice situations, which motivates our quest for a simplified representation for CPT. In this quest, we find that a small sacrifice in precision of weighting yields a big reward in representational transparency.

We propose a simplified weighting function, named the (β, c) model, that achieves highly tractable utility representation of CPT for *risky choice.*[1] This function is linear with slope smaller than 1 on the open interval $(0, 1)$, jumps down to 0 at end point 0, and jumps up to 1 at end point 1. We use β for the slope and c for the intercept. This simplification preserves the key features of the "true" weighting functions, which are

[1]Besides risky choice, CPT also applies to uncertain choice (i.e. choice under ambiguity). In this paper, however, we restrict our discussion of the (β, c) model to risky choice only.

the oversensitiveness to probability changes close to impossibility and certainty, and the undersensitiveness in the interim, whilst only sacrificing precision along the curvature of the “true” functions. With this new weighting function and the original CPT cumulative functional, we systematically examine discrete and continuous lotteries and derive their representation.

The (β, c) representation of a general lottery consists of a linear part that is the "discounted" expected value of the lottery, plus additional value terms of the reference point and the extreme outcomes (if any) of the lottery. Each part in the representation has a transparent interpretation, and the whole representation is highly tractable. The certainty and possibility effects, and reference dependence, for instance, are all shown in concise and intuitive analytical forms.

The (β, c) representation can work as a relatively nice descriptive model of risky choice, because by construction it can explain all four major phenomena of risky choice violating the standard model that CPT was developed to reconcile, including: framing effect, nonlinear preferences, risk seeking and loss aversion.

Besides, the (β, c) representation does two things that CPT does not do. Firstly, it allows for *Bayesian updating of probabilities* with distortions at certainty, due to the linear part of the (β, c) model. CPT, however, cannot accommodate Bayesian updating as the decision weights under CPT are everywhere non-linear in probabilities. In section 5 we provide an application of dynamic decision making that makes use of this feature, where decision makers with (β, c) representation can exhibit time-inconsistent behavior. Secondly, the (β, c) representation *does not converge to the standard expected-utility representation* as a lottery converges to a deterministic outcome, due to discontinuity of the (β, c) model at impossibility and certainty. This feature can change standard results that rely on limiting properties of sequences of lotteries, and we provide an application in section 7 that shows its impact on trembling-hand perfection in finite strategic games.

We also provide two other applications of the (β, c) representation - one on the existence of mixed strategy Nash equilibrium in finite strategic games (section 6) and the other on static investment decision making (appendix). These applications illustrate that the (β, c) model is a useful tool for analyzing implications of preferences exhibiting certainty effect in standard models.

An analogy can be drawn between the role of our (β, c) model in risky choice situations and that of the (β, δ) model (a.k.a. quasi-hyperbolic discounting function, Phelps and Pollak (1968) and Laibson (1997)) in intertemporal choice situations. While there is evidence showing that people’s true intertemporal preference is best described by a hyperbolic discounting function, the (β, δ) model instead uses a simplified quasi-hyperbolic function, which also tremendously reduces the burden in representation whilst preserving the "present bias" that is central to hyperbolic discounting behavior in intertemporal choice.

It was not until the (β, δ) model became the prevalent work horse of behavioral models of intertemporal choice, that the wide implications and applications of present-biased preference were systematically studied and revealed. Using the (β, δ) model, many elegant behavioral theories burgeoned and helped us understand some most interesting intertemporal choice behaviors much more clearly. (For instance, O'Donoghue and Rabin (1999)'s account of procrastination.) We hope the easily tractable and straightforward closed-form representation implied by our (β, c) model will help expand the scope of applications of cumulative prospect theory studied by behavioral economists.

2 A Reminder of the General CPT Model

In cumulative prospect theory (Tversky and Kahneman (1992)), a **lottery**, which is a (discrete) probability measure on X (a bounded subset of $\mathbb{R}$), is defined as:

$$f \equiv (x_{-m}, p_{-m}; ...; x_{-1}, p_{-1}; x_0, p_0; x_1, p_1; ...; x_n, p_n)$$

where outcomes x_i are listed in *ascending* order, i.e. $x_{-m} < ... < x_{-1} < x_0 < x_1 ... < x_n$, and $\sum_{i=-m}^{n} p_i = 1$. We denote by $\mathbf{x}$ the underlying random variable of f.

x_0 is the **reference point** - the neutral outcome. In CPT it is assumed that $x_0 = 0$.

Note lottery f does not necessarily have the neutral outcome x_0, i.e. it may be that $p_0 = 0$. Without loss of generality, we require strictly positive probabilities for all other outcomes, i.e. $p_i > 0 \ \forall i \neq 0$.

Let

$$p^+ \equiv \sum_{k=1}^{n} p_k = 1 - \sum_{k=-m}^{-1} p_k - p_0 = \Pr[\mathbf{x} > x_0]$$
$$p^- \equiv \sum_{k=-m}^{-1} p_k = 1 - \sum_{k=1}^{n} p_k - p_0 = \Pr[\mathbf{x} < x_0]$$

Then $p^+ + p^- + p_0 = 1$. Let

$$f^+ \equiv (x_0, (p_0 + p^-); x_1, p_1; ...; x_n, p_n)$$
$$f^- \equiv (x_{-m}, p_{-m}; ...; x_{-1}, p_{-1}; x_0, (p_0 + p^+))$$

denote the **positive and negative parts** of f, respectively, which are also lotteries. Note that the probabilities that originally correspond to negative (resp. positive) outcomes get re-assigned to x_0 in f^+ (resp. f^-), which will cause a double-counting of outcome x_0 in the final evaluation of f.

CPT says that the decision maker evaluates lottery f according to a **cumulative functional** V :

$$V(f) \equiv V(f^+) + V(f^-) \tag{1}$$

where

$$\begin{aligned} V(f^+) &= \sum_{i=0}^{n} \pi_i^+ v(x_i) \\ V(f^-) &= \sum_{i=-m}^{0} \pi_i^- v(x_i) \end{aligned}$$

where $v(\cdot)$ is a strictly increasing real-valued function with $v(0) = 0$, and π_i^+ (resp. π_i^-) is the **weight** attached to outcome x_i in the evaluation of f^+ (resp. f^-).

In the first version of prospect theory (Kahneman and Tversky (1979)), π_i^+, say, is a function of p_i only and is independent of all other p_k, $k \neq i$. In CPT, however, π_i^+ depends on the (de-)**cumulative probability** of outcome i, i.e. all probabilities $p_i, p_{i+1}, ..., p_n$; π_i^- depends on all probabilities $p_{-m}, p_{-m+1}, ..., p_{i-1}, p_i$. Their particular functional forms are:

$$\begin{aligned} \pi_0^+ &= w^+(1) - w^+(\sum_{k=1}^{n} p_k) \\ \pi_i^+ &= w^+(\sum_{k=i}^{n} p_k) - w^+(\sum_{k=i+1}^{n} p_k), \text{ for } 0 < i < n \\ \pi_n^+ &= w^+(p_n) \\ \pi_0^- &= w^-(1) - w^-(\sum_{k=-m}^{-1} p_k) \\ \pi_i^- &= w^-(\sum_{k=-m}^{i} p_k) - w^-(\sum_{k=-m}^{i-1} p_k), \text{ for } -m < i < 0 \\ \pi_n^- &= w^-(p_{-m}) \end{aligned}$$

where w^+ and w^- are called **weighting functions**, which are both strictly increasing functions mapping $[0, 1]$ into itself, satisfying

$$w^+(0) = w^-(0) = 0, \text{ and } w^+(1) = w^-(1) = 1 \tag{2}$$

Tversky and Kahneman (1992) suggest weighting functions that are *inverse S-shaped*, like those shown in the following figure:

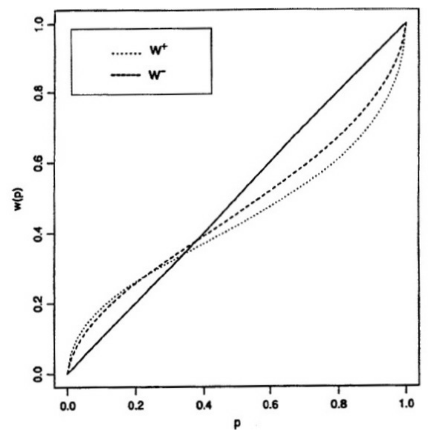


Figure 1: Weighting Functions of CPT

The functional form that they used to fit their experimental data is the following:

$$w(p) = \frac{p^\gamma}{(p^\gamma + (1-p)^\gamma)^{1/\gamma}}$$

where parameter γ takes different values for w^+ and w^-.

Tractability of CPT Representation From the description so far, we see that the preference representation under CPT, the cumulative functional in (1), consists of a complex system of functions. The complexity mainly comes from inverse S-shaped weighting of cumulative probabilities.

As a descriptive model, CPT is quite powerful in reconciling "anomalous" phenomena that violate standard utility models. We provide a brief discussion of its descriptive power as well as that of the (β, c) model in section 3.5.

However, the CPT representation is quite cryptic and it is not easy to comprehend what the representation of a given lottery is in (1), even with the help of the decision weights π^+, π^- and weighting functions w^+, w^-. Without calculating the value of V, it is even harder, and impossible in most cases, to compare two lotteries that do not involve some kind of straightforward dominance (e.g. first-order stochastic dominance).

It is this lack of tractability of the original CPT cumulative functional that motivates our quest for a simplified representation for CPT.

3 The (β, c) Model

In this part, we propose a simplified weighting function for CPT, which we call the (β, c) model. Its relationship to the "true" weighting function is similar to that between the quasi-hyperbolic discounting function and the hyperbolic discounting function. We will show that the (β, c) model yields a highly tractable preference representation that keeps the gist of CPT.

Definition 1 ((β, c) Model) *The (β, c) model has weighting functions satisfying $w^+ = w^- = w$, and*

$$w(p) = \begin{cases} 0 & , \text{ if } p = 0 \\ \beta p + c & , \text{ if } 0 < p < 1 \\ 1 & , \text{ if } p = 1 \end{cases} \tag{3}$$

where β and c are constants and satisfy the following conditions:

$$0 < \beta \leq 1, 0 \leq c < 1, \beta + c \leq 1.$$

This weighting function w only keeps the true probability as the decision weights at the two extremes, 0 and 1.

When $\beta = 1$ and $c = 0$, the weight is exactly the probability, and the resulting CPT model is similar to the expected utility (henceforth EU) model[2]. In other cases, this weighting function will have discontinuities at at least one of the two ends of $[0, 1]$, and it is linear but less "sensitive" in probabilities anywhere in the middle (i.e. flatter than the 45-degree line).

The following is a graph of the (β, c) model:

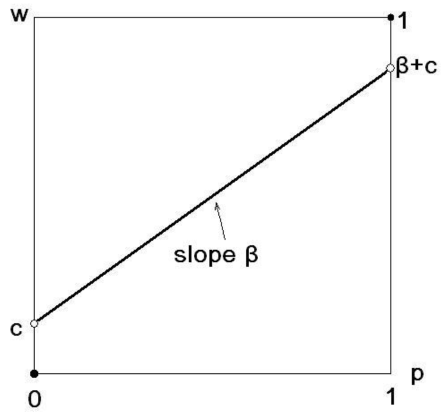


Figure 2: Weighting Function of the (β, c) Model

Interpretation of (β, c) Parameters Evident from the figure above, c (the distance of the "jump" of the value of w at $p = 0$) captures the **possibility effect** at probability 0, or the "*overweighting of very small probabilities*"; $(1 - \beta - c)$ (the distance of the "jump" of the value of w at $p = 1$) captures the **certainty effect** at probability 1, or the "*underweighting of probabilities very close to certainty*". β represents the slope of the linear part, and measures the **lack of sensitivity to intermediate probabilities**, as well as part of the certainty effect captured by $(1-\beta-c)$. The two parameters β and c will continue to manifest themselves in the representations we derive from the (β, c) model.

Notice that when $c = 0$, that is, when there is no possibility effect, the β parameter alone captures both the lack of sensitivity to intermediate probabilities and the certainty effect $(1-\beta)$. A calibrated (β, c) model with $c = 0$ will be used in most of the applications we present later in the paper.

Note that the experiment evidence by Tversky and Kahneman (1992) is consistent with our assumption of $w^+ = w^-$ in the (β, c) model. More discussion of empirics is postponed until section 4.

Degenerate Lotteries Before we derive the (β, c)-representation for general lotteries, it is useful to emphasize the (obvious) fact that the (β, c) model preserves the "original"

[2]Even when $\pi_i^+ = p_i, \pi_j^- = p_j$ for all i, j, $V(f)$ is still not necessarily equivalent to the expected utility of f, unless $v(x_0) = 0$. This is because the outcome x_0 gets double counted in the evaluation of f^+ and f^- in CPT, and we generally have $V(f) = Ev(f) + v(x_0)$ when $\pi_i^+ = p_i, \pi_j^- = p_j$ for all i, j.

preference represented by $v(\cdot)$ in *riskless* situations. That is, for "degenerate" lotteries - those with *sure* outcomes - the (β, c) representation is exactly the same as the value function $v(\cdot)$.

Theorem 0 (Degenerate Lottery) *If $f = (x, p = 1)$, then f is evaluated in the (β, c) model by*

$$V(f) = v(x).$$

Unless otherwise stated, all our following results are about **non-degenerate lotteries** - those with *no* sure outcomes.

3.1 Discrete Lotteries

Consider the discrete lotteries described in section 2. In the (β, c) model, we have the following result.

Theorem 1 (Discrete Lottery) *Given x_0, if discrete lottery f has outcomes y and z such that $y < x_0 < z$, then it is evaluated in the (β, c) model by the following function V:*

$$V(f) = \beta Ev(f) + cv(x_n) + cv(x_{-m}) + 2(1 - \beta - c)v(x_0) + \beta v(x_0) \tag{4}$$

Proof. Under the (β, c) model, from (3) we have

$$\begin{aligned} \pi_0^+ &= \beta(p_0 + p^-) + (1 - \beta - c) \\ \pi_i^+ &= \beta p_i, \text{for } 0 < i < n \\ \pi_n^+ &= \beta p_n + c \\ \pi_0^- &= \beta(p_0 + p^+) + (1 - \beta - c) \\ \pi_i^- &= \beta p_i, \text{for } -m < i < 0 \\ \pi_{-m}^- &= \beta p_{-m} + c \end{aligned} \tag{5}$$

and

$$\begin{aligned} V(f^+) &= \beta \sum_{i=0}^{n} p_i v(x_i) + cv(x_n) + (1 - \beta - c)v(x_0) + \beta p^- v(x_0) \\ V(f^-) &= \beta \sum_{i=-m}^{0} p_i v(x_i) + cv(x_{-m}) + (1 - \beta - c)v(x_0) + \beta p^+ v(x_0) \end{aligned} \tag{6}$$

Substituting into $V(f) = V(f^+) + V(f^-)$ and we get (4). ■

Theorem 1 focuses on the case when $y < x_0 < z$. Other situations are discussed in section 3.2 as special cases.

