Bundling and Nonlinear Pricing in Telecommunications

Yao Luo *

University of Toronto

April 2015

Abstract

This paper studies bundling and price discrimination by a multiproduct firm selling internet and phone services in an imperfect information setting. I derive the optimal selling mechanism, and provide primitive conditions under which different bundling strategies arise. Exploiting the optimality conditions of both the firm and the consumer, I show that the model structure is nonparametrically identified and propose a three-step semiparametric estimation procedure. An application to China Telecom data shows that mixed bundling is beneficial to both the firm and the consumer relative to component pricing.

Key Words: Bundling, Nonlinear Pricing, Multidimensional Screening, Telecommunications, Nonparametric Identification

JEL Codes: L11, L12, L25, L96

*E-mail: *yao.luo@utoronto.ca*. This article is a revised version of the third chapter of my PhD thesis at The Pennsylvania State University. I am very grateful to Isabelle Perrigne and Quang Vuong for advice and support. I thank seminar and conference participants, and Victor Aguirregabiria, Liang Chen, Yongmin Chen, Paul Grieco, Albert Ma, Ali Yurukoglu, Marc Rysman, Xiaobo Tao, Neil Wallace and Julian Wright for constructive comments, as well as China Telecom for providing the data. All remaining errors are mine.

Ah, the Internet! The source of so much goodness. The font of e-mail, news, chat, TV, blogs, books and Facebook. What would we do without it?

– David Pogue (The New York Times. March 21, 2012)

1 Introduction

Bundling and nonlinear pricing have become increasingly prevalent practices in sectors such as travel services, retail and telecommunications. Many telecommunication service providers are now offering bundled service packages with internet, phone and cable TV. Moreover, they often offer nonlinear tariffs by providing additional discounts when subscribers buy larger quantities.

This paper studies bundling and nonlinear pricing by a major telecommunication company in China, China Telecom, that sells internet and phone services to customers through mixed bundling.[1] It aims to evaluate the welfare effects of bundling. While bundling has become a popular selling strategy for many products, the net effects on social welfare depend on a number of factors whose influence can only be determined on the basis of empirical analysis. Since I have data from a single market, there is no natural instruments to solve the endogeneity problems of consumers' purchased quantities. Relying solely on the consumer's optimality condition leads to biased estimates. Instead, I exploit both the firm's and the consumer' optimality conditions to achieve identification.[2]

To this end, I propose a new multiproduct nonlinear pricing model and derive the optimal selling mechanism. While the existing literature mostly addresses bundling and nonlinear pricing separately, my model incorporates both simultaneously as recently done by Armstrong and Vickers (2010). Armstrong (1996) and Rochet and Chone (1998) show that nonlinear pricing for a multiproduct firm is a complex problem because of multidimensional screening. To date there exist only a few scattered papers that allow for multidimensional types, which are needed to capture the basic economics of the environment in my empirical application. My theoretical model explores another tractable way in which multidimensional types can be incorporated in a multiproduct nonlinear pricing problem.

My model endogenizes both the firm's bundling and pricing decisions. It explains which bundling strategy should be adopted by the firm: component pricing, pure bundling, semi-mixed bundling or mixed bundling. My model is general as it allows for various bundling

[1] As the world's largest fixed-line operator as well as the largest wireless broadband operator, China Telecom generated a revenue of about 321.58 billion RMB (51.69 billion US dollar) in 2013.

[2] This is reminiscent of Ekeland, Heckman, and Nesheim (2004) who show that the hedonic model is identified nonparametrically within a single market by exploiting the full economic content of the model.

incentives: utility complementarity, cost saving effects and dependence between the two dimensions of asymmetric information. They have different roles in determining the welfare effects of bundling relative to unbundling. In general, the effects of bundling on consumer surplus and social welfare are ambiguous. On one hand, bundling may benefit the consumers because consumers with a lower taste will not be excluded. On the other hand, bundling can reduce consumer surplus because it provides an additional instrument for the firm to discriminate across consumers.

My model allows for a separable utility into the benefit of consuming phone service only and the complementary benefit of using both internet and phone services, as well as positively dependent types for internet and phone services. I exploit the discrete nature of internet service to solve the multidimensional screening problem. In particular, the number of effective incentive compatibility constraints is significantly reduced. This allows me to characterize the optimal exclusion conditions, the assignment schedules and the tariff functions in an equivalent one-dimensional formulation. Specifically, the provider offers usage-based nonlinear tariffs for phone service and a fixed-fee for internet-only users. The curvature of the tariffs for phone service will differ according to the internet service level. In other words, bundling enables the provider to further discriminate consumers choosing different levels of internet. In addition, I provide the conditions on the primitives under which the firm optimally chooses his bundling strategy, i.e., component pricing, pure bundling, semi-mixed bundling or mixed bundling. Bundling is more likely to dominate component pricing when the cost to provide phone service is lower or consumers value phone service more highly.

I study the identification of the model structure from observables in a single market: the price schedule and consumers' purchased quantities. This part is reminiscent of the nonparametric identification of auction models. See e.g. Guerre, Perrigne, and Vuong (2000) and Athey and Haile (2007). Under a parameterization of the cost function and a multiplicative separability of the utility function in the willingness-to-pay for phone service, I show that the primitives are identified. In particular, the complementary utility function is identified by exploiting the phone usage and tariff variation across consumers adopting different internet levels. While cost saving effects are identified by exploiting the firm's optimal exclusion conditions, the dependence between the two types is identified using the one-to-one mapping between phone usage and the corresponding consumer type. Following the identification results, I propose a three-step semiparametric estimation method and apply it to China Telecom data.

My analysis of China Telecom data shows that (i) internet and phone services tend to be substitutes. Internet offers alternative communication tools such as email, skype and so on, which can explain the substitutability with phone service. Thus the utility of a bundle user is smaller than the sum of the utilities for a phone service user only and a internet user only. Moreover, this substitution effect is stronger with a higher level of internet because a faster internet service allows better alternative communication tools; (ii) the additional fixed cost to bundle internet with phone service is small. China Telecom mainly uses Asymmetric Digital Subscriber Line (ADSL) to provide internet service. Thus internet is transmitted through telephone lines. Moreover, fixed transaction costs such as mailing statement do not increase much because bills are merged if the consumer uses a bundle; and (iii) a higher speed internet adopter tends to have a higher willingness-to-pay for phone service. This is consistent with the Federal Communications Commission 2009 survey which shows that higher speed internet adopters tend to be better educated with higher incomes.

With the estimated model from China Telecom data, I run counterfactual experiments to assess the gain for both the firm and consumers from bundling internet and phone services. My simulation results show that unbundling would lead to a 10.14% decrease in firm's profit and a 17.18% decrease in consumer surplus. This arises from a larger proportion of consumers who would be excluded under component pricing. Relative to mixed bundling, low-end consumers would face more expensive tariff functions under component pricing, while high-end consumers would face a less expensive one. As a result, the two low-end groups would lose consumer surplus by 58.11% and 39.57%, respectively. On the contrary, high-end group would see an increase in their surplus by 9.48%. Thus unbundling would only benefit high-end consumers.

Literature

There is an extensive theoretical and empirical literature on bundling and nonlinear pricing within an incomplete information framework. While the existing literature mostly addresses bundling and nonlinear pricing separately, my model incorporates both simultaneously as recently done by Armstrong and Vickers (2010).

Regarding nonlinear pricing, Armstrong (2006) and Stole (2007) provide recent theoretical surveys. Armstrong (1996) and Rochet and Chone (1998) show that nonlinear pricing for a multiproduct firm is a complex problem because of multidimensional screening. On the empirical side, Leslie (2004), McManus (2007) and Crawford and Shum (2007) render this problem one-dimensional by, e.g., considering unit-demand consumers or homogenizing the products. Using convenient parameterization of model primitives, Ivaldi and Martimort (1994) and Miravete and Röller (2004) endogenize the firm's pricing decision. My paper introduces another tractable way in which multidimensional types can be incorporated in a multiproduct nonlinear pricing problem.

Regarding bundling, the early theoretical literature considers a benchmark case with two single unit products and additive separable utility. See, e.g., Adams and Yellen (1976), McAfee, McMillan, and Whinston (1989) and Salinger (1995). The main condition under which pure or mixed bundling is preferred by the firm over component pricing relates to the correlation of types for the two products. See also Chen and Riordan (2013). Considering a large number of products, Armstrong (1999) and Bakos and Brynjolfsson (1999) show that the firm can approximate the first-best by bundling. Chu, Leslie, and Sorensen (2011) provide a simulation of bundle-size pricing. On the empirical side, Crawford and Yurukoglu (2012) and Ho, Ho, and Mortimer (2012) consider the firm's bundling decision as exogenous and estimate the welfare effects of bundling between upstream and downstream firms. In contrast, my model endogenizes both the firm's bundling and pricing decisions. While the previous studies focused on one or two bundling incentives, my paper allows for utility complementarity, cost saving effects and dependence between multidimensional types.

This paper is also related to the telecommunication literature. Most studies focus on the U.S. market. See, e.g., Miravete (2002) for field experiment data from Kentucky; Liu, Chintagunta, and Zhu (2010) for households' adoption of cable TV, local phone, and broadband Internet access; Crawford and Yurukoglu (2012) for households' choices of cable channels; Grubb and Osborne (2013) for university students' choices of mobile plans. Despite growing public interest in the Chinese telecommunications market, little research has been done on it. This paper is among the first to study consumer-level data from this market. See also Luo, Perrigne, and Vuong (2014).

Recently, Luo, Perrigne, and Vuong (2012) rely on the seminal Armstrong (1996) model to analyze nonlinear pricing of mobile voice and messaging services.[3] However, the same model cannot be applied for my purpose in this paper because Armstrong (1996) requires continuous quantities. As I introduce discrete levels of internet service, either I obtain a non-differentiable cost-based indirect utility function or there is effectively only one level of internet service. This paper proposes a new multiproduct nonlinear pricing model that incorporates product discreteness.

The paper is organized as follows. Section 2 presents the data. Section 3 introduces the model, while Section 4 studies the identification of the model primitives and develops a semi-parametric estimation procedure. Section 5 presents estimation results and counterfactuals. Section 6 concludes. An appendix collects the proofs.

[3]Nonlinear pricing for a single product firm leads to a closed-form solution. See, e.g., Maskin and Riley (1984). Most of the empirical literature, including Leslie (2004) and McManus (2007), uses discrete choice models while considering prices exogenous. While endogenizing the price, Luo, Perrigne, and Vuong (2014) show that the model primitives are identified and develop a nonparametric estimation method.

2 Data

The Chinese telecommunications industry is dominated by three state-owned firms: China Telecom, China Unicom and China Mobile. Table 1 gives the nationwide market structure by the number of subscribers. While China Mobile has dominated mobile services since its inception, China Telecom and China Unicom roughly divide the territory in half for internet and land line services: China Telecom in southern China and China Unicom in northern China.

Table 1: Number of Subscribers as of Dec 31, 2009 (in millions)

	Fixed Line	Broadband	Mobile Service
China Telecom	189	53	56
China Unicom	103	39	145
China Mobile	25	6	522

I collected data from China Telecom in a major metropolitan area in the South, where it enjoys a market share of 85% for both internet and land line subscriptions. The sample is composed of all new residential subscribers in August 2009, who receive internet and phone services through the One Home label. For the month of September 2009, the data contain for each subscriber: his/her choice of internet service, the total number of minutes used and the amount paid. There are two internet speed levels, 1 Mbps or 2 Mbps, resulting in three possible bundles with phone service: phone service only (no internet), a bundle with 1 Mbps internet and a bundle with 2 Mbps internet. Table 2 provides summary statistics on the bill measured in RMB and the number of phone minutes by internet choice with the corresponding number of subscribers. The bill paid by a consumer combines internet, different types of phone calls, extra fees for peak hours usage and several add-ons such as voice mail service, music on hold, ring tones, etc. The data do not provide detailed information on these extra features and the corresponding prices. All these extra features explain the important variability of the per minute rate. The consumption of phone calls tends to increase with the level of internet. I remark that the per minute rate tends also to increase with the level of internet.

China Telecom implements mixed bundling, i.e., internet and phone services are offered either separately or in a bundle. The firm charges a fixed fee to internet-only users and a usage-based tariff to bundle users. Specifically, the monthly fixed fee is 78 and 88 RMB for 1 and 2 Mbps, respectively. The usage-based tariff differs with the level of internet and is nonlinear. Table 3 provides the regression of the bill on the number of total minutes and its square for each bundle. The three tariffs are increasing and concave, i.e., consumers pay

Table 2: Summary Statistics

Internet	Variable	N	Mean	S.D.	Min	Max
0	bill	7683	74.91	76.17	19	972.70
	total minutes	–	600.21	720.84	10	4990
	per minute rate	–	0.2129	0.2981	0.0057	9.2333
1	bill	11206	131.06	48.70	99.92	998.84
	total minutes	–	627.88	717.66	10	4956
	per minute rate	–	0.2496	0.4637	0.0060	10.4996
2	bill	12406	176.47	60.85	108	985.99
	total minutes	–	973.13	1045.65	10	4992
	per minute rate	–	0.2727	0.6037	0.0077	23.5360

a lower price per minute of call time when they consume more. Moreover, subscribers tend to pay more for the same amount of phone calls when they choose a higher internet speed. To see this, I calculate the mean of the per minute rate for the three bundles: 21.29 cents with no internet, 24.96 cents with 1 Mbps internet, and 27.27 cents with 2 Mbps internet. Table 3 suggests that bundling enables the provider to discriminate further phone service users depending on their choice of internet.

Table 3: Regression of the Bill [a]

Internet (Mbps)	0	1	2
$TotalMin$	0.0859	0.0978	0.1049
	(0.0010)	(0.0013)	(0.0011)
$TotalMin^2$	-3.61e-06	-4.12e-06	-7.01e-06
	(1.04e-07)	(1.61e-07)	(1.40e-07)
$Constant$	81.1750	112.1811	150.8603
	(0.8021)	(0.9854)	(1.1487)
Adjusted R^2	0.2711	0.4186	0.4086

[a] Note: standard errors are in parentheses.