Notice in the proof above that, when $0 < i < n$ or $-m < i < 0$, the decision weight attached to outcome x_i is simply βp_i, which does not depend on the cumulative probability. The cumulative probability only affects the weight when $i = 0, -m$, or n, i.e. the reference point or the two extreme outcomes. This is why we also call the weighting function of the (β, c) model **quasi-cumulative**.

Theorem 1 holds for lotteries with $p_0 = 0$ and $p_0 > 0$, i.e. it does not matter whether the reference point x_0 is an outcome of this lottery. Below we provide an interpretation of (4) in the case when $p_0 > 0$, and it is straightforward to extend it to the case when $p_0 = 0$.

Interpretation of the (β, c) Representation for Discrete Lotteries

- The part $\beta Ev(f)$ in (4) reflects the fact that the (β, c) model has a weighting function that is linear in probabilities for all the intermediate outcomes. And because the decision weights attached to outcomes are less sensitive than their respective true probability, the true EU of the lottery is "discounted" by factor $\beta \leq 1$.

- The part $cv(x_n)+cv(x_{-m})+2(1-\beta-c)v(x_0)$ is the direct result of the possibility effect and the certainty effect captured by the (β, c) model. The probability p of an outcome is transformed into $\beta p + c$ in the representation. Evident from (5) above, because of the linearity of the weighting function in intermediate probabilities, combined with the cumulative nature of the weighting function in CPT, each intermediate outcome x_i only gets weight βp_i, which goes into $\beta Ev(f)$ in the representation, and only the reference point and extreme outcomes get different weights (see (6)):

  - In both f^+ and f^-, due to cumulative weighting, the reference point x_0 becomes the "sure" outcome. For instance, in f^+, x_0 is the minimum gain, and its (de-)cumulative probability (i.e. the probability of the outcome of f^+ no smaller than x_0) is 1. Therefore, in the evaluation of f^+, x_0 gets overweighted relative to the other outcomes by $(1 - \beta - c)$, due to certainty effect. Similarly, x_0 is the minimum loss in f^-, and it also gets overweighted by $(1 - \beta - c)$ in the evaluation of f^-. This results in the extra term $2(1 - \beta - c)v(x_0)$.
  - In f^+, the extreme outcome x_n has (de-)cumulative probability (i.e. the probability of the outcome of f^+ no smaller than x_n) close to 0, and therefore it gets overweighted relative to the other outcomes by factor c due to the possibility effect, which results in the additional term $cv(x_n)$.
  - In f^-, the extreme outcome x_{-m} has cumulative probability close to 0 and thus gets overweighted relative to the other outcomes by factor c due to the possibility effect, which results in the additional term $cv(x_{-m})$.

- The last part $\beta v(x_0)$ is due to double-counting of the reference point. We know f^+ (resp. f^-) truncates f from x_0 upwards (resp. downwards), and re-assigns all the remaining probability p^- (resp. p^+) to x_0. This leads to a complete double-counting of x_0 with probabilities p^+, p^-, and p_0 (which sum up to 1) in the final evaluation of x_0. After discounting each probability by β, this results in an addition term $\beta v(x_0)$.

Corollary 1 *If $v(x_0) = 0$, the (β, c) model gives the following representation of discrete lottery f with outcomes y and z such that $y < x_0 < z$:*

$$V(f) = \beta Ev(f) + cv(x_n) + cv(x_{-m})$$

In CPT, it is assumed $x_0 = 0$ and $v(0) = 0$, which leads to $v(x_0) = 0$.

Our discussion of the (β, c) model in this paper also invokes this assumption for expository simplicity in applications, although the (β, c) model does not need it.

We have chosen to present the general (β, c) representation without invoking this assumption in theorems, and the representation that does invoke it in corollaries.

3.2 Special Cases of Discrete Lotteries

3.2.1 Pure Gains

This is the special case when $x_i \geq x_0, \forall i$. In this case, $f = f^+$.

Corollary 2 (Pure Gains) *Discrete lottery f with $x_i \geq x_0$ $\forall i$ (Pure-Gain Lottery) is evaluated in the (β, c) model by*

$$V(f) = \begin{cases} \beta Ev(f) + cv(x_n) + (1-\beta-c)v(x_1) & , \text{ if } p_0 = 0 \\ \beta Ev(f) + cv(x_n) & , \text{ if } p_0 > 0 \end{cases} \tag{7}$$

Proof. Note that $x_i \geq x_0$ $\forall i \Longrightarrow p^- = 0 \Longrightarrow p^+ + p_0 = 1 \Longrightarrow$
$f^- = (x_0, 1) \Longrightarrow V(f^-) = v(x_0) = 0$.
Therefore, when $p_0 = 0$:

$$V(f) = V(f^+) = \beta Ev(f) + cv(x_n) + (1-\beta-c)v(x_1)$$

Some points to notice:

- $p_0 = 0$ means x_1 is the "sure" outcome in this lottery, and thus x_1 gets overweighted by $(1-\beta-c)$ due to the certainty effect;
- the extreme outcome x_n still has an additional term $cv(x_n)$ in the representation due to the possibility effect.

When $p_0 > 0$:

$$\begin{aligned} V(f) &= V(f^+) = \beta Ev(f) + cv(x_n) + (1-\beta-c)v(x_0) \\ &= \beta Ev(f) + cv(x_n) \end{aligned}$$

- In the evaluation, x_0 always gets double-counted, but it does not appear in the representation because $v(x_0) = 0$ is invoked.

■

3.2.2 Pure Losses

The special case when $x_i \leq x_0, \forall i$. In this case, $f = f^-$.

Corollary 3 (Pure Losses) *Discrete lottery f with $x_i \leq x_0$ $\forall i$ (Pure-Loss Lottery) is evaluated in the (β, c) model by*

$$V(f) = \begin{cases} \beta Ev(f) + cv(x_{-m}) + (1-\beta-c)v(x_{-1}) & , \text{if } p_0 = 0 \\ \beta Ev(f) + cv(x_{-m}) & , \text{if } p_0 > 0 \end{cases}$$

Proof. $x_i \leq x_0$ $\forall i \Longrightarrow p^+ = 0 \Longrightarrow p^- + p_0 = 1 \Longrightarrow$
$f^+ = (x_0, 1) \Longrightarrow V(f^+) = v(x_0) = 0.$
Therefore, when $p_0 = 0$:

$$V(f) = V(f^-) = \beta Ev(f) + cv(x_{-m}) + (1-\beta-c)v(x_{-1})$$

When $p_0 > 0$:

$$\begin{aligned} V(f) &= V(f^-) = \beta Ev(f) + cv(x_{-m}) + (1-\beta-c)v(x_0) \\ &= \beta Ev(f) + cv(x_{-m}) \end{aligned}$$

■

3.2.3 Binary Lotteries

Corollary 4 (Binary Lottery) *If $f = (x_1, p_1; x_2, p_2)$, where $x_1 < x_2$ and $p_i > 0$ $\forall i$, then f is evaluated in the (β, c) model by*

$$V(f) = \begin{cases} \beta Ev(f) + cv(x_2) + (1-\beta-c)v(x_1) & , \text{if } x_0 \leq x_1 \\ \beta Ev(f) + cv(x_1) + (1-\beta-c)v(x_2) & , \text{if } x_0 \geq x_2 \\ \beta Ev(f) + cv(x_1) + cv(x_2) & , \text{if } x_1 < x_0 < x_2 \end{cases} \tag{8}$$

Proof. When $x_0 \leq x_1$:
This is a pure-gain lottery, and by Corollary 2 we have

$$V(f) = \beta Ev(f) + cv(x_2) + (1-\beta-c)v(x_1)$$

When $x_0 \geq x_2$:
This is a pure-loss lottery, and by Corollary 3 we have

$$V(f) = \beta Ev(f) + cv(x_1) + (1-\beta-c)v(x_2)$$

When $x_1 < x_0 < x_2$:

In this case, $f^+ = (x_0, p_1; x_2, p_2)$, $f^- = (x_1, p_1; x_0, p_2)$, and by (4) of Theorem 1, we immediately have

$$\begin{aligned} V(f) &= \beta Ev(f) + cv(x_1) + cv(x_2) + 2(1-\beta-c)v(x_0) + \beta v(x_0) \\ &= \beta Ev(f) + cv(x_1) + cv(x_2) \end{aligned}$$

■

3.3 Continuous Lotteries

While CPT only considers discrete lotteries, we want our model to also be applicable to continuous lotteries of the following kind.

Definition 2 (Continuous Lottery) *f is called a continuous lottery if it is an **atomless** probability density function defined on a **convex** support.*

We denote by S the support and only discuss the case when $S \subseteq \mathbb{R}$. Denote by $int(S)$ the interior of S. Note that when S is a singleton, the continuous lottery becomes *degenerate*, and its representation is given in Theorem 0.

We continue to denote by $\mathbf{x}$ the underlying continuous random variable of f, and denote by F its c.d.f.

We define the value and weighting functions for f in a very similar way as for discrete lotteries, in the same spirit as Quiggin (1982).[3]

Consider two weighting functions, $w^+(\cdot)$ and $w^-(\cdot)$, which satisfy property (2) and are strictly increasing and differentiable on $(0,1)$, with derivatives denoted $w^{+\prime}(\cdot)$ and $w^{-\prime}(\cdot)$, respectively.

We assume $\lim_{t\to 0+} w^{+\prime}(t)$, $\lim_{t\to 1-} w^{+\prime}(t)$, $\lim_{t\to 0+} w^{-\prime}(t)$, $\lim_{t\to 1-} w^{-\prime}(t)$ all exist, and denote them $w^{+\prime}(0)$, $w^{+\prime}(1)$, $w^{-\prime}(0)$, $w^{-\prime}(1)$, respectively.

Given x_0, for realization $x \in S$ of $\mathbf{x}$, if $x > x_0$, we define its weight as

$$\begin{aligned} \pi^+(x) &\equiv \lim_{\varepsilon\to 0+} \frac{1}{\varepsilon}[w^+(1-F(x)) - w^+(1-F(x+\varepsilon))] \\ &= w^{+\prime}(1-F(x)) \cdot f(x) \end{aligned}$$

If $x < x_0$, we define its weight as

$$\begin{aligned} \pi^-(x) &\equiv \lim_{\varepsilon\to 0+} \frac{1}{\varepsilon}[w^-(F(x+\varepsilon)) - w^-(F(x))] \\ &= w^{-\prime}(F(x)) \cdot f(x) \end{aligned}$$

[3] See Quiggin (1982) for a discussion of the analogy between the weighting functions for discrete and continuous lotteries.

And

$$\begin{aligned}\pi^+(x_0) &\equiv w^+(1) - w^+(1 - F(x_0))\\ \pi^-(x_0) &\equiv w^-(1) - w^-(F(x_0))\end{aligned}$$

Similar to those for discrete lotteries, these weights are also transformed from the "cumulative probability" of outcome x. Since f is a continuous probability distribution, the weights assigned to its outcomes need to have the format of "probability densities", and hence we use the derivatives $w^{+\prime}$ and $w^{-\prime}$ in the definitions above.

The value of lottery f is then defined as

$$V(f) \equiv V(f^+) + V(f^-)$$

where

$$\begin{aligned}V(f^+) &\equiv \int_{x_0}^{\infty} \pi^+(t)v(t)dt + \pi^+(x_0)v(x_0)\\ V(f^-) &\equiv \int_{-\infty}^{x_0} \pi^-(t)v(t)dt + \pi^-(x_0)v(x_0)\end{aligned}$$

where $v(\cdot)$ is exactly the same utility function as in the case of discrete lotteries.

Theorem 2 (Continuous Lottery) *Continuous lottery f with support S is evaluated in the (β, c) model by*

$$V(f) = \begin{cases} \beta Ev(f) + (2 - \beta - 2c)v(x_0) & , \textit{if } x_0 \in int(S) \\ \beta Ev(f) + v(x_0) & , \textit{if } x_0 \notin int(S) \end{cases} \tag{9}$$

Proof. By the definition of continuous lottery f we know that $\forall x \in S$, $0 < f(x) < 1$.

In the (β, c) model, by (3) we have $w^{+\prime}(t) = w^{-\prime}(t) = w'(t) = \beta$, $\forall t \in [0, 1]$ (recall our definition of $w^{+\prime}(0)$, $w^{+\prime}(1)$, $w^{-\prime}(0)$ and $w^{-\prime}(1)$ above). Therefore we have

$$\pi^+(x) = \pi^-(x) = \beta f(x), \forall x \neq x_0$$

If $x_0 \in int(S)$, we know $F(x_0), 1 - F(x_0) \in (0, 1)$. Therefore

$$\begin{aligned}V(f^+) &= \int_{x_0}^{\infty} \beta f(t)v(t)dt + (1 - \beta - c + \beta F(x_0))v(x_0)\\ V(f^-) &= \int_{-\infty}^{x_0} \beta f(t)v(t)dt + (1 - \beta F(x_0) - c)v(x_0)\end{aligned}$$

Therefore $V(f) = \beta Ev(f) + (2 - 2c - \beta)v(x_0)$.

If $x_0 \notin int(S)$, we know $F(x_0) \cdot (1 - F(x_0)) = 0$, and therefore one of $\pi^+(x_0)$ and $\pi^-(x_0)$ must take value 1 while the other takes value 0, which implies $\pi^+(x_0) + \pi^-(x_0) = 1$.

Therefore $V(f) = \beta Ev(f) + v(x_0)$. ■

Interpretation of the (β, c) Representation for Continuous Lotteries For a continuous lottery f, if $x_0 \notin int(S)$, it must be that x_0 is either weakly bigger than or weakly smaller than all the outcomes of f. The atomless assumption guarantees that there exists no extreme outcome with strictly positive probability in f, thus CPT suggests that all probabilities get discounted due to lack of sensitivity (in intermediate probability "densities"), and in the (β, c) model they are discounted by factor $\beta \leq 1$, which results in the linear part of the representation.

Moreover, x_0 is by definition always double counted in our definition of decision weights ($\pi^+(x_0)$ and $\pi^-(x_0)$) above, and hence the additional term $v(x_0)$.

If $x_0 \in int(S)$, there must exist outcomes y and z of f such that $y < x_0 < z$, and hence we must have $\Pr[\mathbf{x} \leq x_0] > 0$ and $\Pr[\mathbf{x} \geq x_0] > 0$. Thus, just like in the discrete case, x_0 becomes the "sure" outcome in both f^+ and f^- (it is either the minimum gain or minimum loss), and is therefore overweighted by $2(1 - \beta - c)$ due to the certainty effect.

Moreover, x_0 is still double counted in the evaluation, although since it now appears in the interior of support S, it will be discounted by β. Hence we have the additional term $(2 - \beta - 2c)v(x_0)$.

Corollary 5 *If $v(x_0) = 0$, the (β, c) model gives the following representation of continuous lottery f:*

$$V(f) = \beta Ev(f) \tag{10}$$

Intuition: In a continuous lottery (satisfying Definition 2), all outcomes have zero probability measure. Since there is no extreme outcome with strictly positive probability, there is no overweighting of such outcomes. Therefore, in the (β, c) model, all outcomes are (under-)weighted by factor β, resulting in a linear representation of the lottery.

3.4 Summary of the (β, c) Representation

As Theorems 1 and 2 show, the (β, c) model results in a preference representation consisting of a linear part that is the discounted expected value of the lottery, $\beta \cdot Ev(f)$, and additional terms of the reference point and extreme outcomes (if any) that exhibit the certainty effect, the possibility effect, and reference dependence. Each part in the representation has a transparent interpretation, and the whole representation is highly intuitive and tractable.

A natural special case that draws our attention is when $c = 0$ and $\beta < 1$, and when the assumption that $v(x_0) = 0$ is invoked. In this case, Corollaries 1 and 5 show that the (β, c) representation of non-degenerate discrete and continuous lotteries is simply the linear discounted expected value of the lottery, $\beta \cdot Ev(f)$. Combined with Theorem 0, the (β, c) representation of lotteries exhibits the same feature as the (β, c) model itself

- it is *linear* and *discontinuous*. The discontinuity in the (β, c) representation occurs as soon as risk is introduced into a degenerate lottery, where the representation value drops from expected value to a strictly smaller discounted expected value, although linearity in probabilities is maintained for all lotteries.

3.5 The (β, c) Model - What It Does and Does Not Do

Tractable Preference Representation Above all, the (β, c) model achieves highly tractable preference representation for general lotteries. The "gist" of CPT, such as the certainty and possibility effects, and reference dependence are all shown in concise and intuitive analytical forms, as discussed in the previous section.

3.5.1 What It Does that CPT also Does

By construction, the (β, c) model can explain all the major phenomena of risky choice violating the standard model that CPT was developed to reconcile (among other things), as set out by Tversky and Kahneman (1992), including[4]:

Framing Effect/Reference Dependence Variations in the framing of options in terms of gains or losses yield systematically different preferences. The (β, c) model keeps the same assumption on reference dependence and the same framing process of gains and losses as CPT.

Nonlinear Preferences/Certainty Effect CPT reconciles nonlinear preferences (e.g. "Allais" behavior) through changes of curvature of the weighting function (steeper at impossibility and certainty, and flatter in the middle); the (β, c) model achieves non-linearity by discontinuities at impossibility and certainty. In the next section we illustrate that a calibrated (β, c) model can explain Allais Paradox.

Risk Seeking Preference for large prizes with small probabilities over its expected value, and for large losses with substantial probabilities over a sure smaller loss. These can be explained in both CPT and the (β, c) model, where decision makers overweight small probabilities and underweights large probabilities.

Loss Aversion Losses loom larger than gains. The asymmetry between gains and losses can be explained in CPT by allowing for a steeper weighting function for losses $(w^-(\cdot))$ than that for gains $(w^+(\cdot))$. Similarly, we can extend the (β, c) model to allow for different

[4]Besides the four phenomena discussed here, Tversky and Kahneman (1992) also mention a fifth pheonomena (that is related to uncertainty/ambiguity) that CPT can explain - Source dependence. The (β, c) model in this paper does not explain source dependence as we have restricted our discussion to risky choice only.

weighting functions, (β^+, c^+) and (β^-, c^-), for gains and for losses respectively, with a corresponding larger β^- than β^+.[5]

3.5.2 What It Does that CPT Does Not Do

Besides working as a relatively nice descriptive model of risky choice, the (β, c) model also does two things that CPT does not do:

"Bayesian" Updating When we use $c = 0$, the (β, c) model allows for normal Bayesian updating of probabilities strictly smaller than 1 in dynamic decision making. This is because, when $c = 0$, the (β, c) weighting function is *linear* in probabilities except for 1. If such probabilities satisfy the Bayes rule, so do their respective decision weights. However, the decision weights under CPT do not satisfy Bayes rule as they are everywhere non-linear in probabilities. In section 5 we provide an application of dynamic decision making that makes use of this feature.

Non-Convergence Due to the *discontinuities* of the (β, c) model at impossibility and certainty, the decision weights do not converge to 0 or 1 when the probabilities do, unlike the decision weights in CPT that do converge due to continuity of its weighting functions. Therefore the (β, c) representation does not converge to the standard expected-utility representation as a lottery converges to a deterministic outcome, which for instance can be seen from the contrast between Theorems 0 and 1. This feature may change standard results that rely on limiting properties of sequences of lotteries, and we provide an application of this feature in section 7 that shows its impact on trembling-hand perfection in strategic games.

3.5.3 What It Does Not Do - Limitations of the (β, c) Model

The benefits we derive from the two key features of the (β, c) model, linearity and discontinuity, come at a cost. The (β, c) model and representation by construction have their limitations.

Loss of continuity at impossibility and certainty The (β, c) weighting function is neither differentiable nor continuous at impossibility (when $c > 0$) and certainty (when $1 - \beta - c > 0$), and therefore it does not allow estimation of sensitivity (i.e. derivative) in probabilities at these two points.

Some may argue that the non-convergence result we discussed previously can also be viewed as a limitation of the model, as the discontinuities of the (β, c) weighting function imply that even the slightest change in probability from total impossibility or certainty will

[5] The detailed discussion of the (β, c) model with two different weighting functions is omitted as loss aversion is not the focus of the current paper.

result in a big "jump" in evaluation. However, some results we derive in applications rely critically on this discontinuity. A detailed assessment of this issue is provided in section 7.3.

Loss of curvature Due to the linearity in intermediate probabilities, the (β, c) weighting function does not have changing curvature and therefore cannot capture any "fine" effects related to curvature, such as changes in a decision maker's sensitivity towards intermediate probabilities at different levels of probabilities. For this reason, it cannot explain nonlinear preferences in choices that do not involve sure outcomes (e.g. common ratio effect with only intermediate probabilities).

Loss of curvature in the weighting function naturally brings the (β, c) representation closer to EU representation compared to the original CPT representation. In many cases this makes the (β, c) representation coincide with EU representation. Some may view this as "too far" a departure from CPT and thus another limitation of the model.

In particular, when the assumption $v(x_0) = 0$ is invoked (as it should be in most situations), i) for all *non-degenerate continuous lotteries* satisfying Definition 2, Corollary 5 tells us that the (β, c) representation is equivalent to EU representation; ii) for all *non-degenerate discrete lotteries with the same extreme (i.e. largest and smallest) outcomes*, Corollary 1 implies that the (β, c) representation is also equivalent to EU representation. In these cases, the only "useful addition" that the (β, c) model brings to the standard EU representation is the discontinuity when evaluating *degenerate lotteries*, as shown by the contrast between Theorem 0 and the corollaries mentioned previously.

In all other cases, particularly for all *non-degenerate discrete lotteries with different extreme outcomes*, the (β, c) representation lies "strictly between" EU and CPT representations.

4 Empirics and Calibration

Weighting functions with a linear part in intermediate probabilities are sometimes used in ad hoc manners in empirical studies to fit experimental data on risky choice. As such data by definition involve discrete rather than continuous lotteries (probability distributions), they can not logically show what the limits of decision weights are as true probability converges to 0 or 1, and thus continuity of the weighting function at impossibility and certainty does not really matter for empirics. However, if a linear weighting function for intermediate probabilities fit experimental data reasonably well, this will provide support for the (linear part of) the (β, c) model as a good descriptive model, as well as providing estimates of the parameters of the (β, c) model.

Tversky and Fox (1995) estimated among other things a linear weighting function for three studies of risky choice, as an approximation of their more general "subadditive"

(SA) weighting functions. They found that the former fits the data quite well. The median estimates in their study of a set of parameters that correspond to β, c, and $(1-\beta-c)$ in the (β, c) model were 0.76, 0.07, and 0.16, respectively.

In a more theoretical study of the SA weighting functions, Tversky and Wakker (1995) argued that *"Since w* (author: w is their SA weighting function) *is fairly linear in the middle region, ... upper SA usually hold for larger intervals, and for most functions found in the literature ... even $\varepsilon' = 0$ can be chosen"*. In their model, ε' represents the lower bound of the range of probabilities where "upper subadditivity" applies, which by and large captures the certainty effect. $\varepsilon' = 0$ corresponds to $c = 0$ in the (β, c) model.

We take these studies as the basis of our calibration of the (β, c) model in the applications we will discuss in subsequent sections. Before that, we fist illustrate below that a reasonably calibrated (β, c) model can explain Allais Paradox.

4.1 Allais Paradox

Consider the original Allais Paradox:

$$\begin{array}{ll} \text{Question 1:} & \begin{cases} \text{Lottery } A & = (\$1\ Million, 100\%) \\ \text{Lottery } A^* & = (\$0, 1\%; \$1\ Million, 89\%; \$5\ Million, 10\%) \end{cases} \\ \text{Question 2:} & \begin{cases} \text{Lottery } B & = (\$0, 89\%; \$1\ Million, 11\%) \\ \text{Lottery } B^* & = (\$0, 90\%; \$5\ Million, 10\%) \end{cases} \end{array}$$

A typical "Allais behavior" involves the choice of A in Question 1 and the choice of B^* in Question 2, which violates EU theory.

Now we show that such behavior can be explained using the (β, c) model with the estimates from Tversky and Fox (1995): $\beta = 0.76$, and $c = 0.07$.

First suppose the value function for money of the decision maker (DM) is $v(\cdot)$ as in CPT, and take \$0 as the reference point, i.e. $v(0) = 0$.

If DM has (β, c) preference, by Corollary 2 her evaluation of the lotteries in Questions 1 and 2 should be:

$$\begin{aligned} V(A) &= v(1M) \\ V(A^*) &= \beta \times (89\% \times v(1M) + 10\% \times v(5M)) + c \times v(5M) \\ &= 0.89\beta \times v(1M) + (0.1\beta + c) \times v(5M) \\ V(B) &= \beta \times (11\% \times v(1M)) + c \times v(1M) \\ &= (0.11\beta + c) \times v(1M) \\ V(B^*) &= \beta \times (10\% \times v(5M)) + c \times v(5M) \\ &= (0.1\beta + c) \times v(5M) \end{aligned}$$

And her "Allais behavior" can be represented by

$$\begin{aligned} V(A) &> V(A^*) \Leftrightarrow \\ v(1M) &> 0.89\beta \times v(1M) + (0.1\beta + c) \times v(5M); \text{ and} \\ V(B) &< V(B^*) \Leftrightarrow \\ (0.11\beta + c) \times v(1M) &< (0.1\beta + c) \times v(5M) \end{aligned}$$

which together are equivalent to

$$\frac{0.1\beta + c}{1 - 0.89\beta} < \frac{v(1M)}{v(5M)} < \frac{0.1\beta + c}{0.11\beta + c}.$$

Since $\beta = 0.76$, and $c = 0.07$, the condition above reduces to

$$0.46 < \frac{v(1M)}{v(5M)} < 0.95.$$

Therefore, a value function in money $v(\cdot)$ that is concave enough will satisfy this condition. For instance, $v(x) = x^\alpha$ with $\alpha = 0.4$ works, where $\frac{v(1M)}{v(5M)} \approx 0.53$. Smaller positive values of α also work.

5 Application I - Dynamic Decision Making

The preference represented by the (β, c) model has a bias towards certainty. We show in this section that in dynamic situations, this bias can lead to distortions in evaluation when the decision maker *Bayesian updates* probabilities, which may result in inconsistent behaviors.

5.1 A Promotion Example

Suppose there is a CD album that a decision maker (DM) values at $\bar{v} = £5$ for sale at a shop. DM demands at most *one* unit of this CD. The current price posted is $p_1 = £3$. The shop has announced that it will run a promotion on this item in two steps: it will first reduce the price to $p_2 = £2$ next week (week 2), and then finally to $p_3 = £1$ in week 3.[6]

Apart from the price changes, suppose DM has no preference towards time - the value of this CD will be £5 to her no matter when she buys it.[7] Her utility from buying the CD

[6] Real-life examples of such promotions are observed around Christmas, where the price of some product may be slightly reduced close to the holiday as a "Christmas promotion", and then further reduced at the "Boxing Day sale". Whether announced or not, such "step promotions" can in many cases be anticipated in advance by consumers.

[7] Although the intention of such "step promotions" may be related to (some) people's time preference (e.g. Christmas gift purchases), all consumers need not have time preference. What we study here is how

at any price p is simply $\bar{v} - p$. Therefore she would generally want to buy at the lowest final price £1.

However, the shop has limited stock of this CD and there are other shoppers who, knowing about the promotion or not, may buy it before DM. Suppose that DM cannot stay at the shop all the time to monitor the stock. Therefore, if she waits for the promotion prices, there is a risk that it will be sold out. She estimates that there is an 80% chance that stock will last till price drops to £2, and a 72% chance that it will last till price drops to £1. In other words, if she waits till the price drops to £2, there is a 20% chance that the CD is sold out when she goes; conditional on there being still stock at £2, it will last till price drops to £1 with probability $\frac{72\%}{80\%} = 90\%$ (i.e. there is a further 10% chance of sell-out if she waits till the price drops to £1).

Suppose not buying will leave DM with a payoff of 0, which we naturally take as the reference point.

5.1.1 The Dynamic Purchasing Decision Problem

What should DM do now? In other words, at what price should she buy this CD?