Given that I need the tariff functions in view of Sections 3 and 4, I follow Luo (2011) method to estimate the tariff function for each bundle while taking into account unobserved add-ons and features. Details can be found in Appendix B. Figure 9 displays the resulting three tariff functions denoted by $T_0(\cdot), T_1(\cdot), T_2(\cdot)$ for the three bundles. I then construct a quantity of phone usage $q \equiv T_j^{-1}(t)$, where t is his/her payment and j is his/her internet choice. The quantity q aggregates all observed quantity of minutes as well as the unobserved phone services chosen by the consumer. In Section 4 (on identification and estimation), I consider that $(t, q, j, T_j(\cdot))$ are the observables.

3 The Model

3.1 Assumptions and Notations

In view of the discussion in Section 2, I consider a monopoly provider selling internet and phone services as separate products or in a bundle. Internet is offered at several speed levels, denoted by $j \in \mathcal{J}$, where $\mathcal{J} \equiv \{0, 1, 2, \ldots, J\}$ is the possible choice set of internet levels with 0 denoting no internet.[4] Phone service is measured by $q \in \mathbb{R}^+$. Due to implementation difficulties of random contracts, the provider offers non-random nonlinear pricing schedules of the general form $T(q, j)$, with $q \in \mathbb{R}^+$ and $j \in \mathcal{J}$.

A consumer is characterized by a vector of types $(\theta, \beta) \in \Theta \equiv [\underline{\theta}, \overline{\theta}] \times [\underline{\beta}, \overline{\beta}]$, where $0 < \underline{\theta} < \overline{\theta} < \infty$ and $\underline{\beta} \leq 0 < \overline{\beta} < \infty$. The type θ represents his taste or willingness-to-pay for phone service and β defines the minimum internet need above which he will consider buying internet. The latter can be nonpositive because some consumers may have negative perspectives on internet. The vector (θ, β) is private information. That is, the provider does not know the consumer' types but knows the joint distribution $F(\cdot, \cdot)$ on $[\underline{\theta}, \overline{\theta}] \times [\underline{\beta}, \overline{\beta}]$.[5]

The consumer chooses internet speed and phone usage. Following Dubin and McFadden (1984) and Hanemann (1984), I assume that there is no uncertainty in the decision on the continuous variable (phone usage) at the time of the choice for internet. In other words, both choices are made simultaneously. I make the following assumption on the consumer's utility function.

Assumption 1: *A (θ, β) agent consuming (q, j) gets utility*

$$U(q, j; \theta, \beta) = \begin{cases} U(q, j; \theta) & \text{for } j \geq \beta, \\ 0 & \text{for } j < \beta. \end{cases}$$

Assumption 1 prevents the consumer to choose a bundle with an internet speed that falls below his minimum need β. It is analogous to the absolute spending limit assumption in Che and Gale (2000). The value β results from the consumer's internet use such as sending emails, shopping online, playing games or streaming movies. Internet connection speed determines whether these applications will run effectively. Thus I assume failing to meet the

[4]This choice set is exogenous given by technological constraint. See Mazzeo (2002) and Seim (2006) for endogenizing the product decisions of the firm.

[5]Because modeling competition is out of the scope of this paper, we can view θ as a sufficient statistic that summarizes preferences of the consumer for phone service by China Telecom and its competitors. See Ivaldi and Martimort (1994) for an example. Moreover, I do not consider uncertainty on types, which leads to a two-stage model. See, e.g., Miravete (2002), Miravete (2005), Narayanan, Chintagunta, and Miravete (2007), Economides, Seim, and Viard (2008) and Grubb and Osborne (2013).

minimum internet need leads to a very large disutility. Once the minimum internet speed is satisfied, the underlying family of indifference curves of (q, j) bundles can be described by the variation of a single parameter, i.e., the taste for phone service θ.[6]

Assumption 1 has several advantages. First, it makes the optimal selling mechanism with multidimensional types tractable. See e.g. Armstrong (1996), Rochet and Chone (1998) and Rochet and Stole (2003) for mechanism design with multidimensional types. Second, while maintaining multidimensional types, this general utility function allows complementarity between the two products as well as dependence between the types θ and β. Consequently, all the possible scenarios of bundling may arise at the equilibrium such as component pricing, pure bundling, semi-mixed bundling and mixed bundling.[7]

A (θ, β) consumer chooses a quantity of phone service and internet level (q, j) to maximize his payoff

$$\begin{aligned} \max_{q \in \mathbb{R}^+, j \in \mathcal{J}} \quad & U(q, j; \theta) - T(q, j) \\ s.t. \quad & j \geq \beta. \end{aligned}$$

The firm needs to design a price schedule $T(\cdot, \cdot)$ that maximizes his expected profit. Without loss of generality, I apply the Revelation Principle.[8] In particular, any implementable allocation achieved with a price schedule $T(\cdot, \cdot)$ can also be achieved with a truthful direct mechanism of the form $\{t(\cdot, \cdot), q(\cdot, \cdot), j(\cdot, \cdot)\}$. This mechanism specifies the payment made $t(\theta, \beta)$, the quantity of phone service $q(\theta, \beta)$ and internet speed $j(\theta, \beta)$ for a (θ, β) consumer. As information goods, internet and phone services involve very small variable production costs but substantial transaction costs per customer, such as usage recording, billing and customer service. Thus I assume that the provider's total cost is additively separable across consumers. The cost to serve a consumer with a bundle (q, j) is denoted as $c(q, j)$.

[6] An alternative assumption would be to consider the utility as $U(q, j; \theta) - \delta(\beta)$ for $j < \beta$ where $\delta(\beta)$ captures the disutility for not getting the desired amount of internet speed. If $\delta(\beta)$ is large enough, Proposition 1 extends resulting in the same optimal selling mechanism. However, the proofs of Lemmas 1 and 3 wound be significantly longer.

[7] The model can be extended to entertain both q and θ multidimensional relying on Armstrong (1996). The basic idea would be to design a cost-based tariff. To do so, I would need to define the cost-based indirect utility function $V(c, j; \theta) \equiv \max_{c(q_1, q_2, j) \leq c} U(q_1, q_2, j; \theta_1, \theta_2)$, and $V(c, j; \theta) = h(\theta_1, \theta_2)u(c, j) + v(c, j)$. This model is left for future research. See also Luo, Perrigne, and Vuong (2012).

[8] My model is not standard as the consumer's message space is a correspondence which depends on his true taste. However, the Revelation Principle is still valid as the Nested Range Condition in Green and Laffont (1986) is satisfied. I thank Albert Ma for pointing this paper out to me.

The optimal selling mechanism solves

$$\begin{aligned}
\max_{t(\cdot,\cdot),q(\cdot,\cdot),j(\cdot,\cdot)} \quad & \int_{\Theta} \Big[t(\theta,\beta) - c\Big(q(\theta,\beta), j(\theta,\beta)\Big)\Big] f(\theta,\beta) d\theta d\beta \\
s.t. \quad & U(q(\theta,\beta), j(\theta,\beta);\theta) - t(\theta,\beta) \geq U(q(\tilde{\theta},\tilde{\beta}), j(\tilde{\theta},\tilde{\beta});\theta) - t(\tilde{\theta},\tilde{\beta}), \\
& U(q(\theta,\beta), j(\theta,\beta);\theta) - t(\theta,\beta) \geq 0, \\
& j(\theta,\beta) \geq \beta,
\end{aligned}$$

for all $(\theta,\beta) \in \Theta$ and $(\tilde{\theta},\tilde{\beta}) \in \Theta$ such that $j(\tilde{\theta},\tilde{\beta}) \geq \beta$. The first inequality is the incentive compatibility (IC) constraint, which requires that the consumer truthfully reports his types. The second inequality is the individual rationality (IR) constraint, which requires that the consumer has the option of not buying anything from the provider. The outside option is normalized to 0. The third inequality is the minimum need (MN) constraint, which requires that the consumer can use an internet level above his minimum need for internet.

Hereafter, the subscript q (θ) denotes the partial derivative with respect to q (θ) respectively. I make the following assumptions on the model structure.

Assumption 2: *For all $(\theta,\beta) \in \Theta$, $q \in \mathbb{R}^+$, and $j \in \mathcal{J}$, $U(\cdot,\cdot;\cdot)$, $c(\cdot,\cdot)$ and $F(\cdot,\cdot)$ satisfy*

(i) $U(0,0;\theta) = 0$, $U_q(q,j;\theta) \geq 0$, $U_{qq}(q,j;\theta) \leq 0$, $U_\theta(q,j;\theta) > 0$, $U_{\theta\theta}(q,j;\theta) \leq 0$,

(ii) $U_{q\theta}(q,j;\theta) > 0$,

(iii) $\frac{\partial}{\partial\theta}\frac{-U_{qq}(q,j;\theta)}{U_q(q,j;\theta)} \leq 0$,

(iv) $\frac{c_{qq}(q,j)}{c_q(q,j)} > \frac{U_{qq}(q,j;\theta)}{U_q(q,j;\theta)}$,

(v) $H(\theta|j) \equiv \theta - \frac{1-F(\theta|D(\beta)=j)}{f(\theta|D(\beta)=j)}$ *is increasing in θ, where* $D(\beta) \equiv \min\Big\{j \in \mathcal{J} : j \geq \beta\Big\}$,

(vi) $U(0,j;\theta) = v(0,j) \geq c(0,j)$, *for some function* $v(\cdot,\cdot) : \Theta \to \mathbb{R}^+$.

All these assumptions with the exception of (vi) are standard in the nonlinear pricing literature. See e.g. Maskin and Riley (1984). Assumption 2-(i) says that the outside option (not buying) provides a zero utility and the marginal utility from phone service is nonnegative and decreasing. Moreover, a consumer with a higher taste θ gets a larger utility and this increase is diminishing as θ increases. Assumption 2-(ii) is the standard Spence-Mirrlees single-crossing condition, which says that a consumer with a higher taste θ enjoys a larger marginal payoff for phone usage across every (q,j). Assumption 2-(iii) implies nonincreasing absolute risk aversion, while 2-(iv) requires that the cost function is not too concave in q. The latter is satisfied by any linear or convex cost function. Assumption 2-(v) says that the conditional hazard rate does not decline too rapidly as θ increases, where the term $D(\beta)$ represents a (θ, β) consumer's minimum acceptable internet speed. Most commonly used unimodal distributions satisfy the hazard rate assumption. Assumption 2-(vi) says that the willingness-to-pay for phone service θ does not matter to the consumer unless his phone usage is positive. Thus internet-only users have no parameter θ in their utility function. I call this assumption "weak complementarity". Consequently, the firm can charge a fixed fee $v(0, j)$ and leave no rent to internet-only users. Finally, Assumption 2-(vi) also implies that serving internet as a separate product is profitable for the firm.

3.2 Characterization of the Optimal Selling Mechanism

The approach I adopt is close to the separability case discussed in Rochet and Stole (2003). Specifically, I partition the set of types into one-dimensional subsets and reduce the multidimensional problem to a unidimensional problem. To clarify ideas, I first study the case when the provider observes β. Thus the problem becomes unidimensional and I explicitly characterize the optimal selling mechanism. I then show that this mechanism is still optimal under a standard affiliation assumption when β is not observed, thereby solving the multidimensional screening problem.

I make the following assumption on the utility and the cost functions.

Assumption 3: *For all $\theta \in [\underline{\theta}, \overline{\theta}]$, $q \in \mathbb{R}^+$ and $j \in \mathcal{J}$, $U(\cdot, \cdot; \cdot)$ and $c(\cdot, \cdot)$ satisfy*

(i) The utility function is additively separable as follows

$$U(q, j; \theta) = u(q; \theta) + v(q, j),$$

where $u(\cdot; \cdot)$ satisfies $u(0; \theta) = 0$.[9]

(ii) For all $\tilde{j} > j$, where $\tilde{j} \in \mathcal{J}$

$$\begin{aligned} U_q(q, \tilde{j}; \theta) - U_q(q, j; \theta) &\leq c_q(q, \tilde{j}) - c_q(q, j), \\ v(0, \tilde{j}) - v(0, j) &\leq c(0, \tilde{j}) - c(0, j). \end{aligned}$$

Assumption 3-(i) borrows from Sundararajan (2003) and Chen and Luo (2012) in the context of nonlinear pricing with network effects where j is replaced by the total quantity of product used by all consumers in the market $Q = \int_{\underline{\theta}}^{\overline{\theta}} q(\theta) f(\theta) d\theta$. I remark that the cross derivative $U_{qj}(\cdot, \cdot; \cdot)$ becomes independent of θ. Although the interaction between internet and phone services is the same for consumers with different tastes, I allow this interaction to

[9]The utility $U(q, j; \theta) = u(q; \theta) + v(q, j) + \omega(\theta, j)$ is more general but does not satisfy the weak complementarity assumption. See Assumption 2-(vi).

vary with the bundle. Hereafter, I call $u(\cdot;\cdot)$ the intrinsic utility function for phone service, and $v(\cdot,\cdot)$ the complementary utility function for the bundle. However, I do not impose any sign restriction on the cross derivative of $v(\cdot,\cdot)$.[10]

Assumption 3-(ii) says that the increment of marginal cost for phone service is larger than the marginal utility when one increases the internet level. Similarly, the cost increment for serving internet only is larger than the corresponding incremental utility. It implies $U(q,\tilde{j};\theta) - c(q,\tilde{j}) \leq U(q,j;\theta) - c(q,j)$. Hence, when the provider observes (θ,β), he always prefers to assign the minimum internet speed the consumer can accept.