We can represent this decision problem in "extensive form" in Figure 3, where the solid dots and hollow dots represent the decision nodes of the DM and of "nature", respectively, and the numbers in parentheses represent DM's payoff.[8]

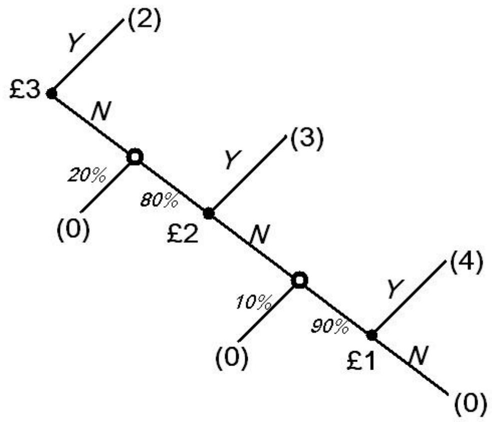


Figure 3: The Dynamic Purchasing Decision Problem

a DM with no time preference should act faced with such promotions, whatever the shop's intention is.

[8] To answer the previous question, we need to compare two lotteries - "buying" and "not buying" - at each of the three prices. Our discussion in the main text focuses on intuition and tries to avoid technical terms and heavy notation. To clarify the correspondence between this example and the terminology of our model, for $i = 1, 2, 3$, "buying" at price p_i is the degenerate lottery: $(\bar{v} - p_i, 1)$. "Not buying" at price p_i is the binary pure-gain lottery: $(0, \pi_i; V_i, 1 - \pi_i)$, where π_i is the sell-out probability at p_i, and V_i is the "continuation value" of the game evaluated at p_i if DM does not buy at p_i. This "continuation value" is DM's *perceived* best alternative option to buying at the current price, which depends on her preference and her options in the remainder of the game. Its particular form becomes clear in our discussion in the main text.

If DM has EU preference, it is straightforward to see that she should wait for the lowest price, because this gives her the highest expected value $(5-1)\times 80\% \times 90\% = 2.88$, whereas buying at £3 gives $5-3=2$ and buying at £2 gives $(5-2)\times 90\% = 2.7$.

5.1.2 (β, c) Preference and Awareness

Suppose instead that DM has (β, c) preference with $\beta = 0.8$ and $c = 0$, then her decision may be different, as her preference suggests that she has a bias towards certainty. Furthermore, her decision may also be affected by whether she is aware of her bias or not. Following the terminology of the self-control literature, we call DM a *naif* if she is oblivious of her certainty bias, and a *sophisticate* if she is fully aware.[9]

Evaluation Distortion in Bayesian Updating In order to clarify the exact meanings of these different cases, sophisticates evaluate lotteries with the (β, c) preference, and know that their bias towards certainty (represented by $\beta < 1$) will cause disproportionate scaling of decision weights in their evaluation of lotteries, particularly when probabilities are scaled up to certainty as compared to scaling up to be merely probable.

For instance, from the perspective at price £3, the probabilities of the stock lasting till prices £2 and £1 are 80% and 72%, respectively, which in the (β, c) model are transformed to decision weights $\beta \cdot 80\%$ and $\beta \cdot 72\%$, respectively. Come price £2, *conditional* on there is still stock, however, the probabilities of stock availability at prices £2 and £1 become $\frac{80\%}{80\%} = 100\%$ and $\frac{72\%}{80\%} = 90\%$, respectively, which are transformed to decision weights 1 (without β) and $\beta \cdot 90\%$, respectively. Therefore, the respective decision weights are not both scaled proportionately when one of the (conditional) probabilities is 1, which in turn leads to evaluation distortion of the outcomes at these prices.

Sophisticates realize that their bias distorts their evaluation when their "Bayesian updating" of probabilities involves sure outcomes. However, naifs incorrectly believe that there is no such distortion in their evaluation when probabilities are scaled up or down, even though they also have (β, c) preference. Therefore, there may be inconsistency between what naifs plan to do and what they end up doing.

5.1.3 Solution

In summary, we want to find the choices of three types of DM: EU, naif and sophisticate. We show that in this example they all indeed make different decisions: EUs buy at the lowest promotion price £1, naifs buy at intermediate price £2, while sophisticates (somewhat surprisingly) buy at the highest price £3.

We use backward induction to analyze the choices of naifs and sophisticates, who share the same (β, c) preference with $\beta = 0.8$ and $c = 0$. Recall from Theorems 0 and 1 that,

[9] DM's decision may also be affected by the degree of her awareness of her bias. Here we only discuss the two extreme cases - total obliviousness and full awareness.

a (β, c) DM will apply multiplier β to the expected utility of any lottery the outcomes of which have probabilities strictly smaller than 1, but will not do so if the lottery has only one sure outcome. Therefore at the lowest price £1, if there is stock, both naifs and sophisticates will buy as this gives them a positive payoff $5-1=4$, higher than the outside option 0. (Recall that EUs will also buy.)

Now consider the situation at the intermediate price £2, where buying would provide payoff $5-2=3$. What is the expected payoff of waiting until the price is £1? Because there is a 10% risk of sell-out, and (β, c) preference underweights any uncertain outcome by $\beta = 0.8$, naifs and sophisticates evaluate the lottery of waiting until £1 to be $0.8 \times (5-1) \times 90\% = 2.88$, which is lower than their payoff from buying at price £2. Thus when the price is £2, both naifs and sophisticates will buy. (Recall that at price £2 EUs still prefer the option of buying at £1 which they evaluate at $(5-1) \times 90\% = 3.6 > 3$.)

Finally, consider the situation at the original price £3, where buying would provide payoff $5-3=2$. What is the payoff from not buying? Sophisticates know that if they do not buy at £3, they will buy at price £2 as long as there is stock. Thus sophisticates only need to know their current evaluation of the lottery of buying at £2, which is $0.8 \times (5-2) \times 80\% = 1.92$, lower than the payoff of buying at £3. Therefore when the price is £3, sophisticates buy right away.

Naifs also get payoff 2 if they buy at £3. What is their payoff from not buying at £3? Recall that they do not realize that their biased preference may lead to evaluation distortions when they scale probabilities. When the price is £3, if you ask naifs at what price they would buy if they did not buy at £3, they will say that they would buy at £1. This is because their current evaluation of buying at £1 is $0.8 \times (5-1) \times 80\% \times 90\% = 2.304$, higher than their current evaluation of buying at £2, which is $0.8 \times (5-2) \times 80\% = 1.92$. Therefore, believing incorrectly that they would buy at £1 and receive a payoff of 2.304, naifs forego the option of buying at £3. However, as our previous analysis shows, naifs will end up buying at £2 come the first price reduction.

We have shown that in this promotion example, EUs wait until the lowest sale price despite the risk of sell-out; (β, c) naifs think they would wait for the lowest price and therefore risk not buying at the original price, but end up buying at the intermediate price because at that price the risk of losing is actually too high for them to take; (β, c) sophisticates, on the other hand, seize the CD at the original price for security, realizing they would otherwise give in to their fear of uncertainty (of sell-out) at the intermediate sale and never reach the final sale.

5.2 Alternative Auction Interpretation

The dynamic decision-making problem in Figure 3 can also be interpreted as that of a bidder in a descending-price auction with a large (perhaps unknown) number of other bidders, such as in a Dutch flower market.

An important feature of the situations that can be described by this problem is that the DM takes the risk in the decision-making process as *given*, by either subjective or objective probability distributions. For example, in a Dutch auction, if the number of bidders is very large and the time for consideration at each price is quite short, it may be a practical "rule of thumb" for a bidder to consider the sell-out probability at any price as given and independent of her own decision, rather than to focus on the interactive aspect of bidding. If a bidder thinks this way, her decision-making process can be described in exactly the same way as illustrated in this promotion example.

5.3 Discussion

5.3.1 Analogy to Time-Inconsistent Behavior

The behavior of (β, c) decision makers in this example resembles *preproperative* behaviors exhibited by quasi-hyperbolic time preference in O'Donoghue and Rabin (1999)'s context of intertemporal choice. We have developed (but have chosen not to present here) examples involving losses that increase in absolute value over the different stages of dynamic decision making where the behavior of (β, c) decision makers resembles *procrastination.* Actually, results exactly parallel to those by O'Donoghue and Rabin (1999) can be developed using the (β, c) model for dynamic decision of risky choice, where decision makers with a bias for certainty exhibit time-inconsistent behavior.

The analogy between some decision patterns in risky choice and intertemporal choice has been investigated empirically by Keren and Roelofsma (1995) and Weber and Chapman (2005), in search for the underlying psychological channels of such analogy, and theoretically by Halevy (2008), who unifies the representation of risk and time preferences and argues that risk works on a more fundamental dimension.

5.3.2 Distorted Bayesian Updating

Unlike these studies, the purpose of this application is to show that the (β, c) model is an easily applicable tool to introduce certainty effect into dynamic decision making. As we have discussed in section 3.5.2, the (β, c) model has a big advantage over CPT in this context as it allows Bayesian updating of probabilities strictly smaller than 1, and creates distortion whenever such updating involves certainty. Though simple, we hope that the promotion example nonetheless illustrates that the (β, c) model is potentially applicable in situations with much more complex information structure and that such "distorted" Bayesian updating can have unexpected implications.

6 Application II - Mixed Strategy Nash Equilibrium in Finite Strategic Games

In the promotion example we discussed previously, the risk is modeled as if it comes from "nature". That is, the DM's choice does not affect the probabilities in the lotteries she faces (i.e. the chances whether or not stock runs out, which is assumed to be given). In this section, we turn to situations where the risk that a DM faces comes from other DMs' choices, and strategic interaction among different DMs affects everyone's decision making process.

Since situations with strategic interaction are generally distinguished from individual decision making, and are modeled by game theory instead of decision theory, in this section we replace the name "DM" with "player". One natural area in game theory to look for the impacts of certainty effect is mixed strategies, as they involve non-deterministic choice of actions and therefore introduces risk in players' payoff. Actually, the sell-out probabilities in the promotion example can be interpreted as a mixed strategy "played" by nature.

In this section, we show that allowing for certainty effect in players' preference can have significant impact on the standard results on mixed strategies of strategic (i.e. normal-form) games (where by definition time no longer plays a role and there are no dynamics at all).

6.1 Definitions

6.1.1 The Game

We consider finite two-player strategic games.[10] Each player $i \in \{1, 2\}$ has a finite *action set* $A_i = \{a_{i1}, a_{i2}, ..., a_{in_i}\}$ where $n_i \in \mathbb{N}$. An action is also called a *pure strategy.* A pure strategy profile of two players is denoted $a = (a_{1j}, a_{2k}) \in A_1 \times A_2$. Given a pure strategy profile a, player i's payoff is defined as $v_i(a)$, where v_i is player i's *preference representation towards deterministic outcomes*, i.e. the value function $v(\cdot)$ in CPT and the (β, c) model we discussed previously.

We denote by $\Gamma = \langle\{1, 2\}, (A_i), (v_i)\rangle$ the finite two-player strategic game we have described above.

6.1.2 (β, c)-Mixed Extension

Now we defined the "mixed extension" of $\Gamma = \langle\{1, 2\}, (A_i), (v_i)\rangle$.

A *mixed strategy* of player i is defined as a probability distribution over A_i, denoted by $\sigma_i = (a_{i1}, p_{i1}; a_{i2}, p_{i2}; ...; a_{in_i}, p_{in_i}) \in \Delta(A_i)$, where $\sum_{j=1}^{n_i} p_{ij} = 1$. We sometimes simplify

[10] The definitions of strategies and equilibria used in this section are standard in game theory. In particular they follow Osborne and Rubinstein (1994). The only change we make is on the preference representation of players.

the notation by writing $\sigma_i = (p_{ij})_{j=1}^{n_i} \in \Delta(A_i)$. In the case that σ_i has $p_{ij} = 1$ for some $a_{ij} \in A_i$, σ_i simply coincides with pure strategy a_{ij}. A *mixed strategy profile* of two players is denoted by $\sigma = (\sigma_1, \sigma_2) \in \Delta(A_1) \times \Delta(A_2)$.

Definition 3 (Lottery Induced by MS Profile) *The lottery induced by mixed strategy profile $\sigma = ((p_{1j})_{j=1}^{n_1}, (p_{2k})_{k=1}^{n_2}) \in \Delta(A_1) \times \Delta(A_2)$ for player $i \in \{1, 2\}$ is the probability distribution that σ implies over player i's payoffs from pure strategy profiles, denoted by*

$$L(\sigma) = (v_i(a_{1j}, a_{2k}), p_{1j} \cdot p_{2k})_{j=1,\ldots,n_1;k=1,\ldots,n_2}. \tag{11}$$

Notice $L(\sigma)$ is a well-defined discrete lottery in the CPT sense because $\sum_{j=1}^{n_1} \sum_{k=1}^{n_2} p_{1j} \cdot p_{2k} = \sum_{j=1}^{n_1} p_{1j} \cdot \sum_{k=1}^{n_2} p_{2k} = \sum_{j=1}^{n_1} p_{1j} \cdot 1 = 1$.

Player i evaluates $L(\sigma)$ using $V_i(\cdot)$, which is her preference representation towards lotteries. $V_i(\cdot)$ can be EU representation, or the cumulative functional $V(\cdot)$ in CPT and the (β, c) model we discussed previously. We denote

$$V_i(\sigma) \equiv V_i(L(\sigma)) = V_i((v_i(a_{1j}, a_{2k}), p_{1j} \cdot p_{2k})_{j=1,\ldots,n_1;k=1,\ldots,n_2})$$

What we have described above is the "mixed extension" of game Γ with preference representation V_i. If the players in Γ have (β, c) preference, V_i will be the (β, c) representation, in which case we call the extension the "(β, c)-mixed extension". We formalize this concept below.

Definition 4 ((β, c)-Mixed Extension) *The (β, c)-mixed extension of game $\Gamma = \langle N, (A_i), (v_i) \rangle$ is the game $\langle N, (\Delta(A_i)), (V_i) \rangle$ where player $i \in N$ has (β_i, c_i) preference with representation V_i.*

Definition 5 *The EU-mixed extension of game $\Gamma = \langle N, (A_i), (v_i) \rangle$ is the game $\langle N, (\Delta(A_i)), (Ev_i) \rangle$ where player $i \in N$ has EU preference with representation Ev_i.*

Note that the (β, c)-mixed extension incorporates the standard notion of "mixed extension" - EU-mixed extension. By Theorems 0 and 1 we know V_i is exactly Ev_i when $(\beta_i, c_i) = (1, 0)$ and $v_i(x_0) = 0$. When all players have EU preference and $v_i(x_0) = 0$ $\forall i \in N$, the (β, c)-mixed extension becomes the standard EU-mixed extension.

6.1.3 Equilibrium

The equilibrium concepts we use in this section are all defined in the standard way on Γ or on its (β, c)-mixed extension. Since Theorem 0 implies that the (β, c) representation always coincides with EU representation for degenerate lotteries, and pure strategies always induces degenerate lotteries, allowing for (β, c) preference should not affect the standard results on *pure strategy Nash equilibria* (PSNE) of finite game Γ at all. Formally, we have

Lemma 1 *(β, c) representation does not affect the set of pure strategy Nash equilibria of game $\Gamma = \langle \{1,2\}, (A_i), (v_i) \rangle$.*

Proof. Consider a general pure strategy profile of game Γ, $a = (a_{1j}, a_{2k})$ for some $a_{1j} \in A_1$ and $a_{2k} \in A_2$. Suppose player $i \in \{1,2\}$ has (β_i, c_i) preference. By (11) we know a induces degenerate lottery $L(a) = (v_i(a_{1j}, a_{2k}), p = 1)$. By Theorem 0 we know $V_i(a) = v_i(a_{1j}, a_{2k})$, irrespective of β_i and c_i. ■

Thus, we omit the discussion of PSNE, and focus on mixed strategies.

Definition 6 (MSNE) *A **mixed strategy Nash equilibrium** (MSNE) of a game $\Gamma = \langle N, (A_i), (v_i) \rangle$ is the Nash equilibrium of its (β, c)-mixed extension $\langle N, (\Delta(A_i)), (V_i) \rangle$, which is a mixed strategy profile $\sigma^* \in \times_{i \in N} \Delta(A_i)$ such that for every player $i \in N$, we have*

$$V_i(\sigma_i^*, \sigma_{-i}^*) \geq V_i(\sigma_i, \sigma_{-i}^*) \text{ for all } \sigma_i \in \Delta(A_i).$$

Lemma 2 (Indifference at MSNE) *Let game $\Gamma = \langle \{1,2\}, (A_i), (v_i) \rangle$ be a finite strategic game. When every player $i \in \{1,2\}$ has $(\beta_i, c_i = 0)$ preference, suppose $\sigma^* \in \times_{i \in \{1,2\}} \Delta(A_i)$ is an MSNE of Γ, then given σ_{-i}^*, every player $i \in \{1,2\}$ is **indifferent** among all the pure strategies in the support of σ_i^*.*

Proof. Denote by S_i the support of σ_i^*. The Lemma holds trivially in the case that S_i is a singleton for every $i \in \{1,2\}$.

In the case that S_i is not a singleton for some player $i \in \{1,2\}$, without loss of generality suppose $a_{i1}, a_{i2} \in S_i$, that is, if we write out $\sigma^* = ((p_{ij})_{j=1}^{n_i}, (p_{-ik})_{k=1}^{n_{-i}})$, we know $p_{i1} > 0$ and $p_{i2} > 0$.

Suppose that, given σ_{-i}^*, player i is not indifferent between a_{i1} and a_{i2}, and assume without loss of generality that

$$V_i(a_{i1}, \sigma_{-i}^*) > V_i(a_{i2}, \sigma_{-i}^*) \tag{12}$$

Now for $j = 1, 2$, we look at profile (a_{ij}, σ_{-i}^*). The lottery it induces for player i is $L((a_{ij}, \sigma_{-i}^*)) = (v_i(a_{ij}, a_{-ik}), p_{-ik})_{k=1,\ldots,n_2}$. By Theorem 1 we know that player i's evaluation of profile (a_{ij}, σ_{-i}^*) and its induced lottery $L(a_{ij}, \sigma_{-i}^*)$ is for $j = 1, 2$,

$$V_i(a_{ij}, \sigma_{-i}^*) = \begin{cases} Ev_i(a_{ij}, \sigma_{-i}^*) & \text{, if } v_i(a_{ij}, a_{-ik}) = \textit{constant} \text{ for all } a_{-ik} \in S_{-i}; \\ \beta_i \cdot Ev_i(a_{ij}, \sigma_{-i}^*) & \text{, otherwise;} \end{cases} \tag{13}$$

where $0 < \beta_i \leq 1$. The reason for this is that the only situation when profile (a_{ij}, σ_{-i}^*) will yield player i a constant payoff is when $v_i(a_{ij}, a_{-ik}) = \textit{constant}$ for all $a_{-ik} \in S_{-i}$.

(13) distinguishes three situations:

(1) $v_i(a_{i1}, a_{-ik}) = \textit{constant}$, $\forall\ a_{-ik} \in S_{-i}$; but $v_i(a_{i2}, a_{-ik}) \neq \textit{constant}$, $\forall\ a_{-ik} \in S_{-i}$. In this case, $V_i(a_{i1}, \sigma_{-i}^*) = Ev_i(a_{i1}, \sigma_{-i}^*)$, and $V_i(a_{i2}, \sigma_{-i}^*) = \beta_i \cdot Ev_i(a_{i2}, \sigma_{-i}^*)$. Therefore

(12) implies that

$$Ev_i(a_{i1}, \sigma^*_{-i}) > \beta_i \cdot Ev_i(a_{i2}, \sigma^*_{-i})$$

In this case, given σ^*_{-i}, no matter how we mix a_{i1} and a_{i2} (with strictly positive probabilities in order to keep them in S_i), the resulting mixed strategy σ^*_i will yield *strictly lower* payoff for player i than (a_{i1}, σ^*_{-i}) does. This is because when we mix a_{i1} and a_{i2} with strictly positive probabilities, we get a payoff that is a mixture between $\beta_i \cdot Ev_i(a_{i1}, \sigma^*_{-i})$ and $\beta_i \cdot Ev_i(a_{i2}, \sigma^*_{-i})$ (note it is not between $Ev_i(a_{i1}, \sigma^*_{-i})$ and $\beta_i \cdot Ev_i(a_{i2}, \sigma^*_{-i})$), since the certainty of payoff $v_i(a_{i1}, \sigma^*_{-i})$ disappears once a_{i2} is played with positive probability (due to $v_i(a_{i2}, a_{-ik}) \neq constant$, $\forall\ a_{-ik} \in S_{-i}$). That is, we have

$$V_i(a_{i1}, \sigma^*_{-i}) = Ev_i(a_{i1}, \sigma^*_{-i}) > \beta_i \cdot Ev_i(\sigma^*_i, \sigma^*_{-i}) = V_i(\sigma^*_i, \sigma^*_{-i})$$

which contradicts that $(\sigma^*_i, \sigma^*_{-i})$ is an MSNE. Therefore case (1) cannot really happen.

(2) $v_i(a_{ij}, a_{-ik}) = constant$, $\forall\ a_{-ik} \in S_{-i}$, for both $j = 1$ and $j = 2$, or $v_i(a_{ij}, a_{-ik}) \neq constant$, $\forall\ a_{-ik} \in S_{-i}$, for both $j = 1$ and $j = 2$. In this case, (12) implies that

$$Ev_i(a_{i1}, \sigma^*_{-i}) > Ev_i(a_{i2}, \sigma^*_{-i}). \tag{14}$$

(3) $v_i(a_{i2}, a_{-ik}) = constant$, $\forall\ a_{-ik} \in S_{-i}$; but $v_i(a_{i1}, a_{-ik}) \neq constant$, $\forall\ a_{-ik} \in S_{-i}$. In this case, $V_i(a_{i1}, \sigma^*_{-i}) = \beta_i \cdot Ev_i(a_{i1}, \sigma^*_{-i})$, and $V_i(a_{i2}, \sigma^*_{-i}) = Ev_i(a_{i2}, \sigma^*_{-i})$. Therefore (12) implies that

$$\beta_i \cdot Ev_i(a_{i1}, \sigma^*_{-i}) > Ev_i(a_{i2}, \sigma^*_{-i})$$

which in turn implies (14) since $\beta_i \leq 1$.

In both cases (2) and (3), we have (14). Now we come back to look at $\sigma^* = ((p_{ij})_{j=1}^{n_i}, (p_{-ik})_{k=1}^{n_{-i}})$, where $p_{i1} > 0$ and $p_{i2} > 0$. The lottery it induces for player i is $L(\sigma^*) = (v_i(a_{ij}, a_{-ik}), p_{ij} \cdot p_{-ik})_{j=1,\ldots,n_1;k=1,\ldots,n_2}$. Therefore, we have

$$V_i(\sigma^*_i, \sigma^*_{-i}) = \begin{cases} Ev_i(\sigma^*_i, \sigma^*_{-i}) & \text{, if } v_i(a_{ij}, a_{-ik}) = constant \text{ for all } (a_{ij}, a_{-ik}) \in S_i \times S_{-i}; \\ \beta_i \cdot Ev_i(\sigma^*_i, \sigma^*_{-i}) & \text{, otherwise;} \end{cases}$$

By (14) we know that it cannot be that $v_i(a_{ij}, a_{-ik}) = constant$ for all $(a_{ij}, a_{-ik}) \in S_i \times S_{-i}$, therefore we have

$$V_i(\sigma^*_i, \sigma^*_{-i}) = \beta_i \cdot Ev_i(\sigma^*_i, \sigma^*_{-i}) = \beta_i \cdot \sum_{j, a_{ij} \in S_i} [p_{ij} \cdot Ev_i(a_{ij}, \sigma^*_{-i})] \tag{15}$$

where $p_{i1} > 0$ and $p_{i2} > 0$.

Therefore by (14) and (15) we know that in both cases (2) and (3), player i can strictly increase $V_i(\sigma^*_i, \sigma^*_{-i})$ by moving some probability from a_{i2} to a_{i1}, without changing any other probabilities in σ^*_i.

This again contradicts the fact that σ^* is an MSNE of Γ. Therefore, given σ^*_{-i}, every player $i \in N$ must be indifferent among all the pure strategies in the support of σ^*_i. ■

Comment: Lemma 2 is due to the linearity of the (β, c) model in probabilities strictly between 0 and 1. It confirms the same result for finite strategic games with standard EU-mixed extensions.

The discontinuities of players' (β, c) representation at impossibility and certainty do not affect indifference at the MSNE because of the MSNE requirement, which is actually a "best response" requirement. If some player does play two or more actions with positive probabilities at MSNE, it means that no single action alone can do better given the opponent's strategy, despite the "bias" she has towards single actions (that yield a constant payoff).

On the other hand, the reverse of Lemma 2 does *not* necessarily hold, exactly due to the discontinuities of players' (β, c) representation at impossibility and certainty. If given the opponent's strategy, a player is indifferent between two pure strategies which are both best responses, mixing them up does not necessarily create another best response, since it could be that one of the original pure strategies yields a constant payoff while the other does not, in which case a mixed strategy between them will destroy the certainty of the constant payoff and result in its underweighting, which in turn leads to a lower overall payoff for the mixed strategy.

Lemma 2 allows us to use the standard "trick" to look for Nash equilibria in the (β, c)-mixed extensions of two-player finite strategic games - identifying the indifference condition of each player.

6.2 Example

Now we use Lemma 2 to find Nash equilibrium of Game A below, which is a two-player game where each player has two actions - (row) player 1 has actions a_{11} and a_{12}, and (column) player 2 has actions a_{21} and a_{22}.

Game A	a_{21}	a_{22}
a_{11}	$\underline{5}$, 0	0, $\underline{4}$
a_{12}	4, $\underline{4}$	$\underline{4}$, 0

All the payoffs shown in the game matrix are the *evaluation results* for pure strategy profiles, according to the preference representation of the relevant player. In particular, the pair of payoffs in row j and column k is defined as $(v_1(a_{1j}, a_{2k}), v_2(a_{1j}, a_{2k}))$.

Notice that player 1 has an action a_{12} that gives her a constant payoff of 4, irrespective of player 2's strategy. Player 2 has no such action.

Game A has been constructed so that it has no PSNE (when both players have EU preference). And by Lemma 1 we know no player's (β, c) preference will change this. We study its MSNE in three cases, and show that in some cases Game A has no MSNE either.

6.2.1 EU Preference

First consider the case when both players are EUs. That is, $V_i(\sigma) = Ev(\sigma)$ for all $\sigma \in \Delta(A_1) \times \Delta(A_2)$, $i = 1, 2$. In this case, the standard results for (EU) mixed strategy apply and it is straightforward to see that Game A has **one** MSNE, $\sigma^{EU} = ((\frac{1}{2}, \frac{1}{2}), (\frac{4}{5}, \frac{1}{5}))$.

6.2.2 Player 1 with (β, c) Preference

Now suppose player 1 has (β, c) preference and layer 2 still has EU.

Claim 1 *Game A has no MSNE when player 1 has (β, c) preference with $\beta \leq 0.8$ and $c = 0$, and player 2 has EU.*

Proof. Suppose on the contrary that there exists an MSNE $\sigma^* = (\sigma_1^*, \sigma_2^*)$, where $\sigma_i^* = (p_i, 1 - p_i)$, $p_i \in [0, 1]$.

First observe that Game A has no PSNE.

Second observe that σ_2^* cannot be a pure strategy, otherwise player 1 would strict prefer a_{11} (if $\sigma_2^* = a_{21}$) or a_{12} (if $\sigma_2^* = a_{22}$), and σ^* would be a PSNE, which does not exist.

Therefore σ_2^* must assign positive probabilities to both a_{21} and a_{22}, i.e. $p_2 \in (0, 1)$.

By Lemma 2 we know $\sigma_1^* = (p_1, 1 - p_1)$ must make player 2 (who is EU) indifferent between actions a_{21} and a_{22}, that is

$$\begin{aligned} V_2(\sigma_1^*, a_{21}) &= V_2(\sigma_1^*, a_{22}) \Leftrightarrow \\ V_2(0, p_1; 4, (1 - p_1)) &= V_2(4, p_1; 0, (1 - p_1)) \Leftrightarrow \\ 4(1 - p_1) &= 4p_1 \Leftrightarrow \\ p_1 &= \frac{1}{2} \end{aligned}$$

Therefore $\sigma_1^* = (\frac{1}{2}, \frac{1}{2})$, which means that σ_1^* is not a pure strategy either.