Solving the Model when β is Observed

When β is observed, the provider solves the following problem for each value of β

$$\begin{aligned}
\max_{t(\cdot,\beta),q(\cdot,\beta),j(\cdot,\beta)} \quad & \int_{\underline{\theta}}^{\overline{\theta}} \Big[t(\theta,\beta) - c(q(\theta,\beta),j(\theta,\beta))\Big] f(\theta|\beta)d\theta \\
s.t. \quad & U(q(\theta,\beta),j(\theta,\beta);\theta) - t(\theta,\beta) \geq U(q(\tilde{\theta},\beta),j(\tilde{\theta},\beta);\theta) - t(\tilde{\theta},\beta), \\
& U(q(\theta,\beta),j(\theta,\beta);\theta) - t(\theta,\beta) \geq 0, \\
& j(\theta,\beta) \geq \beta,
\end{aligned}$$

for all (θ,β) and $(\tilde{\theta},\beta)$ such that $j(\tilde{\theta},\beta) \geq \beta$. Since the consumer cannot misreport β, the two-dimensional IC and IR constraints reduce to one-dimensional constraints. Moreover, the information structure reduces to the conditional density $f(\theta|\beta)$. The problem then becomes a multiproduct nonlinear pricing problem in which a consumer's private information is one-dimensional. I denote the optimal selling mechanism as $\{t^*(\cdot,\cdot), q^*(\cdot,\cdot), j^*(\cdot,\cdot)\}$.

I can now show that the problem reduces to several standard single product nonlinear pricing problems. I need first to show that the provider will always assign the minimum internet speed $D(\beta)$ that the consumer can accept when only β is observed. This gives the following lemma.

Lemma 1: *Under Assumptions 1 and 3,* $j^*(\theta,\beta) = D(\beta)$, $q^*(\theta,\beta) = q^*\Big(\theta, D(\beta)\Big)$, *and* $t^*(\theta,\beta) = t^*\Big(\theta, D(\beta)\Big)$ *for all* $(\theta,\beta) \in \Theta$.

Lemma 1 says that β only affects the optimal selling mechanism through the step function $D(\beta)$.[11] Thus consumers having the same minimum acceptable internet speed $D(\beta)$ face the same phone service assignment $q^*(\cdot,\cdot)$, the same internet assignment $j^*(\cdot,\cdot)$ and price schedule $t^*(\cdot,\cdot)$. The following lemma characterizes the optimal mechanism.

[10] Liu, Chintagunta, and Zhu (2010) find evidence of strong complementarity between local phone consumption and DSL internet. In view of their results, I call $v(\cdot,\cdot)$ the complementary utility function.

[11] By considering non-random nonlinear pricing schedules of the form $T(x,j)$, consumers' report of β is constrained to belong to $\mathcal{J}$.

Lemma 2: *Under Assumptions 1, 2 and 3, for any given value of β, the optimal phone service assignment $q^*(\cdot,\beta)$ and price schedule $t^*(\cdot,\beta)$ satisfy:*

(i) There exists a cutoff taste $\theta^c_\beta \in [\underline{\theta},\overline{\theta}]$, above which consumers are assigned a bundle with internet $D(\beta)$ and phone service, and below which they are assigned internet $D(\beta)$ only.

(ii) If $\theta \in [\theta^c_\beta,\overline{\theta}]$, $q^*(\cdot,\beta)$ and $t^*(\cdot,\beta)$ are solution of

$$U_q\Big[q^*(\theta,\beta),D(\beta);\theta\Big] = c_q\Big[q^*(\theta,\beta),D(\beta)\Big] + U_{q\theta}\Big[q^*(\theta,\beta),D(\beta);\theta\Big]\frac{1-F[\theta|D(\beta)]}{f[\theta|D(\beta)]}, \tag{1}$$

$$t^*(\theta,\beta) = U\Big[q^*(\theta,\beta),D(\beta);\theta\Big] - \int_{\theta^c_\beta}^{\theta} U_\theta\Big[q^*(x,\beta),D(\beta);x\Big]dx. \tag{2}$$

Moreover, the cutoff taste θ^c_β is defined as

$$\theta^c_\beta \equiv \min\Big\{\theta \in [\underline{\theta},\overline{\theta}] : M\Big(\theta,D(\beta)\Big) \geq 0\Big\}, \tag{3}$$

where $M(\theta,j) \equiv [U(q^*(\theta,j),j;\theta) - v(0,j)] - [c(q^*(\theta,j),j) - c(0,j)] - U_\theta(q^*(\theta,j),j;\theta)\frac{1-F(\theta|j)}{f(\theta|j)}$.

(iii) If $\theta \in [\underline{\theta},\theta^c_\beta)$, $q^*(\theta,\beta) = 0$ and $t^*(\theta,\beta) = v\Big(0,D(\beta)\Big)$.

The proof is in two steps. In a first step, I derive the optimal selling mechanism conditionally on serving consumers with a willingness-to-pay equal or above an arbitrary cutoff value θ^c. Thus the optimal selling mechanism is defined by (1) and (2) by replacing θ^c_β with θ^c. An important feature is that the phone service assignment $q(\cdot,\cdot)$ does not depend on θ^c while the price schedule $t(\cdot,\cdot)$ is decreasing in θ^c. In a second step, I find the optimal cutoff value θ^c that maximizes the profit. It is defined by (3). Intuitively, when the firm slightly lowers θ^c, the term $[U(q^*(\theta^c,j),j;\theta^c) - v(0,j)]$ is the incremental utility for a (θ^c,j) consumer switching from internet only to a bundle, $[c(q^*(\theta^c,j),j) - c(0,j)]$ is the incremental cost for the firm, and $U_\theta(q^*(\theta^c,j),j;\theta^c)$ is the additional informational rent everyone in the customer base gets. Therefore, the term $\{[U\Big(q^*(\theta^c,j),j;\theta^c) - v(0,j)] - [c(q^*(\theta^c,j),j) - c(0,j)]\}f(\theta^c|j)$ is the marginal gain for expanding the customer base by lowering θ^c, while $U_\theta(q^*(\theta^c,j),j;\theta^c)[1-F(\theta^c|j)]$ is the corresponding marginal loss for reducing the tariff to every consumer above θ^c. Equation (3) balances these two effects. In addition, it is easy to see that $\theta^c_\beta = \theta^c_j$, where $j = D(\beta)$. The term β only affects the cutoff taste through the step function $D(\beta)$. My results are in the spirits of Armstrong and Rochet (1999) recommendation where they advise to discretize the type space to simplify the multidimensional screening problem.

Equations (1) and (2) define the phone service assignment $q^*(\cdot,\cdot)$ and price schedule $t^*(\cdot,\cdot)$ for the bundle users. In particular, (1) says that the marginal payoff for each type equals the marginal cost plus a nonnegative distortion term due to incomplete information. Intuitively,

$\{U_q[q^*(\theta,\beta),D(\beta);\theta]-c_q[q^*(\theta,\beta),D(\beta)]\}f(\theta|D(\beta))$ represents the firm's desire to implement an efficient allocation weighted by the density while $U_{q\theta}(q^*(\theta,\beta),D(\beta);\theta)[1-F(\theta|D(\beta))]$ represents the informational rent the firm has to give up to the consumer for revealing their private information. Equation (2) says that the payment equals the consumer's utility minus some informational rent. Following Assumption 2, the resulting usage-based tariffs $T^*(\cdot,j)$ are increasing and concave for all j.

If a consumer is excluded from consuming phone service, then his utility function becomes $U(q,j;\theta)=v(0,j)$ by Assumption 2-(vi). Since there is no asymmetric information, the firm can take all the consumer surplus by charging $T^*(0,j)=v(0,j)$, leading to $t^*(\theta,\beta)=v(0,D(\beta))$, for any $(\theta,\beta)\in\Theta$ such that $\theta\in[\underline{\theta},\theta^c_\beta)$.

Solving the Model when both θ and β are Unobserved

I now show that the previous mechanism is still optimal when both θ and β are unobserved. This constitutes an interesting result given the complexity of multidimensional screening problems.

Before showing that the second-best mechanism is the one given in Lemma 2, I make an affiliation assumption on the joint distribution of θ and β.

Assumption 4: *$\forall\theta\in[\underline{\theta},\overline{\theta}]$ and $\forall j\in\mathcal{J}$, $\frac{1-F(\theta|D(\beta)=j)}{f(\theta|D(\beta)=j)}$ is increasing in j.*

Assumption 4 follows Che and Gale (2000). It is equivalent to assume $H(\theta|j)$ be decreasing in j. Intuitively, a consumer is relatively more likely to have a higher willingness-to-pay for phone service if he needs a higher speed internet. According to Horrigan (2010), the Federal Communications Commission 2009 survey shows that higher speed internet adopters tend to be better educated with higher incomes. Since phone service is a normal good, I consider it as a reasonable assumption.

First, I establish the desirability of the minimum acceptable internet speed. Let $\{t^{sb}(\cdot,\cdot),q^{sb}(\cdot,\cdot),j^{sb}(\cdot,\cdot)\}$ be the optimal selling mechanism when both θ and β are private information. The following lemma parallels Lemma 1.

Lemma 3: *Under Assumptions 1 and 3, $j^{sb}(\theta,\beta)=D(\beta)$, $q^{sb}(\theta,\beta)=q^{sb}\Big(\theta,D(\beta)\Big)$, and $t^{sb}(\theta,\beta)=t^{sb}\Big(\theta,D(\beta)\Big)$ for all $(\theta,\beta)\in\Theta$.*

Second, I simplify the firm's problem under Assumption 1. The basic idea is to reduce the number of binding IC constraints. Since the consumer can either overreport, underreport or report truthfully each parameter of his private information. The number of potential deviations increases to eight in a two-dimensional problem from two in a unidimensional problem. However, I show that only three deviations matter as stated in the next lemma.

Lemma 4: *Following Assumption 1, when both θ and β are private information, the IC constraints are satisfied if and only if the following two one-dimensional IC constraints are satisfied. Namely,*

$$U(q(\theta,\beta),D(\beta);\theta)-t(\theta,\beta)\geq U(q(\tilde{\theta},\beta),D(\beta);\theta)-t(\tilde{\theta},\beta) \quad \forall\theta,\beta,\tilde{\theta},$$
$$U(q(\theta,\beta),D(\beta);\theta)-t(\theta,\beta)\geq U(q(\theta,\tilde{\beta}),D(\tilde{\beta});\theta)-t(\theta,\tilde{\beta}) \quad \forall\theta,\beta,\tilde{\beta} \text{ such that } D(\tilde{\beta})\geq\beta.$$

Following Assumption 1, the (θ,β) and $(\theta,\tilde{\beta})$ consumers have the same preferences over outcomes as long as their minimum needs for internet are satisfied. If the two one-dimensional IC constraints above are true, then the (θ,β) consumer does not want to pretend to be $(\theta,\tilde{\beta})$, and the $(\theta,\tilde{\beta})$ consumer does not want to pretend to be $(\tilde{\theta},\tilde{\beta})$. Therefore, by transitivity, the (θ,β) consumer has no incentive to pretend to be $(\tilde{\theta},\tilde{\beta})$. Thus consumers report truthfully. Intuitively, if one considers θ on the x-axis and β on the y-axis, the potential deviations can be horizontal for θ and vertical for β. Lemma 4 tells us that the only binding constraints are only upward for β and upward and downward for θ, while the other deviations are redundant, thereby reducing the number of binding constraints to three.

Third, I show that if the firm implements $\{t^*(\cdot,\cdot),q^*(\cdot,\cdot),j^*(\cdot,\cdot)\}$, the consumer has no incentive to misreport his/her internet need β given a willingness-to-pay θ.

Lemma 5: *Under Assumptions 1, 2, 3 and 4, $q^*(\theta,\beta)$ is decreasing in β and θ^c_β is increasing in β. Moreover, $T^*(q,j)-v(q,j)$ is increasing in j.*

The intuition is as follows. Following Assumption 4, the firm knows that a consumer is more likely to have a higher taste for phone service when he needs a higher speed internet. By exploiting this positive dependence, it can charge the consumer more, adjusted for complementary utility, when the consumer chooses a higher j. This in turn implies that the subscriber would consume less phone service giving a decreasing assignment $q^*(\cdot,\cdot)$ in β. While the optimal exclusion is the result of a trade-off between the marginal gain and the loss of expanding the customer base, Assumption 4 favors the latter as one moves from a low speed internet to a high speed one. Therefore, it becomes more profitable to exclude a larger range of low taste consumers if they adopt a higher speed internet. This explains why the cutoff value θ^c_β is increasing in β. Under Assumption 3, a (θ,β) consumer solves

$$\max_{(q,j):j\geq\beta} \quad u(q;\theta)+v(q,j)-T^*(q,j).$$

Following Lemma 5, because $T^*(q,j)-v(q,j)$ is increasing in j, the additional payment for a higher internet level is larger than the additional utility it brings. Therefore, the consumer will choose the minimum internet speed meeting his needs.

Finally, we are now in a position to show that $\{t^*(\cdot,\cdot), q^*(\cdot,\cdot), j^*(\cdot,\cdot)\}$ is the optimal mechanism when both θ and β are private information. When β is observed, the firm's profit is weakly higher than when β is unobserved. Thus, I only need to show that the two-dimensional IC and IR constraints still hold if the provider uses $\{t^*(\cdot,\cdot), q^*(\cdot,\cdot), j^*(\cdot,\cdot)\}$. Regarding the IR constraints, they are satisfied automatically because they are the same. Following Lemma 4, the IC constraints are satisfied as long as the two one-dimensional IC constraints are. On one hand, misreporting θ is not profitable because of $\{t^*(\cdot,\cdot), q^*(\cdot,\cdot), j^*(\cdot,\cdot)\}$. On the other hand, misreporting β is not profitable either following Lemma 5. The following proposition summarizes these results.