Therefore by Lemma 2 we know $\sigma_2^* = (p_2, 1 - p_2)$, $p_2 \in (0, 1)$ must make player 1 (who has $\beta = 0.8$ and $c = 0$) indifferent between actions a_{11} and a_{12} (since both are assigned positive probabilities in σ_1^*). Notice that $v_1(a_{12}, a_{21}) = v_1(a_{12}, a_{22}) = 4$. Therefore

$$\begin{aligned} V_1(a_{11}, \sigma_2^*) &= V_1(a_{12}, \sigma_2^*) \Leftrightarrow \\ V_1(5, p_2; 0, (1 - p_2)) &= V_1(4, p = 1) \end{aligned}$$

By Theorems 0 and 1 we have

$$\begin{aligned}\beta(5p_2) &= 4 \Leftrightarrow \\ p_2 &= \frac{4}{5\beta} \geq 1 \text{ when } \beta \leq 0.8.\end{aligned}$$

This contradicts that $p_2 \in (0,1)$. Therefore there exists no mixed strategy of player 2 that makes player 1 indifferent between actions a_{11} and a_{12}, and therefore Game A has no MSNE. ■

This means that Game A has **no** Nash equilibrium in either pure or mixed strategy in this case.

6.2.3 Both with (β, c) Preference

Our discussion of the previous case also applies when both players have (β, c) preference.

Claim 2 *Game A has no MSNE when player 1 has (β_1, c_1) preference with $\beta_1 \leq 0.8$ and $c_1 = 0$, and player 2 has any (β_2, c_2) preference with $c_2 = 0$.*

The reason for this result is that, unlike player 1, player 2 has no action that gives her a certain payoff level. Therefore, whenever player 1 plays a completely mixed strategy, the strategy profile will induce a non-degenerate lottery for player 2, no matter which action player 2 chooses. Since the MSNE condition of player 2 is an indifference equation between her two actions (each combined with player 1's mixed strategy), player 2 is always comparing two non-degenerate lotteries, whose evaluation will both be affected by the certainty effect implied by the (β_2, c_2) preference, that is, both will be "discounted" by multiplier β_2 (when $c_2 = 0$). But multiplying any $\beta_2 > 0$ on both sides does not change the condition at all. Therefore the MSNE condition for player 2 remains the same whether she has EU or (β, c) preference.

Proof. Compared to the proof of the previous Claim, the fact that player 2 has (β_2, c_2) preference with $c_2 = 0$ instead of EU only changes her indifference condition from $4(1 - p_1) = 4p_1$ to an equivalent condition $4\beta_2(1 - p_1) = 4\beta_2 p_1$. And all the rest of the proof remains the same. ■

6.3 Existence of Mixed Strategy Nash Equilibrium (MSNE)

We provide more general results in this section.

Consider the two-player finite game $\Gamma = \langle \{1, 2\}, (A_i), (v_i) \rangle$.

For $i \in \{1, 2\}$, $a_i \in A_i$, define

$$\begin{aligned} \bar{v}_i(a_i) &\equiv \max_{a_{-i} \in A_{-i}} (v_i(a_i, a_{-i})) \\ \underline{v}_i(a_i) &\equiv \min_{a_{-i} \in A_{-i}} (v_i(a_i, a_{-i})) \\ \Delta v_i(a_i) &\equiv \bar{v}_i(a_i) - \underline{v}_i(a_i) \end{aligned}$$

When $\Delta v_i(a_i) = 0$, we denote $v_i(a_i) \equiv v_i(a_i, a_{-i})$ (which by definition remains constant for any $a_{-i} \in A_{-i}$).

Proposition 1 *In game* $\Gamma = \langle \{1, 2\}, (A_i), (v_i) \rangle$*, suppose*

(i) every player $i \in N$ *has* $(\beta_i, c_i = 0)$ *preference and 0 is the reference point;*

(ii) $A_2 = \{a_{21}, a_{22}\}$*;*

(iii) $v_i(a_i, a_{-i}) \geq 0$*, for any* $i \in \{1, 2\}$*, any* $a_i \in A_i$*, and any* $a_{-i} \in A_{-i}$*;*

(iv) $v_i(a_i, a_{-i}) \neq v_i(a_i, a'_{-i})$ *for any* $a'_{-i} \neq a_{-i}$ *except for* $i = 1$ *and* $a_i = a_{11}$*;*

(v) $v_1(a_{11}, a_{21}) = v_1(a_{11}, a_{22}) = v_1(a_{11}) < \max_{a_1 \in A_1} (\bar{v}_1(a_1))$*, and* $v_2(a_{11}, a_{21}) \neq v_2(a_{11}, a_{22})$*;*

(vi) game Γ *has no PSNE.*

Then Γ*'s* (β, c)*-mixed extension* $\langle \{1, 2\}, (\Delta(A_i)), (V_i) \rangle$ *has **no** Nash equilibrium (or* Γ *has **no** MSNE) if*

$$\beta_1 \leq \frac{v_1(a_{11})}{\max_{a_1 \in A_1} (\bar{v}_1(a_1))}.$$

Proof. Suppose on the contrary that game Γ has an MSNE $\sigma^* = (\sigma_1^*, \sigma_2^*)$, where $\sigma_i^* \in \Delta(A_i)$. We consider three cases.

(1) Suppose for each player $i \in \{1, 2\}$, the support of σ_i^*, S_i, has at least two actions. Denote $\sigma_i^* = (p_{ij})_{j=1}^{n_i}$ where $p_{ij} > 0$, for $i \in \{1, 2\}$ and $j = 1, ..., n_i$. Then $n_i \geq 2$, for $i \in \{1, 2\}$.

Assumption (iv) and (v) implies that player 1 has only one action, a_{11}, that yields her a constant payoff, $v_1(a_{11})$, irrespective of the opponent's action; and player 2 has no action that yields her a constant payoff. Therefore by Definition 3, (a_{11}, σ_2^*) induces a degenerate lottery $L(a_{11}, \sigma_2^*) = (v_1(a_{11}), p = 1)$ for player 1. By Theorem 0 we know player 1 evaluates $L(a_{11}, \sigma_2^*)$ with $V_1(a_{11}, \sigma_2^*) = v_1(a_{11})$.

Since S_1 has at least two actions, suppose without loss of generality that $a_{1j} \in S_1$, for some $j \neq 1$. By Definition 3, (a_{1j}, σ_2^*) induces the following lottery for player 1:

$$L(a_{1j}, \sigma_2^*) = (v_i(a_{1j}, a_{2k}), p_{2k} > 0)_{a_{2k} \in S_2}$$

By assumption (iv), a_{1j} does not yield player 1 a constant payoff. Since player 1 has

$(\beta_1, 0)$ preference, by Theorem 1 we know player 1 evaluates $L(a_{1j}, \sigma_2^*)$ with

$$\begin{aligned} V_1(a_{1j}, \sigma_2^*) &= \beta_1 \cdot Ev_1(a_{1j}, \sigma_2^*) = \beta_1 \cdot \sum\nolimits_{k, a_{2k} \in S_2} v_1(a_{1j}, a_{2k}) \cdot p_{2k} \\ &< \beta_1 \cdot \bar{v}_1(a_{1j}) \\ &\leq \beta_1 \cdot \max_{a_1 \in A_1} (\bar{v}_1(a_1)) \\ &\leq \frac{v_1(a_{11})}{\max\limits_{a_1 \in A_1} (\bar{v}_1(a_1))} \cdot \max_{a_1 \in A_1} (\bar{v}_1(a_1)) \\ &= v_1(a_{11}) = V_1(a_{11}, \sigma_2^*) \end{aligned} \tag{16}$$

Notice that a_{1j} can be any action in S_1 as long as $a_{1j} \neq a_{11}$. By Lemma 2, we know $V_1(a_{1j}, \sigma_2^*) = V_1(a_{1h}, \sigma_2^*)$, $\forall a_{1h} \in S_1$. Therefore we have

$$V_1(a_{11}, \sigma_2^*) > V_1(a_{1h}, \sigma_2^*), \forall a_{1h} \in S_1$$

Lastly, since (σ_1^*, σ_2^*) induces a lottery over all $V_1(a_{1h}, \sigma_2^*)$ for $a_{1h} \in S_1$, and all $V_1(a_{1h}, \sigma_2^*)$ are equal for $a_{1h} \in S_1$, by Theorem 0 we have

$$V_1(\sigma_1^*, \sigma_2^*) = V_1(a_{1h}, \sigma_2^*), \forall a_{1h} \in S_1$$

which implies that

$$V_1(a_{11}, \sigma_2^*) > V_1(\sigma_i^*, \sigma_2^*).$$

which contradicts the fact that (σ_1^*, σ_2^*) is an MSNE of Γ.

(2) Suppose $S_1 = \{a_{1j}\} \subset A_1$, and $S_2 = A_2 = \{a_{21}, a_{22}\}$.

In this case, by the same comparison as in (16), we have $S_1 = \{a_{11}\}$. By Lemma 2, we know a_{11} makes player 2 indifferent among all actions in S_2, but this contradicts assumption (v), $v_2(a_{11}, a_{21}) \neq v_2(a_{11}, a_{22})$.

(3) Suppose S_1 has at least two actions, and $S_2 = \{a_{2k}\} \subset A_2$. By Lemma 2, we know a_{2k} makes player 1 indifferent among all actions in S_1, which implies that $v_1(a_{1j}, a_{2k}) \equiv \bar{v}$, for all $a_{1j} \in S_1$. Therefore (σ_1^*, a_{2k}) induces a degenerate lottery $(\bar{v}, p = 1)$ for player 1, and we have $V_1(\sigma_1^*, a_{2k}) = \bar{v}$.

Since (σ_1^*, a_{2k}) is an MSNE, we have

$$V_1(\sigma_1^*, a_{2k}) \geq V_1(\sigma_1, a_{2k}), \forall \sigma_1 \in \Delta(A_1) \Rightarrow$$

$$v_1(a_{1j}, a_{2k}) \geq V_1(\sigma_1, a_{2k}), \forall a_{1j} \in S_1, \forall \sigma_1 \in \Delta(A_1) \tag{17}$$

By (17), we know any $a_{1j} \in S_1$ is player 1's best response to a_{2k}. Since game Γ has no PSNE (assumption (vi)), a_{2k} cannot be player 2's best response to any $a_{1j} \in S_1$. Since A_2 has only two actions, then the other action in A_2, a_{2l} must be player 2's best response to all $a_{1j} \in S_1$. That is,

$$v_2(a_{1j}, a_{2l}) > v_2(a_{1j}, a_{2k}), a_{2l} \neq a_{2k}, \forall a_{1j} \in S_1 \tag{18}$$

By Definition 3, (σ_1^*, a_{2k}) induces the following lottery for player 2:

$$L(\sigma_1^*, a_{2k}) = (v_2(a_{1j}, a_{2k}), p_{1j} > 0)_{a_{1j} \in S_1}$$

and (σ_1^*, a_{2l}) induces the following lottery for player 2:

$$L(\sigma_1^*, a_{2l}) = (v_2(a_{1j}, a_{2l}), p_{1j} > 0)_{a_{1j} \in S_1}$$

By assumption (iv), neither $L(\sigma_1^*, a_{2k})$ nor $L(\sigma_1^*, a_{2l})$ is a degenerate lottery. Therefore by (18) we have

$$\begin{aligned} Ev_2(\sigma_1^*, a_{2l}) &> Ev_2(\sigma_1^*, a_{2k}) \Rightarrow \\ V_2(\sigma_1^*, a_{2l}) &> V_2(\sigma_1^*, a_{2k}) \end{aligned}$$

which contradicts that (σ_1^*, a_{2k}) is an MSNE.

We have shown that in any of the three possible cases, there is a contradiction. Therefore Γ has no MSNE.

Recall that Game A satisfies all assumptions (i) through (vi), and we have shown that it has an MSNE when both players are EUs, but not when player 1 has (β, c) preference with $\beta \leq 0.8$ and $c = 0$. ■

6.4 Summary

In finite two-player strategic games, we have the following general results about the impact of certainty effect (CE, represented by the $(\beta, c = 0)$ preference) on equilibrium:

i) CE does not affect the existence of pure strategy Nash equilibrium (PSNE);

ii) CE does not affect the existence of mixed strategy Nash equilibrium (MSNE), if no (β, c)-player has a strategy that yields her a constant payoff, irrespective of the opponent's strategies;

iii) CE may affect the existence of MSNE if some (β, c)-player has a strategy that yields her a constant payoff. If this is the case, then for β small enough, there exists no MSNE.

A crucial reason for result iii) is that the payoff functional V of a (β, c) player ($\beta < 1$) is not continuous in mixed strategies, due to discontinuity of the (β, c) model at certainty.

The general existence of MSNE in finite strategic games is a very strong result in standard game theory. We have shown that, with strong CE (represented by small β), this result can be weakened in two-player finite games. We have shown an example, Game

A, which has no Nash equilibrium in either pure or mixed strategy when player 1 has (β, c) preference with $\beta \leq 0.8$ and $c = 0$. Proposition 1 provides a more general result on non-existence of MSNE.

7 Application III - Trembling-Hand Perfection in Finite Strategic Games

Another part of game theory where mixed strategies are most useful is trembling-hand perfection, which is motivated as a refinement of Nash equilibrium where players' rationality with respect to out-of-equilibrium events are treated as "the result of each player's taking into account that the other players could make uncorrelated mistakes (their hands may tremble) that lead to these unexpected events. The basic idea is that each player's actions be optimal not only given his equilibrium beliefs but also given a perturbed belief that allows for the possibility of slight mistakes. These mistakes are not modeled as part of the description of the game. Rather, a strategy profile is defined to be stable if it satisfies sequential rationality given some beliefs that are generated by a strategy profile that is a perturbation of the equilibrium strategy profile, embodying 'small' mistakes (Osborne and Rubinstein, 1994)."

In a finite strategic game, a **trembling hand perfect equilibrium** (THPE) is defined as a mixed strategy profile σ with the property that there exists a sequence of completely mixed strategy profiles $(\sigma^k)_{k=0}^{\infty}$ that converges to σ such that for each player i the strategy σ_i is a best response to σ_{-i}^k for all values of k.

Since we have discussed the existence of MSNE previously, in this section we focus our attention on trembling-hand perfection of PSNE only, which will allow us to keep a "cleaner" example which suffices to show the point.

7.1 Example

Trembling-hand perfection eliminates Nash equilibria with weakly dominated strategies. This can be illustrated by the following Game B of two-players.

Game B	L	R
U	$\underline{5}, \underline{2}$	1, 0
D	$\underline{5}, 0$	$\underline{4}, \underline{2}$

7.1.1 EU Preference

When both players have EU preference, Game B has two PSNE $(U, L), (D, R)$, and only (D, R) is trembling-hand perfect. The reason is that action U of player 1 is weakly dominated by action D, and therefore U cannot be the best response to any completely mixed strategy of player 2.