Proposition 1: *Under Assumptions 1, 2, 3 and 4, we have* $t^*(\cdot,\cdot) = t^{sb}(\cdot,\cdot)$, $q^*(\cdot,\cdot) = q^{sb}(\cdot,\cdot)$ *and* $t^*(\cdot,\cdot) = j^{sb}(\cdot,\cdot)$.

Armstrong and Rochet (1999) remark that phenomena such as bunching and exclusion that arise in multiproduct nonlinear pricing models create technical difficulties, making it hard to generate closed-form solutions. Following Proposition 1, the optimal selling mechanism reduces to a combination of optimal selling mechanisms for a series of one-dimensional problems. Therefore, I characterize explicitly the optimal exclusion, the assignment schedules and the tariff functions. In my model, bunching arises at equilibrium through $D(\beta)$ and because people with low taste for phone service will consume internet only.

Bundling Decisions

In view of Lemma 2 and Proposition 1, consumers are segmented into several groups based on their internet needs and tastes for phone service. All consumers with β such that $D(\beta) = j$ uses the same internet level, j. I refer to them as group j. Consumers in group j are further divided into two parts according to their values of θ, namely, $[\underline{\theta}, \theta_j^c)$ and $[\theta_j^c, \overline{\theta}]$. The former or low taste subscribers will consume internet service j only, while the latter or high taste subscribers consume the bundle with $q > 0$ and internet service j. I call these two groups internet j users and bundle j users, respectively. In addition, the firm proposes J usage-based tariffs to bundle users ($q > 0$ and $j > 0$) and phone-only users ($q > 0$ and $j = 0$). It proposes $J - 1$ fixed fees to internet-only users ($q = 0$ and $j > 0$). These correspond to the data I will analyze in Section 5.

The results on optimal exclusion have important implications on bundling choices. For instance, the firm will sell internet j separately if $\theta_j^c = \overline{\theta}$. Similarly, he will sell internet only in a bundle if $\theta_j^c = \underline{\theta}$. For any other value of $\theta_j^c \in (\underline{\theta}, \overline{\theta})$, the firm will propose both. Thus, my model is general as it allows the three possible incentives to bundle, namely utility complementarity, cost efficiency and correlation between θ and β. To some extent, my model confirms Schmalensee (1984) and Fang and Norman (2006) results, i.e. the higher the cost or the lower the valuation, the less likely that bundling dominates component pricing. Moreover, my model admits a variety of bundling outcomes, including component pricing, pure bundling, semi-mixed bundling and mixed bundling.

When it is too costly to provide phone service to internet users, i.e. $\underline{\beta} \leq 0$ and $\theta_j^c = \overline{\theta}$ for all $j > 0$, the firm can exclude all internet users from consuming phone service. In this case, the firm would sell internet and phone services separately, which is known as component pricing (CP). When it is optimal for the firm to exclude some but not all internet users from consuming phone service, the provider wound sell internet and phone services not only as separate products but also in a bundle. If all the possible combinations of internet and phone services are offered, i.e. $\underline{\beta} \leq 0$ and $\theta_j^c \in (\underline{\theta}, \overline{\theta})$ for all $j \in \mathcal{J}$, the firm implements mixed bundling (MB). If some combination is not offered, the firm implements semi-mixed bundling (SMB). Moreover, if we let $\underline{\beta} > 0$, pure bundling (PB) arises when it is optimal for the firm to serve phone service to everyone, i.e. $\underline{\beta} > 0$ and $\theta_j^c = \underline{\theta}$ for all $j > 0$.

4 Identification and Estimation

In view of Section 3, the optimal mechanism is defined by (1), (2) and (3). Because the data display mixed bundling, I focus on this case. However, the results below can be readily adapted to the other cases of bundling. It is useful to recall the model structure and the observables. The model primitives are $\{u(\cdot;\cdot), v(\cdot,\cdot), F(\cdot,\cdot), c(\cdot,\cdot)\}$, which are the intrinsic utility function from consuming phone service, the complementary utility function from consuming internet and phone services, the joint distribution of consumers' types and the firm's cost function. Because j can take values in $\{0,1,2\}$, the model primitives become $\{u(\cdot;\cdot), v_j(\cdot), F_j(\cdot), c_j(\cdot)\}$, where $v_j(\cdot) \equiv v(\cdot, j)$, $F_j(\cdot) \equiv F(\cdot|D(\beta) = j)$ and $c_j(\cdot) \equiv c(\cdot, j)$ for $j = 0,1,2$. Regarding the observables, following Section 2, I observe the tariff $T_j(\cdot)$ for $j = 0,1,2$. Moreover, data on internet-only users provide information on $T_j(0)$ for $j = 1,2$. Since I observe q, we have the distribution of consumption $G_j^*(\cdot)$ for $j = 0,1,2$ and $q > 0$ as well as $G_j(0)$ from internet-only users. To summarize, the observables are $\{T_j(\cdot), G_j^*(\cdot)\}$ for $q > 0$ and $j = 0,1,2$ and $\{T_j(0), G_j(0)\}$ for $j = 1,2$.

4.1 Identification

To identify the model primitives, I exploit the variation offered by the data across the different groups of users, including those using either internet or phone only, in addition to the first-order conditions (1), (2) and (3). I proceed in several steps. First, I study which primitives the data on internet-only users and phone-only users will allow me to identify.

While assuming multiplicative separability of the intrinsic utility function and linearity of the cost function, I will show that the intrinsic utility function, the marginal cost parameter as well as the conditional density of θ for $j = 0$ are identified. Second, I investigate the identification of the complementary utility function and the conditional densities of θ for $j = 1, 2$ from the bundle users data. Third, I show how to exploit the firm's optimal exclusion conditions to identify the fixed cost parameters. Therefore, the optimality of tariffs and bundling as well as one-to-one mapping at the equilibrium between the consumption of phone service and the unknown type θ will play crucial roles. [12]

Identifying Assumptions

I make the following identifying assumptions on the model primitives. Hereafter, the prime denotes a derivative with respect to q.

Assumption 5: *For all $\theta \in [\underline{\theta}, \overline{\theta}]$ and $q \in \mathbb{R}^+$,*

(i) The intrinsic utility $u(q;\theta)$ satisfies

$$u(q;\theta) = \theta u_0(q),$$

with $u_0(0) = 0$, $u_0'(q) \geq 0$ and $u_0''(q) \leq 0$.

(ii) The cost function is of the form

$$c_j(q) = \kappa_0 \mathbb{1}(q > 0) + \kappa_j \mathbb{1}(j > 0) - \Delta_j \mathbb{1}(qj > 0) + \gamma q,$$

where $\gamma > 0$, $\kappa_0 > 0$, $\kappa_j > 0$, and $\Delta_j \geq 0$ for $j = 1, 2$.

(iii) $v_0(q) = 0$.

Following the literature, I assume multiplicative separability of the intrinsic utility function in the type θ as stated in Assumption 5-(i). Thus, I interpret $u_0(\cdot)$ as the base intrinsic utility function. However, Assumption 5-(i) will not be sufficient to achieve identification. I provide an intuitive argument. Equations (1) and (2) provide $2J$ one-to-one mappings between θ and q and between t and q, respectively. On the other hand, I have to identify J cost functions $c_j(\cdot)$, J complementary utility functions $v_j(\cdot)$, and J conditional type distributions $F_j(\cdot)$ as well as the base intrinsic utility function $u_0(\cdot)$. It is clear that additional restrictions need to be imposed.

[12]From the auction literature, the one-to-one mapping at the equilibrium plays a crucial role to identify the model primitives. See Guerre, Perrigne, and Vuong (2000) and Athey and Haile (2007). In a single product nonlinear pricing model, Luo, Perrigne, and Vuong (2014) show that the optimality of tariff in addition to the one-to-one mapping between the observed quantity and the unobserved taste are both needed to identify the model primitives. The multidimensional screening problem adds additional difficulties. See e.g. Luo, Perrigne, and Vuong (2012). In the context of insurance, Aryal, Perrigne, and Vuong (2009) exploit a repeated outcome, i.e. the number of accidents, to identify the model structure.

Several identifying assumptions can be entertained. Following Section 2, the production of telecommunication services tends to involve high fixed costs and small marginal costs. Therefore, I assume a linear cost function as stated in Assumption 5-(ii). The term κ_0 is a fixed cost associated to any positive production of phone service, while κ_j is a fixed cost to provide internet level j. The term Δ_j is the difference between the sum of these two fixed costs and the fixed cost to sell a bundle. It measures the cost saving effect of selling internet and phone services as a bundle.

Lastly, Assumption 5-(iii) says that there is no complementary utility when there is no consumption of internet. It implies that $v_0'(q) = 0$ for all $q \in \mathbb{R}^+$.[13] I can then rewrite the first-order conditions (1) and (2) defining the optimal selling mechanism. Namely,

$$\theta u_0'(q) + v_j'(q) = \gamma + u_0'(q)\frac{1 - F_j(\theta)}{f_j(\theta)}, \qquad \forall q \in [\underline{q}_j, \overline{q}_j] \tag{4}$$

$$T_j'(q) = \theta u_0'(q) + v_j'(q), \qquad \forall q \in [\underline{q}_j, \overline{q}_j] \tag{5}$$

where $\underline{q}_j = q^*(\theta_j^c, j)$ and $\overline{q}_j = q^*(\overline{\theta}, j)$ for $j = 0, 1, 2$. Together with the boundary conditions $T_j(\underline{q}_j) = \theta_j^c u_0(\underline{q}_j) + v_j(\underline{q}_j)$ and the cutoff tastes in (3), (4) and (5) define the optimal mechanism.

Identification of γ, $v_j(0)$, $F_j(\theta_j^c)$, $F_0(\cdot)$ and $u_0(\cdot)$

Under Assumption 5-(ii), the marginal variable cost enters in (4). Thus, γ is identified from (4) and (5) evaluated at the maximum phone usage. This gives $\gamma = T_j'(\overline{q}_j)$, $\forall j \in \{0, 1, 2\}$. The identification of the fixed cost parameters κ_0, κ_j and Δ_j will be shown later.

From the model of Section 3, internet-only users pay $T_j(0) = v_j(0)$ for any $j = 1, 2$, which renders the identification of $v_j(0)$ immediate. Moreover, the proportion of individuals using internet only among their group of users gives $F_j(\theta_j^c)$ for $j = 1, 2$. These results are summarized in the following lemma while the identification of θ_j^c will be addressed later.

Lemma 6: *γ is identified. $v_j(0)$ and $F_j(\theta_j^c)$ are identified for $j = 1, 2$.*

I now turn to data from phone-only users. Identification in this group reduces to the single product nonlinear pricing model studied by Luo, Perrigne, and Vuong (2014). In particular, the first-order conditions (4) and (5) for $j = 0$ are equivalent to

$$u_0'(q) = T_0'(q)\xi(q)/\overline{\theta}, \tag{6}$$

$$\theta_0(q) = \overline{\theta}/\xi(q), \tag{7}$$

[13] An alternative normalization would be to assume a similar value for another internet level instead.

where

$$\xi(q) = \Big[1 - G_0^*(q)\Big]^{1-\frac{\gamma}{T_0'(q)}} \exp\Big\{\gamma \int_q^{\overline{q}_0} \frac{T_0''(x)}{T_0'(x)^2} \log\Big[1 - G_0^*(x)\Big] dx\Big\}, \tag{8}$$

with $\gamma = T_0'(\overline{q}_0)$ and $q \in [\underline{q}_0, \overline{q}_0]$.

Equations (6) and (7) show that the marginal intrinsic utility $u_0'(\cdot)$ and the unobserved taste for phone service θ for phone-only users are identified up to a constant. In view of this, a natural normalization is $\overline{\theta}$=1.

Assumption 6: $\overline{\theta} = 1$.

Under such a normalization, $u_0(\cdot)$ can be interpreted as the intrinsic utility function for the highest taste. Since $\theta_0^c = \theta_0(\underline{q}_0)$, θ_0^c is identified. Moreover, I can further identify $u_0(\cdot)$ using the boundary condition $T_0(\underline{q}_0) = \theta_0^c u_0(\underline{q}_0)$. Namely,

$$u_0(q) = \frac{T_0(\underline{q}_0)}{\theta_0^c} + \int_{\underline{q}_0}^{q} u_0'(x) dx, \tag{9}$$

for all $q \in [\underline{q}_0, \overline{q}_0]$. The next proposition summarizes these results.

Proposition 2: *Under Assumptions 1-6, the base intrinsic utility function $u_0(\cdot)$ and the type $\theta_0(\cdot)$ are identified on $[\underline{q}_0, \overline{q}_0]$. Moreover, the truncated conditional taste distribution $F_0^*(\cdot)$ is identified on $[\theta_0^c, \overline{\theta}]$, while the conditional taste distribution $F_0(\cdot)$ is identified up to a constant on $[\theta_0^c, \overline{\theta}]$.*

The distribution $F_0(\cdot)$ is identified up to a constant because I do not observe the proportion of consumers who do not buy internet and phone services. Moreover, usage and payment data do not provide any variation to identify $u(\cdot)$ and $F_0(\cdot)$ on $[0, \underline{q}_0)$ and $[\underline{\theta}, \theta_0^c)$, respectively.

Identification of $v_j(\cdot)$ and $F_j(\cdot)$

I now turn to data from bundle j users to address the identification of $v_j(\cdot)$ and $F_j(\cdot)$ for $j = 1, 2$. My proof of identification is contructive. I exploit the one-to-one mapping between phone usage q and taste θ for bundle users, which implies for each $q \in [\underline{q}_j, \overline{q}_j]$, $G_j^*(q) = [F_j(\theta) - F_j(\theta_j^c)]/[1 - F_j(\theta_j^c)]$. Taking the derivative gives $g_j^*(q) = \theta_j'(q) f_j(\theta)/[1 - F_j(\theta_j^c)]$. Thus the inverse hazard rate becomes $[1 - F_j(\theta)]/f_j(\theta) = \theta_j'(q)[1 - G_j^*(q)]/g_j^*(q)$.

From (4), replacing the left-hand side by $T_j'(q)$ and $[1 - F_j(\theta)]/f_j(\theta)$ by $\theta_j'(q)[1 - G_j^*(q)]/g_j^*(q)$ in the right-hand side gives

$$\theta_j'(q) = \frac{T_j'(q) - \gamma}{u_0'(q)} \frac{g_j^*(q)}{1 - G_j^*(q)}.$$

Integrating both sides from q to $\overline{q}_j$ leads to

$$\theta_j(q) = 1 - \int_q^{\overline{q}_j} \frac{T_j'(x) - \gamma}{u_0'(x)} \frac{g_j^*(x)}{1 - G_j^*(x)} dx, \tag{10}$$

where the normalization of Assumption 6 is used. Equation (10) shows that $\theta_j(\cdot)$ is identified wherever both $u_0'(\cdot)$ is identified and $g_j^*(\cdot)$ is observed. While $\overline{q}_j \leq \overline{q}_0$ by Lemma 5, it is not necessary that $\underline{q}_j \geq \underline{q}_0$. For convenience, I assume that $\underline{q}_j \geq \underline{q}_0$ hereafter.[14]

Once $\theta_j(\cdot)$ is identified, $v_j'(\cdot)$ is identified by plugging (10) into (5). That is,

$$v_j'(q) = T_j'(q) - \theta_j(q) u_0'(q). \tag{11}$$

By Equation (10), θ_j^c is identified as $\theta_j^c = \theta_j(\underline{q}_j)$. Using the boundary condition $T_j(\underline{q}_j) = \theta_j^c u_0(\underline{q}_j) + v_j(\underline{q}_j)$, the complementary utility function $v_j(\cdot)$ is identified as

$$v_j(q) = T_j(\underline{q}_j) - \theta_j^c u_0(\underline{q}_j) + \int_{\underline{q}_j}^q v_j'(x) dx. \tag{12}$$

The following proposition summarizes these results.

Proposition 3: *Under Assumptions 1-6, the complementary utility function $v_j(\cdot)$ and the type $\theta_j(\cdot)$ are identified on $[\underline{q}_j, \overline{q}_j]$, where $j = 1, 2$. Moreover, the truncated conditional taste distribution $F_j^*(\cdot)$ is identified on $[\theta_j^c, \overline{\theta}]$, while the conditional taste distribution $F_j(\cdot)$ is identified on $[\theta_j^c, \overline{\theta}]$.*

The type distribution $F_j(\cdot)$ can be recovered from $F_j^*(\cdot)$ on $[\theta_j^c, \overline{\theta}]$ because (i) I observe the proportion of consumers whose phone usage is less than $\underline{q}_j$, $F_j(\theta_j^c)$, and (ii) $F_j(\cdot) = F_j(\theta_j^c) + [1 - F_j(\theta_j^c)] F_j^*(\cdot)$. On the other hand, usage and payment data do not provide any variation to identify $v_j(\cdot)$ and $F_j(\cdot)$ on $[0, \underline{q}_j)$ and $[\underline{\theta}, \theta_j^c)$, respectively.

To understand better the intuition behind these results, I consider alternative expressions for (10) and (11). Using (5), (10) can be rewritten equivalently as

$$\theta_j(q) = 1 - \int_q^{\overline{q}_j} \frac{T_j'(x) - \gamma}{T_0'(x)} \frac{g_j^*(x)}{1 - G_j^*(x)} \theta_0(x) dx. \tag{13}$$

Intuitively, the difference between a consumer's taste and the highest taste is the weighted average of $\theta_0(\cdot)$ over $[q, \overline{q}_j]$, where the weight is determined by the slope of the tariff functions and the conditional distribution of phone usage.

[14] My data confirm this. Otherwise, bounds can be derived.

Similarly, (11) can be rewritten equivalently as

$$v_j'(q) = T_j'(q) - \frac{\theta_j(q)}{\theta_0(q)} T_0'(q).$$

Here again, the weighted shape difference between the two tariff functions $T_j(\cdot)$ and $T_0(\cdot)$ is used to recover the complementary utility function $v_j(\cdot)$. The weight is determined by the ratio of corresponding tastes $\theta_j(q)/\theta_0(q)$.

Identification of κ_0, κ_j and Δ_j

It remains to address the identification of the fixed cost parameters κ_0, κ_j and Δ_j. Among the equilibrium conditions, only the cutoff tastes involve the fixed costs. First, I use information from phone-only user data to identify κ_0. Second, I exploit information from bundle users data to identify Δ_1 and Δ_2. However, κ_1 and κ_2 remain not identified. By Assumption 2-(vi), κ_1 and κ_2 are bounded by the monthly fees for internet-only users. The following proposition formalizes these results.

Proposition 4: *Under Assumptions 1-6, we have*

(i) The parameter κ_0 is identified as $\kappa_0 = \frac{\gamma}{T_0'(\underline{q}_0)} T_0(\underline{q}_0) - \gamma \underline{q}_0$.

(ii) The parameters Δ_1 and Δ_2 are identified as

$$\Delta_j = \left[\frac{\gamma}{T_0'(\underline{q}_0)} T_0(\underline{q}_0) - \gamma \underline{q}_0\right] - \left[T_j(\underline{q}_j) - T_j(0) - \gamma \underline{q}_j\right] + \theta_0(\underline{q}_j) u_0(\underline{q}_j) \frac{T_j'(\underline{q}_j) - \gamma}{T_0'(\underline{q}_j)}, \quad \textit{for } j = 1, 2$$

(iii) The parameters κ_1 and κ_2 are bounded, i.e. $\kappa_j \leq T_j(0)$ for $j = 1, 2$.

Regarding the identification of the fixed cost parameters, following (3), some consumers switch from bundle-j to internet-j as the firm lowers the cutoff taste. As a result, the difference in the fixed cost $(\kappa_0 + \kappa_j - \Delta_j) - \kappa_j$ affects the optimality of the cutoff tastes. Thus κ_0 and Δ_j relate to the utility, cost and inverse hazard rate at the cutoff values. The latter are identified from following Propositions 2 and 3 as discussed above, thereby identifying κ_0 and Δ_j.

4.2 Estimation

My proof of identification is constructive and can be used to derive a semiparameteric estimator. Since the estimation of $v_j(0)$ and $F(\theta_j^c)$ can be calculated directly from internet-only users data, I focus on the estimation of all the other primitives. I then use data from phone-only and bundle users. For convenience, I order them lexicographically by their consumption bundles. For all the consumers having a positive quantity of phone service, I group those using the same internet level $j = 0, 1, 2$ and then order them according to their phone usage q. I denote N_j^* as the number of users for group j. This would give $\{(q_j^i, t_j^i)\}_{i=1,\ldots,N_j^*}$, where $0 < q_j^1 \leq q_j^2 \leq \ldots \leq q_j^{N_j^*}$ and $0 < t_j^1 \leq t_j^2 \leq \ldots \leq t_j^{N_j^*}$.

I propose a three-step estimation procedure. First, I estimate γ and $\xi(\cdot)$ using $\gamma = T_0'(\overline{q}_0)$ and (8), respectively. An estimate for $\xi(\cdot)$ will allow me (i) to obtain an estimate of the marginal intrinsic utility function $u_0'(\cdot)$ using (6) and (ii) to construct a sample of pseudo tastes for phone-only users from (7). To complete the estimation of $u_0(\cdot)$, I will estimate θ_0^c and $u_0(\underline{q}_0)$ using $\theta_0^c = \theta_0(\underline{q}_0)$ and $T_0(\underline{q}_0) = \theta_0^c u_0(\underline{q}_0)$, respectively. Second, the estimated marginal intrinsic utility function is used to (i) estimate the marginal complementary utility functions $v_j'(\cdot)$ using (11) and (ii) to construct a sample of pseudo tastes for bundle-j users from (13). To complete the estimation of $v_j(\cdot)$, I will estimate θ_j^c, $u_0(\underline{q}_j)$ and $v_j(\underline{q}_j)$ using $\theta_j^c = \theta_j(\underline{q}_j)$ and $T_j(\underline{q}_j) = \theta_j^c u_0(\underline{q}_j) + v_j(\underline{q}_j)$, respectively. Third, I use the estimated pseudo tastes to estimate the conditional taste densities. Details can be found in Appendix B.

5 Empirical Analysis of China Telecom Data

5.1 Estimation Results

As discussed in Section 2, I estimate the tariff functions $T_0(\cdot)$, $T_1(\cdot)$ and $T_2(\cdot)$ and the resulting phone usage q. See Appendix B and Figure 3 displaying $T_0(\cdot)$, $T_1(\cdot)$ and $T_2(\cdot)$.

Regarding the cost, I obtain an estimate for the marginal variable cost γ which is equal to 5.95 cents. This is approximately a fourth of the average price charged per minute. The fixed cost for phone service is 6.58 RMB, while the estimate of the bound for the fixed cost of 1 Mbps (2 Mbps) internet κ_1 (κ_2) is 78 RMB (88 RMB). The estimate of the cost saving parameter for 1 Mbps internet Δ_1 is 2.60 RMB while this value increases to 3.27 RMB for Δ_2. The fixed cost for phone service seems to be small. It is approximately a tenth of the average bill of phone-only subscribers. These cost parameters suggest that the firm has a comfortable profit margin as discussed later. Relative to providing internet only, providing a bundle does not impose much additional cost to the firm as suggested by the estimates of Δ_1 and Δ_2. China Telecom mainly uses Asymmetric Digital Subscriber Line (ADSL) to provide internet service. Thus internet is transmitted through telephone lines. Moreover, fixed transaction costs such as mailing statement do not increase much because bills are merged if the consumer uses a bundle.

I then obtain estimates of the marginal intrinsic utility $u_0'(\cdot)$ and the marginal complementary utility functions $v_1'(\cdot)$ and $v_2'(\cdot)$. The first is displayed in Figure 1 while the latter two are displayed in Figure 2. The estimated marginal intrinsic utility $u_0'(\cdot)$ is positive and decreasing, thereby satisfying Assumption 5-(i). The estimated marginal complementary utility functions $\hat{v}_1'(\cdot)$ and $\hat{v}_2'(\cdot)$ are both negative and increasing with $\hat{v}_1'(\cdot)$ above $\hat{v}_2'(\cdot)$, thereby satisfying Assumption 3-(ii). Since both are negative, internet and phone services seem to be substitutes. Internet offers alternative communication tools such as email, skype and so on, which can explain the substitutability with phone service. Thus the utility of a bundle user is smaller than the sum of the utilities for a phone service user only and a internet user only. Moreover, this substitution effect is stronger with a higher level of internet because a faster internet service allows better alternative communication tools.

Figure 1: Marginal Intrinsic Utility $\hat{u}_0'(\cdot)$

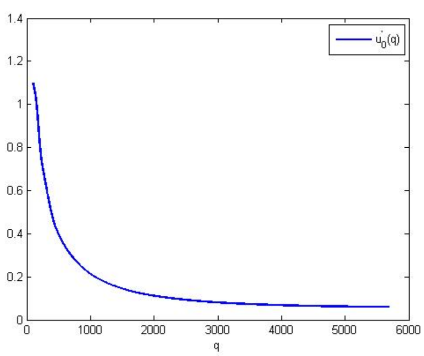


Figure 2: Marginal Complementary Utilities $\hat{v}_1'(\cdot)$, $\hat{v}_2'(\cdot)$

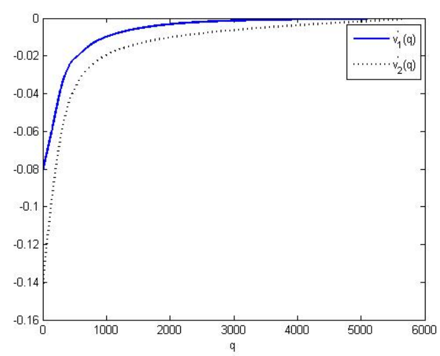


Figure 3 displays the inverse of the estimated $\theta_0(\cdot)$, $\theta_1(\cdot)$ and $\theta_2(\cdot)$. They are increasing in the type and decreasing in the internet choice j, thereby satisfying Lemma 5. As internet speed increases, a larger range of low taste (for phone service) consumers are excluded from using phone service. These observations satisfy the predictions of my model in Section 3. Figure 4 displays the estimated type densities $f_0^*(\cdot)$, $f_1^*(\cdot)$ and $f_2^*(\cdot)$. As the internet speed increases, the density function becomes less skewed to the left, thereby implying that consumers are more likely to have a higher taste for phone service. Here again, one sees the increase in the cutoff taste as the level of internet increases. Figure 5 displays the hazard rate functions $H_0(\cdot)$, $H_1(\cdot)$ and $H_2(\cdot)$. They are increasing in the type and decreasing in the internet choice j, thereby satisfying Assumptions 2-(v) and 4.