Now we look at the following Game C, which only reduces player 1's payoff of (D, L) from 5 in Game B to 4 here.

Game C	L	R
U	$\underline{5}, \underline{2}$	1, 0
D	4, 0	$\underline{4}, \underline{2}$

When both players have EU preference, Game C also has two PSNE $(U, L), (D, R)$.

Is either of them trembling-hand perfect? *Yes, both are*!

As action D no longer dominates U in Game C (nor vice versa), both PSNE are trembling-hand perfect.

However, if we interpret all the payoffs in this game as the players' evaluation results of deterministic money amounts in terms of millions of dollars (e.g. take $v(x) = x$ as the value function in money and let the numbers in the game matrix represent millions of dollars), self inspection as well as a potential analogy to the Allais Paradox both suggest that, strategy D can be more appealing to player 1 than strategy U and therefore the equilibrium (D, R) may be more likely to happen than (U, L).

One plausible argument of reasoning by player 1 is: D is a "safe strategy" as it gives a high payoff of 4 for sure, no matter what player 2 plays, whereas although U may possibly give a higher payoff if player 2 plays L, there is always risk that player 2 may play R instead, at least by mistake, which leads to a payoff too low for player 1.

We may hope that trembling-hand perfection, as a refinement of Nash equilibrium that is designed to incorporate players' concerns of "mistakes" into equilibrium, can eliminate the "unfavorable" equilibrium (U, L). However, because neither action of either player dominates the other, the concept of trembling-hand perfection alone cannot achieve this goal.

Now we show that by adding into the model another crucial factor from the argument of player 1 mentioned above - the bias towards certainty - trembling-hand perfection can indeed eliminate the unfavorable equilibrium (U, L).

7.1.2 Player 1 with (β, c) Preference

Now suppose player 1 has (β, c) preference with $\beta = 0.8$ and $c = 0$. Player 2 still has EU.

In order to show (U, L) is no longer trembling-hand perfect, we consider any completely mixed strategy of player 2, $\sigma_2^p = (L, p; R, 1-p)$, $p \in (0, 1)$. Now we need to show that U is never a best response to any sequence of σ_2^p that converges to L.

When player 1 plays U, the strategy profile is (U, σ_2^p), by Definition 3 in the previous section we know her payoff is

$$\begin{aligned}V_1((U,\sigma_2^p)) &= \beta\times(v_1(U,L)\times p + v_1(U,R)\times(1-p))\\ &= \beta\times(5p+1-p) = \beta\times(4p+1)\end{aligned}$$

When player 1 plays D, the strategy profile is (D,σ_2^p), and her payoff is

$$V_1((D,\sigma_2^p)) = 4$$

as $v_1(D,L) = v_1(D,R) = 4$.

Since $\beta = 0.8$, we know that for all $p < 1$

$$V_1((U,\sigma_2^p)) < V_1((D,\sigma_2^p))$$

That is, U is never a best response to any completely mixed strategy σ_2^p, and hence it is never a best response to any sequence of σ_2^p that converges to L, either. Therefore we have the following result.

Claim 3 *In Game C, (U,L) is **not** trembling-hand perfect when player 1 has (β,c) preference with $\beta = 0.8$ and $c = 0$, and player 2 has EU.*

Now we check whether (D,R) is still trembling-hand perfect. The argument above already shows that D is always a best response of player 1 to any completely mixed strategy of player 2. It remains to be shown that R is a best response of player 2 to some sequence of completely mixed strategy of player 1, $\sigma_1^q = (U,q;D,1-q)$, $q\in(0,1)$ that converges to U. To see this, we have

$$\begin{aligned}V_2((\sigma_1^q,R)) &= Ev_2((\sigma_1^q,R))\\ &= v_2(U,R)\times q + v_2(D,R)\times(1-q)\\ &= 2(1-q)\end{aligned}$$

and

$$\begin{aligned}V_2((\sigma_1^q,L)) &= Ev_2((\sigma_1^q,L))\\ &= v_2(U,L)\times q + v_2(D,L)\times(1-q)\\ &= 2q\end{aligned}$$

Therefore for all $q > \frac{1}{2}$

$$V_2((\sigma_1^q,R)) > V_2((\sigma_1^q,L))$$

which means that R is a best response to the sequence of completely mixed strategies $(\sigma_1^q)_q$ for all $q > \frac{1}{2}$, which converges to U as $q\to 1$.

Therefore (D, R) is still trembling-hand perfect.

Trembling-hand perfection has successfully eliminated only the unfavorable equilibrium (U, L) when we allow player 1 to have (β, c) preference with $\beta = 0.8$ and $c = 0$.

7.1.3 Both with (β, c) Preference

It is easy to see that our discussion of the previous case also applies here and therefore the result is the same as before.

The reason for this result is again that only player 1 has an action (D) that gives her a constant payoff level, irrespective of the opponent's action. Player 2 always faces risky payoffs whenever player 1 plays a completely mixed strategy, and therefore player 2's evaluation of the payoff from playing either of her actions will be affected by the certainty effect implied by (β, c) preference, that is, "discounted" by multiplier β. Since multiplying $\beta = 0.8$ on the payoffs from both player 2's actions does not change their comparison, player 2's best response remains the same whether she has EU or (β, c) preference.

7.2 Generalization

In finite two-player strategic games, we have the following general results about the impact of certainty effect (CE) on trembling-hand perfect equilibrium (THPE):

i) CE does not affect THPE if no (β, c)-player has a strategy that gives her a constant payoff, irrespective of the opponent's strategies;

ii) CE may affect THPE if some (β, c)-player has a strategy that gives her a constant payoff. If this is the case, the set of pure strategy THPE (PS-THPE) under CE is a subset of the original set of PS-THPE; in particular, if $c = 0$, then for β small enough, the set of PS-THPE under CE is a *strict* subset of the original set of PS-THPE.

Our analysis of Game C when player 1 has (β, c) preference serves as a "proof" of point ii) above. The intuition of this result is: As we only consider PSNE, and we know that (β, c) preference does not affect PSNE at all, trembling-hand perfection can not expand the set of qualifying PSNE. When player 2 plays a completely mixed strategy, it favors the "safe strategy" of player 1 (i.e. the strategy that gives her a constant payoff) as player 1 discounts any uncertain payoffs with β. For β small enough, such "discounting" is so severe that in the limit of player 2's sequence of completely mixed strategies, player 1's payoff from the "risky strategy" is strictly lower than that of the safe strategy.

7.3 Discussion

In the result we have shown in this section, the discontinuity of the (β, c) weighting function at $p = 1$ has played a crucial role.

As Theorems 0 and 1 show, when $c = 0$ and $\beta < 1$, the discontinuous weighting function results in a *discontinuity in preference representation* - any non-degenerate lottery

f's evaluation $V(f)$ is equal to its expected value $Ev(f)$ "discounted" by β, which makes $V(f)$ strictly smaller than $Ev(f)$ (at least in absolute value).

When a player (e.g. player 1 in Game C) has this kind of (β, c) preference, as well as an action (e.g. D) that yields a certain payoff, her evaluation of strategy profiles consisting of an own action that does not yield a certain payoff (e.g. U), along with the opponent's any completely mixed strategy (e.g. σ_2^p), will always be the "discounted" expected value. Even though the opponent's sequence of completely mixed strategies converges to a pure strategy (i.e. an action, e.g. L), and hence the sequence of strategy profiles converges to a deterministic action profile (e.g. (U, L)), the (β, c) player's evaluation of the sequence of strategy profiles still does **not** converge to the expected value of the limiting action profile (e.g. $v_1(U, L)$). Rather, it converges to its expected value discounted by β, which is strictly lower in absolute value. Therefore, although the limiting action profile itself (e.g. (U, L)) may have a higher evaluation than the profile consisting of the own action yielding the certain payoff and the same action of the opponent (e.g. (D, L)), the sequence of strategy profiles (e.g. (U, σ_2^p)) can still all have lower evaluations than the latter. If this is the case for all completely mixed strategies by the opponent, the limiting action profile (e.g. (U, L)) will not be the best response to any sequence of completely mixed strategies by the opponent, and therefore will not be trembling-hand perfect.

The key intuition of the argument above is that, given even the slightest chance that the opponent may make a mistake, a (β, c) player will strictly prefer the action that leads to a certain payoff to an alternative action that results in a *lower evaluation* compared to the certain payoff, even if the alternative action has a strictly *higher expected value*. This exactly reflects the distortion in evaluation due to the certainty effect of the (β, c) preference.

8 Conclusion

There have been so many "anomalous" empirical findings of people's risky choice behavior that violate the expected utility theory as a normative rule. Prospect theory and cumulative prospect theory are among the best behavioral economic theories that attempt to provide a unified way to understand and explain such findings. With complexity in the functional forms of weighting functions, CPT gains descriptive power and accuracy in approximating empirical evidence. However, what is lost is tractability in its utility representation, which hinders its application to wider risky choice situations.

The (β, c) model we propose achieves highly tractable utility representation of CPT by simplifying weighting functions whilst preserving the basic tenets of CPT. We hope it would become an easier work horse for behavioral economists to study risky choice behaviors, and help expand the scope of applications of CPT. The applications we have discussed illustrate its useful role in what we think are the most natural situations where stand models can be improved. There are potentially many other interesting topics where the (β, c) model can be most helpful, including risky choice situations with both dynamics and strategic interaction, such as in extensive-form games.

References

[1] Conlisk, J. (1989), "Three Variants on the Allais Example", *American Economic Review*, Vol. 79, No. 3 (Jun., 1989), pp. 392-407.

[2] Kahneman, D., A. Tversky (1979), "Prospect Theory: An Analysis of Decision under Risk", *Econometrica*, 47, pp. 263–291.

[3] Keren, G. and Roelofsma, P. (1995): "Immediacy and Certainty in Intertemporal Choice." *Organizational Behavior and Human Decision Processes*, 63(3), pp. 287-297.

[4] Halevy, Y. (2008): "Strotz Meets Allais: Diminishing Impatience and the Certainty Effect." *American Economic Review*, 2008, 98(3), pp. 1145–1162.

[5] Laibson, D. (1997), "Golden Eggs And Hyperbolic Discounting", *Quarterly Journal Of Economics*, March 1997, pp 443-477.

[6] O'Donoghue, T., M. Rabin (1999a), "Doing It Now or Later", *American Economic Review*, LXXXIX 1999, 103-124.

[7] Osborne, M. J., A. Rubinstein (1994), "A Course in Game Theory", MIT Press.

[8] Phelps, E. S., R. A. Poflak (1968), "On Second-Best National Saving and Game-Equilibrium Growth", *Review of Economic Studies*, April 1968, 35(2), pp. 185-99.

[9] Quiggin, J. (1982), "A Theory of Anticipated Utility", *Journal of Economic Behavior and Organization*, 3 (1982), pp 323-343.

[10] Tversky, A., C. Fox (1995), "Weighing Risk and Uncertainty", *Psychological Review*, 102 (2), pp 269-283.

[11] Tversky, A., D. Kahneman. (1992), "Advances in Prospect Theory: Cumulative Representation of Uncertainty", *Journal of Risk and Uncertainty*, 5, pp 297–323.

[12] Tversky, A., P. Wakker. (1995), "Risk Attitudes and Decision Weights", *Econometrica*, 63(6), pp 1255–1280.

[13] Weber, B. and Chapman, G. (2005): "The Combined Effects of Risk and Time on Choice: Does Uncertainty Eliminate the Immediate Effect? Does Delay Eliminate the Certainty Effect?" *Organizational Behavior and Human Decision Processes*, 96, pp. 104-118.

9 Appendix - Application IV - Static Investment Decision

In this section we use simple binary lotteries to illustrate the implications of certainty effect as represented by the (β, c) model in simple static investment choices.

As a first step, consider an investor with preference represented by the (β, c) model where $c = 0$ and $\beta = 0.9$ (whom we call the (β, c)-investor). That is, her weighting function *underweights* the probabilities of all *uncertain* outcomes by a factor 0.9, and only the *certain* outcome gets a full weight 1. Her preference is therefore biased towards certainty.

9.1 Under-Investment with Certain Cost and Uncertain Return

Suppose this investor needs to choose *one* between two investment projects, A and B. Project A requires an initial investment (i.e. cost) of $-C_A = -10$ and yields in the next period a return of either 0 with probability $\frac{1}{2}$ or $R_A = 20$ with probability $\frac{1}{2}$. Project B requires an initial cost of $-C_B = -5$ and yields in the next period a return of either 0 with probability $\frac{1}{2}$ or $R_B = 10$ with probability $\frac{1}{2}$.

We assume the investor evaluates different periods separately, but there is no time discounting. That is, whilst evaluating a project, she does not subtract relevant costs from returns across periods for each "state of the world" to get the "net payoff" in that state. Rather, she evaluates the value of the project in each period, and then sum up all periods. This assumption is consistent with the "mental accounting" type of behavior found in experiments. A "period" simply serves as a reference of "accounting" in project evaluation.

We further suppose the underlying utility function is $v(x) = x$, which implies that an EU-investor with this utility function is risk neutral; and we set the reference point at $x_0 = 0$.

We have constructed this example so that an EU-investor would be indifferent between projects A and B, since

$$Ev(A) = Ev(B) = 0$$

Which project would the (β, c)-investor choose?

Since both projects are binary lotteries, we can evaluate them under the (β, c) model by (8) in Corollary 4, and we have

$$\begin{aligned} V(A) &= \frac{1}{2}\beta R_A - C_A = -1 \\ V(B) &= \frac{1}{2}\beta R_B - C_B = -0.5 \end{aligned}$$

Therefore the (β, c)-investor strictly prefers project B, the one with lower cost (in absolute value) and lower expected return, although both projects provide the same expected utility.

Why is this the case? The reason is exactly the certainty bias in the (β, c)-investor's preference, which says that she overweights certain outcomes relative to uncertain ones.

When the cost of a project is certain but the return is random, the same cost "hurts" a (β, c)-investor more than an EU-investor (due to overweighting). Thus, given the same choice, a (β, c)-investor would try "harder" to avoid the certain cost - she achieves this by choosing the lower-cost project when an EU-investor feels indifferent between the two.

We can summarize this point as: **Certainty effect implies under-investment in projects with certain cost and uncertain return**. And we can actually prove this for quite general cases.