Using these estimated values, I can access empirically the firm's profit as well as the consumers' informational rents. The informational rent is estimated by $\hat{\theta}_j^i \hat{u}_0(q_j^i) + \hat{v}_j(q_j^i) - T_j(q_j^i)$. The ratio of the total informational rent across all consumers by the total amount paid is 29.76%. This measures the cost of asymmetric information. When considering by group of users, I find 53.27% for phone-only users, 27.02% for bundle users with 1 Mbps of internet level and 27.66% for bundle users with 2 Mbps. These overall rents tend to decrease with the level of internet. I recall that internet-only users do not enjoy any rent since the

Figure 3: Phone Service Assignments $\hat{q}_0(\cdot)$, $\hat{q}_1(\cdot)$, $\hat{q}_2(\cdot)$

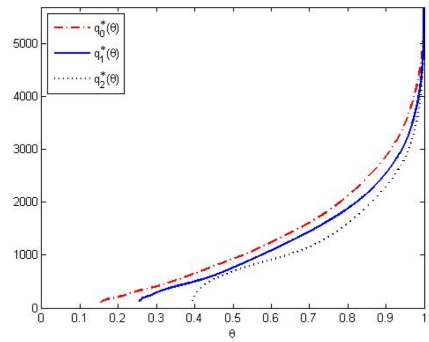


Figure 4: Conditional Type Densities $\hat{f}_0^*(\cdot)$, $\hat{f}_1^*(\cdot)$, $\hat{f}_2^*(\cdot)$

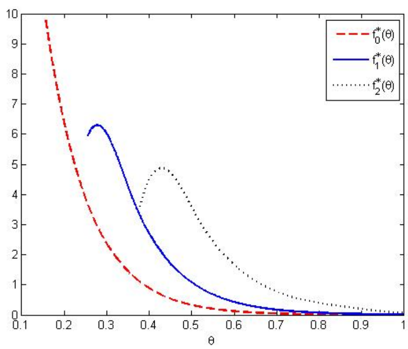
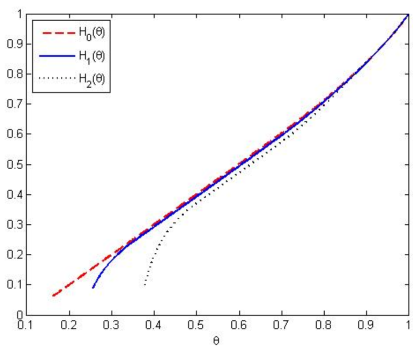


Figure 5: $\hat{H}_0(\cdot)$, $\hat{H}_1(\cdot)$, $\hat{H}_2(\cdot)$



firm can extract all their rents by charging them a fixed fee. Regarding the firm's profit, because I obtain only bounds for the cost parameters κ_1 and κ_2, namely 78 RMB for the former and 88 RMB for the latter, the firm's profit margin ranges from 39.46% (if κ_1=78 and κ_2=88) to 74.06% (if κ_1=0 and κ_2=0). The profit margins for different groups are of the following: group 0 at 54.36%, group 1 between 33.24% (if κ_1=78) and 76.43% (if κ_1=0) and group 2 between 40.96% (if κ_2=88) and 75.78% (if κ_2=0). Overall, China Telecom seems to be making a comfortable profit margin.

5.2 The Welfare Effects of Bundling

With structural estimates at hand, I can perform a counterfactual to evaluate the effects of bundling on firm's profit, consumer surplus and social welfare relative to component pricing. In particular, I simulate the case where the firm offers instead two fixed-fee contracts for internet and one usage-based contract for phone service. I assume that the firm does not change the internet speeds it offers.

A Theoretical Discussion

In general, the effects of bundling on consumer surplus and social welfare are ambiguous. However, the literature offers a consensus that lower prices or higher output levels are necessary for welfare improvement. For instance, in a discrete choice framework, Salinger (1995) shows that bundling can increase consumer surplus when it results in lower prices. Schmalensee (1981) and Schwartz (1990) formally show that welfare must fall if output does not rise with third degree price discrimination. While the previous literature has mainly focused on the use of bundling as a price discrimination device (See, e.g., Crawford and Yurukoglu (2012)), my model allows utility complementarity, cost saving effects and dependence between the two dimensions of asymmetric information. They may have different roles in determining the welfare effects of mixed bundling relative to unbundling. I will construct two examples below showing the ambiguity of the results in my model.

Under component pricing, the firm's problem is to maximize its profit by designing a tariff function while the consumers' taste are distributed according to a mixture of three conditional distributions. Since these three distributions differ in location and shape, their mixture is obtained by *shifting* and *reshaping* them. To isolate the role of these two procedures, I consider two examples. Let $U(q, j, \theta) = \theta q - \frac{1}{2}q^2$ if $q \leq \theta$ and $\theta^2/2$ otherwise, while the cost function is $c(q, j) = \gamma q$.

First, I consider two groups whose tastes for phone service are uniformly distributed on $[0, 1]$ and $[1, 2]$, respectively. Group 1 accounts for a proportion of 50%. If the firm can discriminate among the two groups, the optimal phone service assignments would be $q_1^*(\theta) = 2\theta - \gamma - 1$ and $q_2^*(\theta) = 2\theta - \gamma - 2$, while the cutoff tastes would be $\theta_1^c = (1 + \gamma)/2$ and $\theta_2^c = (2 + \gamma)/2$. If the firm cannot discriminate, it proposes a single assignment $q^*(\theta) = 2\theta - \gamma - 2$ with a cutoff taste $\theta^c = (2 + \gamma)/2$. I remark that in this case, the firm does as it was facing only consumers with a higher need of internet. Thus bundling benefits the consumers because consumers with a lower taste will not be excluded. This would results an increase in the consumer surplus. Similarly, since the firm would get the same profit from the consumers with a higher taste of internet, it will get a larger profit as bundling will allow it to get profit from the other group of consumers as well.

Second, I consider two groups whose tastes for phone service are distributed on the same interval $[0, 1]$ with densities $f_1(\theta) = 2(1 - \theta)$ and $f_2(\theta) = 1$. Assume $\gamma = 0$ and group 1 accounts for a proportion of 50%. Under mixed bundling and component pricing, firm's profit is 0.0602 and 0.0584, respectively. Consumer surplus is 0.0332 and 0.0341, respectively. In this case, the firm benefits from bundling while the consumer are penalized. Thus bundling can reduce consumer surplus because it provides an additional instrument for the firm to discriminate across consumers.

Counterfactual Simulation

Solving the firm's problem under unbundling is a hard task because the model does not lead to a closed-form solution, I will propose instead a numerical approximation of the solution. In particular, I will search numerically an optimal usage-based tariff function that is approximated by quadratic splines $T(\cdot;\delta) = \sum_{k=1}^{K} \delta_k \psi_k(\cdot)$, where $\psi_k(\cdot)$ is a quadratic basis function. The tariff $T(\cdot;\delta)$ is non-negative and increasing if and only if the coefficients δ_k are non-negative. Moreover, as I do not identify the type densities $f_0(\cdot)$, $f_1(\cdot)$ and $f_2(\cdot)$ below the cutoff tastes, I assume

$$\hat{f}_j(\theta) = \begin{cases} \hat{f}_j^*(\hat{\theta}_j^c)[1 - F_j(\theta_j^c)](\frac{\theta}{\hat{\theta}_j^c})^k & \text{if } \theta < \hat{\theta}_j^c, \\ \hat{f}_j^*(\hat{\theta}_j^c)[1 - F_j(\theta_j^c)] & \text{if } \theta \geq \hat{\theta}_j^c, \end{cases}$$

where $k = [\hat{\theta}_j^c \hat{f}_j^*(\hat{\theta}_j^c)(1 - F_j(\theta_j^c))/F_j(\theta_j^c)] - 1$. This approximation satisfies all the assumptions of Section 3. It allows continuity at the cutoff point $(\hat{\theta}_j^c, \hat{f}_j^*(\hat{\theta}_j^c)[1 - F_j(\theta_j^c)])$ and its integration from 0 to $\hat{\theta}_j^c$ equals $F_j(\theta_j^c)$.

The estimated optimal tariff function solves the following problem

$$\begin{aligned} \max_{\delta \geq 0} \quad & N_0 \int_{\underline{\theta}}^{\overline{\theta}} \Big(T(q_0(\theta;\delta);\delta) - (\hat{\kappa}_0 + \hat{\gamma} q_0(\theta;\delta))\Big) \hat{f}_0(\theta) d\theta \\ & + N_1 \int_{\underline{\theta}}^{\overline{\theta}} \Big(T(q_1(\theta;\delta);\delta) - (\hat{\kappa}_0 + \kappa_1 - \hat{\Delta}_1 + \hat{\gamma} q_1(\theta;\delta))\Big) \hat{f}_1(\theta) d\theta \\ & + N_2 \int_{\underline{\theta}}^{\overline{\theta}} \Big(T(q_2(\theta;\delta);\delta) - (\hat{\kappa}_0 + \kappa_2 - \hat{\Delta}_2 + \hat{\gamma} q_2(\theta;\delta))\Big) \hat{f}_2(\theta) d\theta, \end{aligned}$$

where $q_j(\theta;\delta) \equiv \arg\max_q \{\theta \hat{u}_0(q) + \hat{v}_j(q) - T(q;\delta)\}$ and N_j is the number of group-j consumers. Since κ_1 and κ_2 are not identified, I set them to zero. I use equally-spaced knots and increase number of parameters K until the marginal benefit of adding one more knot is less than 0.1%. The resulting estimated tariff function captures the optimal nonlinear price schedule under unbundling.

My simulation results show that unbundling would lead to a 10.14% decrease in firm's profit and a 17.18% decrease in consumer surplus, resulting in a 12.16% decrease in social welfare. These can be explained by the fact that unbundling would exclude too many consumers as discussed previously in my first numerical example. Figure 6 compares the tariff functions under mixed bundling (dashed lines) and component pricing (solid lines). Relative to mixed bundling, groups 0 and 1 would face more expensive tariff functions under component pricing, while group 2 would face a less expensive one. As a result, group 0 would lose by 58.11% of consumer surplus, while group 1 would lose 39.57%. On the contrary, group

2 would see an increase in their surplus by 9.48%. Thus unbundling would only benefit those who value highly internet.

Figure 6: Tariff Functions under Mixed Bundling and Component Pricing

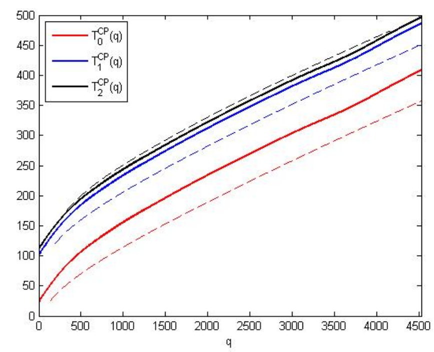


Figure 7 displays the breakdown of expected social welfare into consumer surplus and firm profit, while Figure 8 displays the breakdown of expected bill into cost and firm profit. I treat mixed bundling as the benchmark and normalize its corresponding welfare and bill to 100. Since the cost function is linear, changes in expected cost reflect changes in expected phone usage. Figure 7 confirms that groups 0 and 1 are losing the most in terms of consumer surplus under component pricing. The firm is losing profit as well. The loss is decreasing with internet speed. Figure 8 provides a justification, namely the production cost much decreases under component pricing because of a dramatic decrease in consumption of phone service. For instance, group 0 users' expected phone usage would drop from 497.62 to 253.09 minutes, while their expected indirect utility would decrease from 34.54 to 14.47 RMB. On the contrary, group 2 users would use 7.09% more of phone calls and thus would see their consumer surplus increasing by 9.48%. Finally, the ratio of the total informational rent by the total of bill would be 27.71% in component pricing, which represents a decrease in the cost of asymmetric information relative to bundling. This arises from a larger proportion of consumers who would be excluded under component pricing.

6 Conclusion

This paper studies bundling and price discrimination by a multiproduct firm selling internet and phone services in an imperfect information setting. Consumers are characterized by a taste for phone service and a minimum need for internet, thereby leading to a multidimensional screening problem. I derive the optimal selling mechanism, as well as the conditions on the model primitives under which different bundling strategies arise. I show that the model

Figure 7: Breakdown of Welfare

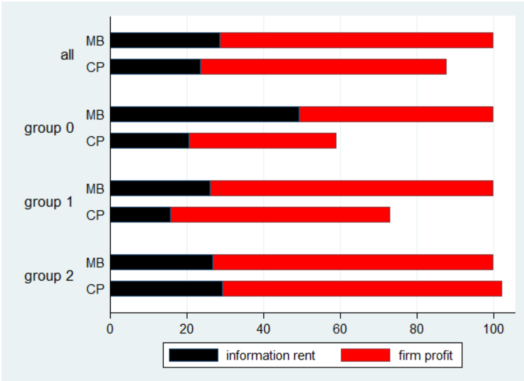
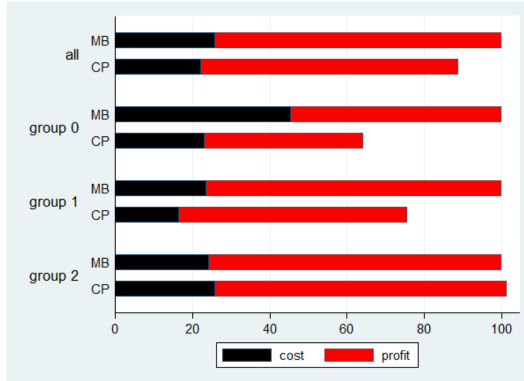


Figure 8: Breakdown of Bill



primitives are identified under parameterization of the cost function and multiplicative separability of the utility function. I develop a semiparametric estimator involving kernel density estimation and sieve estimators. The empirical analysis of China Telecom data suggests that both the firm and consumers benefit from bundling internet and phone services.

With the methodology I develop in this paper, a class of nonlinear pricing models with both discrete and continuous products/attributes can be solved theoretically and estimated empirically. Other goods featuring both minimum need and nonlinear pricing include water, energy, food and insurance. See e.g. Attanasio and Pastorino (2011) for nonlinear pricing of food in Mexican villages. While firms usually offer nonlinear pricing of insurance coverage, some states require a vehicle owner to carry some minimum level of insurance. Potential applications also include insurance contracts in which insurees bundle automobile and home insurance, and also a large number of products from manufacturing industries such as automobiles or computers where each product can be viewed as a bundle of various customized attributes. See e.g. Luo, Perrigne, and Vuong (2013). The results I developed in this paper can also be used to analyze products under nonlinear pricing with important network effects. See e.g. Chen and Luo (2012).

References

Adams, W. and J. Yellen (1976): "Commodity Bundling and the Burden of Monopoly," *The Quarterly Journal of Economics*, 90, 475–498.

Armstrong, M. (1996): "Multiproduct Nonlinear Pricing," *Econometrica*, 64, 51–75.

——— (1999): "Price Discrimination by a Many-Product Firm," *Review of Economic Studies*, 66, 151–168.

——— (2006): "Recent Developments in the Economics of Price Discrimination," In R. Blundell, W. Newey and T. Persson, eds., *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress*, Volume 2, Cambridge University Press.

Armstrong, M. and J.-C. Rochet (1999): "Multi-Dimensional Screening: A User's Guide," *European Economic Review*, 43, 959–979.

Armstrong, M. and J. Vickers (2010): "Competitive Non-Linear Pricing and Bundling," *The Review of Economic Studies*, 77, 30–60.

Aryal, G., I. Perrigne, and Q. Vuong (2009): "Identification of Insurance Models with Multidimensional Screening," Working Paper, Pennsylvania State University.

Athey, S. and P. Haile (2007): "Nonparametric approaches to auctions," In J. Heckman and E. Leamer, eds., *Handbook of Econometrics*, Volume VI, Amsterdam: North Holland.

Attanasio, O. and E. Pastorino (2011): "Nonlinear Pricing of Food in Village Economies," Working Paper, University of Minnesota.

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

Che, Y. and I. Gale (2000): "The Optimal Mechanism for Selling to a Budget-Constrained Buyer," *Journal of Economic Theory*, 92, 198–233.

Chen, L. and Y. Luo (2012): "Nonlinear Pricing with Network Effects in Yellow Pages," Working Paper, Pennsylvania State University.

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

Chu, C., P. Leslie, and A. Sorensen (2011): "Bundle-Size Pricing As an Approximation to Mixed Bundling," *American Economic Review*, 101, 263–303.

Crawford, G. and M. Shum (2007): "Monopoly Quality Degradation and Regulation in Cable Television," *Journal of Law and Economics*, 50, 181–219.

Crawford, G. S. and A. Yurukoglu (2012): "The Welfare Effects of Bundling in Multichannel Television Markets," *American Economic Review*, 102, 643–85.

Dole, D. (1999): "Cosmo: A Constrained Scatterplot Smoother for Estimating Convex, Monotonic Transformations," *Journal of Business & Economic Statistics*, 17, 444–455.

Dubin, J. and D. McFadden (1984): "An Econometric Analysis of Residential Electric Appliance Holdings and Consumption," *Econometrica*, 52, 345–362.

Economides, N., K. Seim, and V. Viard (2008): "Quantifying the Benefits of Entry Into Local Phone Service," *RAND Journal of Economics*, 39, 699–730.

Ekeland, I., J. J. Heckman, and L. Nesheim (2004): "Identification and Estimation of Hedonic Models," *Journal of Political Economy*, 112, S60–S109.

Fang, H. and P. Norman (2006): "To Bundle or Not to Bundle," *Rand Journal of Economics*, 37, 946–963.

Green, J. R. and J.-J. Laffont (1986): "Partially verifiable information and mechanism design," *The Review of Economic Studies*, 53, 447–456.

Grubb, M. and M. Osborne (2013): "Cellular Service Demand: Biased Beliefs, Learning, and Bill Shock," *American Economic Review*, forthcoming.

Guerre, E., I. Perrigne, and Q. Vuong (2000): "Optimal Nonparametric Estimation of First-price Auctions," *Econometrica*, 68, 525–574.

Hanemann, W. (1984): "Discrete/Continuous Models of Consumer Demand," *Econometrica*, 52, 541–561.

Ho, K., J. Ho, and J. H. Mortimer (2012): "The Use of Full-Line Forcing Contracts in the Video Rental Industry," *American Economic Review*, 102, 686–719.

Horrigan, J. (2010): "Broadband Adoption and Use in America," OBI Working Paper, Federal Communications Commision.

Ivaldi, M. and D. Martimort (1994): "Competition Under Nonlinear Pricing," *Annales d'Economie et de Statistique*, 34, 71–114.

Leslie, P. (2004): "Price Discrimination in Broadway Theater," *RAND Journal of Economics*, 35, 520–541.

Liu, H., P. Chintagunta, and T. Zhu (2010): "Complementarities and the Demand for Home Broadband Internet Services," *Marketing Science*, 29, 701–720.

Luo, Y. (2011): "Nonlinear Pricing with Product Customization in Mobile Service Industry," Working Paper, NET Institute.

Luo, Y., I. Perrigne, and Q. Vuong (2012): "Multiproduct Nonlinear Pricing: Mobile Voice Service and SMS," Working Paper, Pennsylvania State University.

——— (2013): "A General Framework for Nonlinear Pricing Data," Working Paper, New York University.

——— (2014): "Structural Analysis of Nonlinear Pricing," Working Paper, New York University.

Maskin, E. and J. Riley (1984): "Monopoly with Incomplete Information," *RAND Journal of Economics*, 15, 171–196.

Mazzeo, M. (2002): "Product Choice and Oligopoly Market Structure," *RAND Journal of Economics*, 33, 221–242.

McAfee, R., J. McMillan, and M. Whinston (1989): "Multiproduct Monopoly, Commodity Bundling, and Correlation of Values," *The Quarterly Journal of Economics*, 371–383.

McManus, B. (2007): "Nonlinear Pricing in an Oligopoly Market: The Case of Specialty Coffee," *RAND Journal of Economics*, 38, 512–532.

Miravete, E. (2002): "Estimating Demand for Local Telephone Service with Asymmetric Information and Optional Calling Plans," *Review of Economic Studies*, 69, 943–971.

——— (2005): "The Welfare Performance of Sequential Pricing Mechanisms," *International Economic Review*, 46, 1321–1360.

Miravete, E. and L. Röller (2004): "Competitive Nonlinear Pricing in Duopoly Equilibrium: The Early US Cellular Telephone Industry," Working Paper, University of Texas at Austin.

Narayanan, S., P. Chintagunta, and E. Miravete (2007): "The Role of Self Selection, Usage Uncertainty and Learning in the Demand for Local Telephone Service," *Quantitative Marketing and Economics*, 5, 1–34.

Rochet, J.-C. and P. Chone (1998): "Ironing, Sweeping, and Multidimensional Screening," *Econometrica*, 66, 783–826.

Rochet, J.-C. and L. Stole (2003): "The Economics of Multidimensional Screening," In D. Mathias, L.P. Hansen and S.J. Turnovsk, eds., *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress*, Volume 1, Cambridge University Press.

Salinger, M. (1995): "A Graphical Analysis of Bundling," *Journal of Business*, 68, 85–98.

Schmalensee, R. (1981): "Output and Welfare Implications of Monopolistic Third-Degree Price Discrimination," *American Economic Review*, 71, 242–247.

——— (1984): "Gaussian demand and commodity bundling," *Journal of Business*, 211–230.

Schwartz, M. (1990): "Third-Degree Price Discrimination and Output: Generalizing a Welfare Result," *American Economic Review*, 80, 1259–1262.

Seim, K. (2006): "An Empirical Model of Firm Entry with Endogenous Product-Type Choices," *RAND Journal of Economics*, 37, 619–640.

Stole, L. (2007): "Price Discrimination and Competition," In M. Armstrong and R. Porter, eds., *Handbook of Industrial Organization*, Volume 3, Amsterdam: North Holland.

Sundararajan, A. (2003): "Network Effects, Nonlinear Pricing and Entry Deterrence," New York University.

——— (2004): "Nonlinear Pricing of Information Goods," *Management Science*, 50, 1660–1673.

Appendix A: Proofs

Proof of Lemma 1: Due to minimum internet need, $j(\theta,\beta) \geq D(\beta)$ at equilibrium. I now prove by contradiction that the optimal internet assignment $j^*(\theta,\beta) = D(\beta)$. Consider the optimal mechanism $\{t(\cdot,\cdot), q(\cdot,\cdot), j(\cdot,\cdot)\}$ satisfying the IC, IR and MN constraints. Suppose that there exists some (θ,β) such that $j(\theta,\beta) > D(\beta)$. Now consider a new mechanism $\left\{\tilde{t}(\cdot,\cdot), \tilde{q}(\cdot,\cdot), \tilde{j}(\cdot,\cdot)\right\}$ where $\tilde{t}(\theta,\beta) = t(\theta,\beta) + U(\tilde{q}(\theta,\beta), \tilde{j}(\theta,\beta);\theta) - U(q(\theta,\beta), j(\theta,\beta);\theta)$, $\tilde{q}(\theta,\beta) = q(\theta,\beta)$, and $\tilde{j}(\theta,\beta) = D(\beta)$.

First, I show that the new mechanism satisfies the IC, IR and MN constraints. By definition, the consumer surplus keeps the same under both mechanisms. Thus, the IR and MN constraints hold under the new mechanism. I now show that the IC constraints hold under the new mechanism. Consider the original IC constraints:

$$U(q(\theta,\beta), j(\theta,\beta);\theta) - t(\theta,\beta) \geq U(q(\tilde{\theta},\tilde{\beta}), j(\tilde{\theta},\tilde{\beta});\theta) - t(\tilde{\theta},\tilde{\beta}).$$

By definition, the left-hand side equals $U(\tilde{q}(\theta,\beta), \tilde{j}(\theta,\beta);\theta) - \tilde{t}(\theta,\beta)$. The right-hand side equals $U(\tilde{q}(\tilde{\theta},\tilde{\beta}), \tilde{j}(\tilde{\theta},\tilde{\beta});\theta) - \tilde{t}(\tilde{\theta},\tilde{\beta})$ because

$$\begin{aligned}&\left[U(q(\tilde{\theta},\tilde{\beta}), j(\tilde{\theta},\tilde{\beta});\theta) - t(\tilde{\theta},\tilde{\beta})\right] - \left[U(\tilde{q}(\tilde{\theta},\tilde{\beta}), \tilde{j}(\tilde{\theta},\tilde{\beta});\theta) - \tilde{t}(\tilde{\theta},\tilde{\beta})\right]\\ =&U(q(\tilde{\theta},\tilde{\beta}), j(\tilde{\theta},\tilde{\beta});\theta) - U(\tilde{q}(\tilde{\theta},\tilde{\beta}), \tilde{j}(\tilde{\theta},\tilde{\beta});\theta) + U(\tilde{q}(\tilde{\theta},\tilde{\beta}), \tilde{j}(\tilde{\theta},\tilde{\beta});\tilde{\theta}) - U(q(\tilde{\theta},\tilde{\beta}), j(\tilde{\theta},\tilde{\beta});\tilde{\theta}) = 0,\end{aligned}$$

where the last equation is true if Assumption 3-(i) is satisfied. Thus, the IC constraints hold under the new mechanism. Second, the new mechanism is more profitable than the original one because of Assumption 3-(ii), a contradiction.

I now turn to $q^*(\cdot,\cdot)$ and $t^*(\cdot,\cdot)$. Given $j^*(\theta,\beta) = D(\beta)$, the consumer's problem becomes $\max_{q\in\mathbb{R}^+} U(q, D(\beta);\theta) - T^*(q, D(\beta))$. This implies that β only affects phone usage through $D(\beta)$. Hence, $q^*(\theta,\beta) = q^*(\theta, D(\beta))$. Given that I consider non-random nonlinear pricing schedules, it further implies that $t^*(\theta,\beta) = T^*(q^*(\theta, D(\beta)), D(\beta)) = t^*(\theta, D(\beta))$. □

Proof of Lemma 2: For a given cutoff taste θ^c, the optimal mechanism $\{q^*(\cdot,\cdot;\theta^c), j^*(\cdot,\cdot;\theta^c), t^*(\cdot,\cdot;\theta^c)\}$ can be derived following Sundararajan (2004) and is defined by (1), $j^*(\theta,\beta;\theta^c) = D(\beta)$, and (2) by replacing θ^c_j with θ^c. An important feature is that the allocation $\{q^*(\cdot,\cdot;\theta^c), j^*(\cdot,\cdot;\theta^c)\}$ does not depend on θ^c, while the optimal price schedule $t^*(\cdot,\cdot;\theta^c)$ does. In particular, $t^*(\theta,\beta;\theta^c) = U(q^*(\theta,\beta), D(\beta);\theta) - \int_{\theta^c}^{\theta} U_\theta(q^*(x,\beta), D(\beta);x)dx$. The provider's problem is then to find an optimal θ^c to maximize its expected profit

$$\int_{\theta^c}^{\overline{\theta}} \left[t^*(\theta,\beta;\theta^c) - c(q^*(\theta,\beta), D(\beta))\right] f(\theta|D(\beta))d\theta + \int_{\underline{\theta}}^{\theta^c} \left[v(0, D(\beta)) - c(0, D(\beta))\right] f(\theta|D(\beta))d\theta,$$

which is the summation of the profits collected from consumers buying internet and phone services and from consumers buying only internet. Differentiating the expected profit with respect to θ^c gives $-f(\theta^c|D(\beta))M(\theta^c, D(\beta))$, which leads to the boundary condition (3). $\square$

Proof of Lemma 4: If the two-dimensional IC constraints hold, the two one-dimensional constraints hold automatically. I now establish that, if the two one-dimensional IC constraints hold, the two-dimensional IC constraints hold as well. Consider any two pairs (θ, β) and $(\tilde{\theta}, \tilde{\beta})$, such that $D(\tilde{\beta}) \geq \beta$. The first one-dimensional IC constraint at $(\theta, \tilde{\beta})$ implies

$$U(q(\theta, \tilde{\beta}), D(\tilde{\beta}); \theta) - t(\theta, \tilde{\beta}) \geq U(q(\tilde{\theta}, \tilde{\beta}), D(\tilde{\beta}); \theta) - t(\tilde{\theta}, \tilde{\beta}).$$

The second one-dimensional IC constraint at (θ, β) implies

$$U(q(\theta, \beta), D(\beta); \theta) - t(\theta, \beta) \geq U(q(\theta, \tilde{\beta}), D(\tilde{\beta}); \theta) - t(\theta, \tilde{\beta}).$$

These two inequalities imply $U(q(\theta, \beta), D(\beta); \theta) - t(\theta, \beta) \geq U(q(\tilde{\theta}, \tilde{\beta}), D(\tilde{\beta}); \theta) - t(\tilde{\theta}, \tilde{\beta})$. Therefore, the two-dimensional IC constraints are satisfied. $\square$

Proof of Lemma 5: First, I show that $q^*(\theta, \beta)$ is decreasing in β. Since $q^*(\theta, \beta) = q^*(\theta, D(\beta))$, without loss of generality I show that $\partial q^*(\theta, j)/\partial j \leq 0$. Taking the total derivative of (1) with respect to j gives[15]

$$\frac{\partial q(\theta, j)}{\partial j} = \frac{-U_{qj} + c_{qj} + U_{q\theta j}\frac{1-F}{f} + U_{q\theta}\frac{\partial (1-F)/f}{\partial j}}{U_{qq} - c_{qq} - U_{qq\theta}\frac{1-F}{f}} \leq 0,$$

where the inequality holds since $c_{qj} \geq U_{qj}$ (Assumption 3-(ii)), $U_{q\theta j} = 0$ (Assumption 3-(i)), $U_{q\theta} > 0$ (Assumption 2-(ii)), $\partial \frac{1-F}{f}/\partial j \geq 0$ (Assumption 4). The denominator is negative under Assumption 2. See Sundararajan (2004) for details.