Proposition 2 (Under-Investment with Certain Cost and Uncertain Return) *In investment choice with certain cost and uncertain (discrete or continuous) pure-gain return, when $c = 0$ and $\beta < 1$, a (β, c)-investor will strictly prefer the project with the **lower cost** (in absolute value) between two projects with the same expected utility.*

Proof. Step 1: Discrete pure-gain return

Suppose each project $i \in \{A, B\}$ requires a *certain* (occurring with probability 1) cost in the first period and yields a *random* return in the second period. Since we only consider "pure-gain" return, we suppose the lowest possible returns that both these projects yield in the second period are the same, which we take as the reference point x_0 and normalize to $x_0 = 0$.

We denote by R_i the "return lottery" of project i (note here R_i is not an outcome but a random variable), whose outcomes are $r_{i0}(= 0) < r_{i1} < ... < r_{in_i}$, with respective probabilities $p_{ij} \in (0, 1)$, $j \in \{0, 1, 2, ..., n_i\}$, and $\sum_{j=0}^{n_i} p_{ij} = 1$.

Denote C_i the (absolute value of) cost of project i and assume without loss of generality $C_A > C_B > 0$.

Suppose $v(x) = x$ (risk neutrality), which is consistent with our general assumption $v(x_0) = 0$.

We want to show that when $c = 0$ and $\beta < 1$, a (β, c)-investor will always strictly prefer project B (the one with the lower certain cost) when A and B provide the same expected utility.

The EU of lottery R_i is

$$Ev(R_i) = \sum_{j=0}^{n_i} p_{ij} r_{ij}$$

The EU of project i is

$$Ev(i) = Ev(R_i) - C_i$$

Suppose an EU-investor is indifferent between A and B:

$$Ev(A) = Ev(B)$$

which implies

$$C_A - C_B = Ev(R_A) - Ev(R_B) > 0$$

Now we study the choice of a (β, c)-investor with $c = 0$ and $\beta < 1$.

Since the returns are pure-gain lotteries, from (7), we have

$$V(i) = \beta Ev(R_i) - C_i \quad (19)$$

Therefore

$$\begin{aligned} & V(A) - V(B) \\ = \ & (\beta Ev(R_A) - C_A) - (\beta Ev(R_B) - C_B) \\ = \ & \beta(Ev(R_A) - Ev(R_B)) - (C_A - C_B) \\ = \ & (C_A - C_B)(\beta - 1) \end{aligned}$$

Since $C_A > C_B$ we have $V(A) < V(B)$ if and only if $\beta < 1$. Therefore the (β, c)-investor will always strictly prefer the project with the lower cost.

Step 2: Continuous pure-gain return

Suppose the returns of both projects are continuous lotteries satisfying Definition 2. Then by Theorem 2 we know (19) still holds. Therefore the proof above follows through. We are done. ∎

Notice this result and the next one are not driven by the risk-averse or risk-seeking attitudes in the EU theory, as we have used a risk neutral utility function $v(x) = x$. These results are simply due to the certainty effect imbedded in the (β, c) model.

9.2 Over-Investment with Uncertain Cost and Certain Return

When we swap the roles of costs and returns in the previous example, we will find that **certainty effect also implies over-investment in projects with certain return and uncertain cost**.

To see this, suppose A now provides a sure return of $R_A = 10$ in the second period, but its cost in the first period may be either 0 or $-C_A = -20$ with equal probability. Project B provides a sure return of $R_B = 5$ in the second period and costs either 0 or $-C_B = -10$ with equal probability in the first period.

Buying a house on floating-rate mortgage might be a good real-life example of such investments if we consider the "return" as the benefit of living in the house (which is thus certain) while the cost is the monthly mortgage payment (which is uncertain due to the floating rate). Other similar choice situations include a foreigner planning a trip to see the London 2012 Olympic Games, where he is quite certain about the reward from this trip (i.e. seeing the Games) but the cost (e.g. flight price) may well be subject to changes until he makes up his mind (e.g. deciding to go and booking the flights).

Faced with this choice, an EU-investor still feels indifferent as $Ev(A) = Ev(B) = 0$. How would the (β, c)-investor choose?

Again by Corollary 4, we have

$$\begin{aligned} V(A) &= R_A - \frac{1}{2}\beta C_A = 1 \\ V(B) &= R_B - \frac{1}{2}\beta C_B = 0.5 \end{aligned}$$

Therefore the (β, c)-investor strictly prefers project A, the one with higher expect cost (in absolute value) and higher return, although both projects provide the same expected utility.

When the return of a project is certain but the cost is random, the certainty effect implies that a (β, c)-investor would overweight the return relative to the cost, and therefore would try "harder" to get a high return than suggested by EU theory. In other words, due to underweighting of the uncertain costs, a (β, c)-investor can "bear" higher cost than an EU-investor. Therefore she chooses the higher-cost and higher-return project when an EU-investor feels indifferent between two projects.

We can also prove this point for general cases.

Proposition 3 (Over-Investment with Uncertain Cost and Certain Return) *In investment choice with uncertain (discrete or continuous) pure-loss cost and certain return, when* $c = 0$ *and* $\beta < 1$, *a* (β, c)*-investor will strictly prefer the project with the* ***higher expected cost*** *(in absolute value) between two projects with the same expected utility.*

Proof. Step 1: Discrete pure-loss cost

Suppose each project $i \in \{A, B\}$ requires a random cost in the first period which is a pure-loss lottery and yields a certain return in the second period. We suppose the lowest possible absolute value of costs in both projects are the same, which we take as the reference point x_0 and normalize to $x_0 = 0$.

Since we use $v(x) = x$, there is no asymmetry between gains and losses. Therefore we can treat the absolute values of costs as pure-gain lotteries to lighten notation in the following proof. In particular, we denote the "cost lottery" of project i as C_i, whose outcomes are $c_{i0}(= 0) < c_{i1} < ... < c_{in_i}$, with respective probabilities $p_{ij} \in (0, 1)$, $j \in \{0, 1, 2, ..., n_i\}$, and $\sum_{j=0}^{n_i} p_{ij} = 1$.

Denote R_i the return of project i and assume $R_A > R_B > 0$.

Suppose $v(x) = x$ (risk neutrality), which is consistent with our general assumption $v(x_0) = 0$.

We want to show that when $c = 0$ and $\beta < 1$, a (β, c)-investor will always strictly prefer project A (with the higher return and expected cost) when A and B provide the same expected utility.

The EU of lottery C_i is

$$Ev(C_i) = \sum_{j=0}^{n_i} p_{ij} c_{ij}$$

The EU of project i is

$$Ev(i) = R_i - Ev(C_i)$$

Suppose an EU-investor is indifferent between A and B:

$$Ev(A) = Ev(B)$$

which implies

$$Ev(C_A) - Ev(C_B) = R_A - R_B \tag{20}$$

Now we study the choice of a (β, c)-investor with $c = 0$ and $\beta < 1$.

Since the (absolute values of) costs are pure-gain lotteries, from (7) in Corollary 2 we have

$$V(i) = R_i - \beta Ev(C_i)$$

Therefore

$$\begin{aligned} & V(A) - V(B) \\ = \; & (R_A - \beta Ev(C_A)) - (R_B - \beta Ev(C_B)) \\ = \; & (R_A - R_B) - \beta(Ev(C_A) - Ev(C_B)) \\ = \; & (R_A - R_B)(1 - \beta) \end{aligned}$$

Since $R_A > R_B$ we have $V(A) > V(B)$ if and only if $\beta < 1$. Notice when A and B provide the same expected utility, by (20) we know $R_A > R_B$ is equivalent to $Ev(C_A) > Ev(C_B)$. Therefore the (β, c)-investor will always strictly prefer the project with the higher expected cost.

Step 2: Continuous pure-loss cost

Suppose the costs of both projects are continuous lotteries satisfying Definition 2. Then by Theorem 2 we know the proof still follows through. We are done. ■

In the next section we discuss investment choice with more general values of β and c.

9.3 Investment Choice with General (β, c) Values

In Propositions 2 and 3, we have kept $c = 0$ for simplicity. When $c > 0$, from Theorem 1 we know the extreme outcomes of the return lottery or the cost lottery will get overweighted relative to the other outcomes by factor c. This means the comparison between two investment projects would depend not only on their expected return or expected cost, but also on the distribution of their outcomes. Therefore the range of cases in which we can draw comparison conclusions as general as Propositions 2 and 3 are limited. While we can still show the conditions required for such conclusions given specific distributions, in this part we illustrate the intuition only using the special case of *binary lotteries*. We continue to use the risk-neutral utility function $v(x) = x$, and set reference point at $x_0 = 0$.

Certain cost and uncertain return Consider two investment projects A and B, where project $i \in \{A, B\}$ requires a certain cost $-C_i$ in the first period and yields a random pure-gain return $(0, 1-p; R_i, p)$ in the second period, $R_i > 0$, $p \in (0,1)$. Then we have the following conclusion:

Proposition 4 *In the investment choice with certain cost and uncertain return described in this section, a (β, c)-investor strictly prefers the project with the* ***lower cost*** *(in absolute value) between two projects with the same expected utility* ***if and only if*** $\beta < 1 - \frac{c}{p}$.

Proof. First note that p is the probability of the higher outcome in the return lottery, and $p \in (0, 1)$.

Without loss of generality, suppose $C_A > C_B$.

Since $Ev(A) = Ev(B)$ and $Ev(i) = pR_i - C_i$, we have

$$C_A - C_B = p(R_A - R_B) > 0$$

Now we find the condition for a (β, c)-investor to choose project B (the one with lower cost) instead of A, that is

$$V(B) > V(A)$$

Since the returns are pure-gain lotteries, i.e. $0 = x_0 < R_i$, $i \in \{A, B\}$, from Corollary 2 we have

$$V(i) = (p\beta + c)R_i - C_i$$

Therefore

$$\begin{aligned}
V(B) &> V(A) \\
&\Leftrightarrow (p\beta + c)R_B - C_B > (p\beta + c)R_A - C_A \\
&\Leftrightarrow C_A - C_B > (p\beta + c)(R_A - R_B) \\
&\Leftrightarrow p(R_A - R_B) > (p\beta + c)(R_A - R_B) \\
&\Leftrightarrow p > p\beta + c \\
&\Leftrightarrow \beta < 1 - \frac{c}{p}
\end{aligned}$$

For instance, suppose $p = 0.5$. Then the parameter values $(\beta = 0.7, c = 0.1)$ would guarantee that a (β, c)-investor chooses B. ■

Corollary 6 *In the investment choice situation with certain cost and uncertain return described in this section, a (β, c)-investor strictly prefers the project with the* ***lower cost*** *(in absolute value) between two projects with the same expected utility* ***if and only if*** *the weighting function w satisfies $w(p) < p$, where $p \in (0, 1)$ is the probability of the higher outcome in the return lottery.*

Proof. In the last step of the proof of Proposition 4, we have found a sufficient and necessary condition

$$p > p\beta + c$$

But since $p \in (0, 1)$, by (3) we know $w(p) = p\beta + c$. Therefore $w(p) < p$ is also a sufficient and necessary condition. ■

Intuition: The intuition here is still under-investment with certain cost due to certainty effect. Notice the condition $w(p) < p$ in Corollary 6 is consistent with the empirical evidence in CPT.

Uncertain cost and certain return Now we reverse the roles of cost and returns. Consider two investment projects A and B, where project $i \in \{A, B\}$ requires a random cost in the first period which is a pure-loss lottery $(0, 1 - p; -C_i, p)$, with $p \in (0, 1)$, and yields a certain return R_i in the second period. Then we have the following conclusion:

Proposition 5 *In the investment choice with uncertain cost and certain return described in this section, a (β, c)-investor strictly prefers the project with the* ***higher expected cost*** *(in absolute value) between two projects with the same expected utility* ***if and only if*** $\beta < 1 - \frac{c}{p}$.

Proof. Without loss of generality, suppose $C_A > C_B$.

Since $Ev(A) = Ev(B)$ and $Ev(i) = R_i - pC_i$, we have

$$R_A - R_B = p(C_A - C_B) > 0$$

Now we find the condition for a (β, c)-investor to choose project A (the one with higher expected cost) instead of B, that is

$$V(A) > V(B)$$

Since the costs are pure-loss lotteries, i.e. $0 = x_0 > -C_i$, $i \in \{A, B\}$, from Corollary 3 we have

$$V(i) = R_i - (p\beta + c)C_i$$

Therefore

$$\begin{aligned} V(A) &> V(B) \\ &\Leftrightarrow R_A - (p\beta + c)C_A > R_B - (p\beta + c)C_B \\ &\Leftrightarrow R_A - R_B > (p\beta + c)(C_A - C_B) \\ &\Leftrightarrow p(C_A - C_B) > (p\beta + c)(C_A - C_B) \\ &\Leftrightarrow p > p\beta + c \\ &\Leftrightarrow \beta < 1 - \frac{c}{p} \end{aligned}$$

Again, if $p = 0.5$, the parameter values $(\beta = 0.7, c = 0.1)$ would guarantee that a (β, c)-investor chooses A. ■

Corollary 7 *In the investment choice with uncertain cost and certain return described in this section, a (β, c)-investor strictly prefers the project with the* ***higher expected cost*** *(in absolute value) between two projects with the same expected utility* ***if and only if*** *the weighting function w satisfies $w(p) < p$, where $p \in (0, 1)$ is the probability of the higher outcome in the return lottery.*

Proof. The same as the proof of Corollary 6. ■

Conclusion of Thesis

The theories we have presented in this thesis are relevant and applicable in many real-life situations.

On the industry level, the theory of mixed two-sided markets models one of the fastest growing and most promising industries - the "platform" industry. The past two decades have seen the emergence and dominance of gigantic platforms including Windows, Google and Facebook. Other less obvious platforms include a wide range of telecommunication networks and financial intermediaries. Understanding the two defining characteristics of such markets, the two-sidedness and the mixedness, is key to understanding their present and future development.

On the firm level, pricing strategies are of ultimate importance to platforms, as well as to other "ordinary" businesses in general, such as fitness clubs and shopping malls. The theory of multiproduct pricing with two-part tariffs is therefore useful in a wider context. The two-part-tariff effect we have found in the first two parts of the thesis helps us understand the prevalence of two-part tariffs in real life. The simplicity and implementability of two-part tariffs make them very practical ways to increase profit over conventional separate pricing.

On the individual level, decision making of risky choice occurs on a daily basis. From static investment decisions to dynamic purchasing decisions, and from individual decision problems to strategic interaction between decision makers, our simplified model of decision weighting provides a transparent way of understanding a wide range of "anomalous" phenomena that violate standard utility models.

Given the relevance and range of applications, we hope the theories presented in this thesis prove a useful addition to the microeconomic literature as well as an adequate foundation for future research.