Second, I show that θ_j^c is increasing in j. Differentiating $M(\theta, j)$ with respect to j gives

$$\begin{aligned}\frac{\partial M(\theta, j)}{\partial j} &= [U_q - U_{q\theta}\frac{1-F}{f} - c_q]\frac{\partial q}{\partial j} + [U_j - v_j(0, j)] - [c_j - c_j(0, j)] - U_\theta \frac{\partial (1-F)/f}{\partial j} \\ &= \Big[U_j - v_j(0, j)\Big] - \Big[c_j - c_j(0, j)\Big] - U_\theta \frac{\partial (1-F)/f}{\partial j} \leq 0,\end{aligned}$$

where the first equation follows from $U_{\theta j} = 0$ (Assumption 3-(i)); the second equality uses Equation (1); the inequality holds since $U_\theta \geq 0$ (Assumption 2-(i)), $[U_j - v_j(0, j)] - [c_j -$

[15]For notation convenience, I use differentiation as if the variable j is continuous, suppress the arguments of functions and omit the asterisk superscript in this proof.

$c_j(0,j)] \leq 0$ (Assumption 3-(ii)), and $\partial \frac{1-F}{f}/\partial j \geq 0$ (Assumption 4). This implies that $M(\theta^c_{j'}, j) \geq M(\theta^c_{j'}, j') \geq 0$ if $j' > j$. Thus, $\theta^c_j \leq \theta^c_{j'}$ by the definition of θ^c_j.

Third, I show that $T^*(q,j) - v(q,j)$ is increasing in j. Lemma 2 implies that $T^*(q,j) - v(q,j) = u(q,\theta(q,j)) - \int_{\theta^c_j}^{\theta(q,j)} u_\theta(q(x,j),x)dx$, where $\theta(\cdot,j)$ is the inverse function of $q^*(\cdot,j)$. Differentiating with respect to j gives

$$\frac{\partial\Big[T^*(q,j) - v(q,j)\Big]}{\partial j} = u_\theta\Big(q\Big(\theta^c_j, j\Big), \theta^c_j\Big)\frac{\partial \theta^c_j}{\partial j} - \int_{\theta^c_j}^{\theta(q,j)} u_{q\theta}\Big(q(x,j),x\Big)\frac{\partial q(x,j)}{\partial j}dx \geq 0,$$

where the inequality holds since $u_\theta > 0$, $\partial\theta^c_j/\partial j \geq 0$, $u_{q\theta} > 0$ and $\partial q(\theta,j)/\partial j \leq 0$. $\square$

Proof of Proposition 4: I show that fixed cost parameters $\kappa_0 - \Delta_j$ can be identified for $j = 1,2$. The cutoff consumer receives no informational rent. Namely,

$$T_j(\underline{q}_j) = \theta^c_j u_0(\underline{q}_j) + v_j(\underline{q}_j). \tag{A.1}$$

By definition of the cutoff taste, I have

$$\Big[\theta^c_j u_0(\underline{q}_j) + v_j(\underline{q}_j) - v_j(0)\Big] - u_0(\underline{q}_j)\frac{1-F_j(\theta^c_j)}{f_j(\theta^c_j)} = \Big(\kappa_0 + \kappa_j - \Delta_j + \gamma\underline{q}_j\Big) - \kappa_j. \tag{A.2}$$

Note that $\frac{1-F_j(\theta^c_j)}{f_j(\theta^c_j)} = \frac{T'_j(\underline{q}_j)-\gamma}{u'_0(\underline{q}_j)} = \theta_0(\underline{q}_j)\frac{T'_j(\underline{q}_j)-\gamma}{T'_0(\underline{q}_j)}$. Equations (A.1) and (A.2) imply

$$\kappa_0 - \Delta_j = T_j(\underline{q}_j) - T_j(0) - \gamma\underline{q}_j - \theta_0(\underline{q}_j)u_0(\underline{q}_j)\frac{T'_j(\underline{q}_j) - \gamma}{T'_0(\underline{q}_j)}, \tag{A.3}$$

Similarly, if $j = 0$, $\kappa_0 = T_0(\underline{q}_0) - 0 - \gamma\underline{q}_0 - T_0(\underline{q}_0)\frac{T'_0(\underline{q}_0)-\gamma}{T'_0(\underline{q}_0)} = \frac{\gamma}{T'_0(\underline{q}_0)}T_0(\underline{q}_0) - \gamma\underline{q}_0$, where the first equality follows from $\theta_0(\underline{q}_0)u_0(\underline{q}_0) = T_0(\underline{q}_0)$. Thus κ_0 is identified, leading to the identification of Δ_1 and Δ_2 by (A.3). $\square$

Appendix B: Estimation

This appendix describes how I estimate my model semiparametrically.

Estimation of Tariff Functions and Construction of Phone Usage

The data provide the quantity of phone calls Q measured in minutes, the internet speed j measured in Mbps and the payment t measured in RMB. Following Luo (2011), I aggregate phone call minutes, add-ons and additional features into a single index $q = Q \times \epsilon$ to capture phone usage. The term ϵ captures the add-ons and additional features, which are unobserved by the analyst. Thus the tariff for group-j becomes $t = T_j(Q\epsilon)$, where $j \in \mathcal{J}$, and $T_j(\cdot)$ is strictly increasing and concave. Considering the inverse and taking the natural logarithm gives

$$\log Q = \log T_j^{-1}(t) - \log \epsilon. \tag{B.1}$$

Following Luo (2011), I assume that $\epsilon \perp \theta$. The tariff function $T_j(\cdot)$ is identified.

To estimate $T_j(\cdot)$, I approximate its inverse function with splines and find the optimal approximate spline that minimizes the sum of squared errors in (B.1). Since $T_j^{-1}(\cdot)$ is increasing and convex, I use constrained smoothing regression splines proposed by Dole (1999) to approximate it

$$\psi(\cdot;\delta_j) \equiv \sum_{l=1}^{n_j} \delta_j^l s_j^l(\cdot),$$

where δ_j is a vector of parameters δ_j^l, s_j^l is a cubic basis function, and n_j is the number of interior knots. The function $\psi(\cdot;\delta_j)$ is increasing and positive if and only if $\delta_j \geq 0$.

I then solve the following problem:

$$\min_{\delta_0,\delta_1,\delta_2 \geq 0} \quad \sum_{j\in\mathcal{J}} \sum_{i=1}^{N_j^*} \Big[\log Q_j^i - \log \psi(t_j^i;\delta_j) \Big]^2.$$

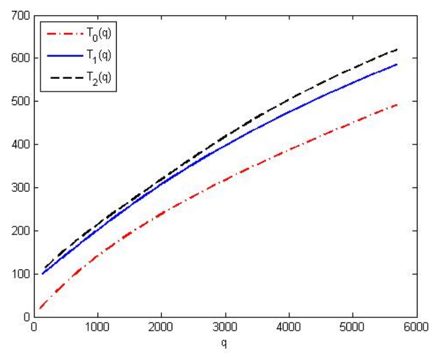
I estimate $T(\cdot)$ as $\hat{T}(\cdot) = \psi^{-1}(\cdot;\hat{\delta}_j)$. Figure 9 displays the estimated tariff functions $T_0(\cdot), T_1(\cdot), T_2(\cdot)$. I construct a bundle-$j$ user's phone usage as $q = \hat{T}_j^{-1}(t) = \psi(t;\hat{\delta}_j)$ for all $t \in [\underline{t}, \overline{t}]$. The data on bundle-$j$ users are $\Big\{(q_j^i, t_j^i)\Big\}_{i=1}^{N_j^*}$ and $\hat{T}_j(\cdot)$.

Estimation of γ, $u_0(\cdot)$ and $\theta_0(\cdot)$

In this subsection, I use phone-only users data, i.e. $\{(q_0^i, t_0^i)\}_{i=1,\ldots,N_0^*}$. To obtain an estimate of γ, I need to estimate $\overline{q}_0$. A convenient estimator that converges very fast is to take the maximum value, i.e. $\hat{\underline{q}}_0 = q_0^{N_0^*}$, leading to an estimator of γ, i.e. $\hat{\gamma} = T_0'(q_0^{N_0^*})$.

Following (8), I need to estimate $G_0^*(\cdot)$. I use the following empirical distribution estima-

Figure 9: Tariffs $\hat{T}_0(\cdot)$, $\hat{T}_1(\cdot)$, $\hat{T}_2(\cdot)$



tor leading to

$$\hat{G}_0^*(q) = \frac{1}{N_0^*} \sum_{i=1}^{N_0^*} \mathbb{1}(q_0^i \leq q), \tag{B.2}$$

for an arbitrary value of $q \in [\underline{q}_0, \overline{q}_0]$. An estimator of $\xi(\cdot)$ is obtained by replacing $G_0^*(\cdot)$ by its empirical distribution $\hat{G}_0^*(\cdot)$ and γ by $\hat{\gamma}$. Since $\hat{G}_0^*(\cdot)$ is a step function, the integral in (8) can be rewritten as a finite sum of integrals. Since in each of these integrals, $\log(1 - G_0^*(\cdot))$ is a constant and the primitive of $T_0''(\cdot)/T_0'(\cdot)$ is $-1/T_0'(\cdot)$, $\hat{\xi}(q)$ is equivalent to

$$\begin{aligned}\hat{\xi}(q) = & \Big[1 - \hat{G}_0^*(q)\Big]^{1 - \frac{\hat{\gamma}}{T_0'(q)}} \\ & \times \exp\left\{\hat{\gamma}\Big[\frac{1}{T_0'(q)} - \frac{1}{T_0'(q_0^{l+1})}\Big] \log(1 - \hat{G}_0^*(q_0^l)) + \hat{\gamma} \sum_{k=l+1}^{N_0^*-1} \Big[\frac{1}{T_0'(q_0^k)} - \frac{1}{T_0'(q_0^{k+1})}\Big] \log(1 - \hat{G}_0^*(q_0^k))\right\},\end{aligned}$$

for $q \in [q_0^l, q_0^{l+1})$, where $l = 0, 1, \ldots, N_0^* - 1$. For $q \in [q_0^{N_0^*}, \overline{q}_0]$, I have $\hat{\xi}(q) = 1$.

Using (6) and (7), I then estimate $u_0'(\cdot)$ and $\theta_0(\cdot)$ by

$$\begin{aligned}\hat{u}_0'(q) &= T_0'(q)\hat{\xi}(q), \\ \hat{\theta}_0(q) &= 1/\hat{\xi}(q),\end{aligned}$$

for an arbitrary value of $q \in [\underline{q}_0, \overline{q}_0]$. Lastly, I estimate θ_0^c by $\hat{\theta}_0^c = \hat{\theta}_0(q_0^1)$ and $u_0(\underline{q}_0)$ by $\widehat{u_0(\underline{q}_0)} = T_0(\underline{q}_0)/\hat{\theta}_0^c$. These estimates allow me to obtain an estimate of $u_0(\cdot)$ following (9).

Estimation of $v_j(\cdot)$, $\theta_j(\cdot)$, κ_0, κ_1, κ_2, Δ_1 and Δ_2

In this subsection, I am now using the bundle users data $\{(q_1^i, t_1^i)\}_{i=1,\ldots,N_1^*}$ and $\{(q_2^i, t_2^i)\}_{i=1,\ldots,N_2^*}$. Following (10), since estimates for γ and $u_0'(\cdot)$ have been obtained previously, I need an estimate of $g_j^*(\cdot)$ and $G_j^*(\cdot)$. I use the empirical distribution for $G_1^*(\cdot)$ and $G_2^*(\cdot)$ from (B.2) by replacing 0 by 1 and 2, respectively. For the density, I use a kernel density estimator to estimate $g_1^*(\cdot)$ and $g_2^*(\cdot)$. Following (10) and (11), the pseudo type $\theta_j(\cdot)$ and the marginal complementary utility $v_j'(\cdot)$ can be estimated as

$$\hat{\theta}_j(q) = 1 - \int_q^{\overline{q}_j} \frac{T_j'(x) - \hat{\gamma}}{\hat{u}_0'(x)} \frac{\hat{g}_j^*(x)}{1 - \hat{G}_j^*(x)} dx,$$
$$\hat{v}_j'(q) = T_j'(q) - \hat{\theta}_j(q)\hat{u}_0'(q),$$

where $q \in [\underline{q}_j, \overline{q}_j]$ and $j = 1, 2$. Lastly, I estimate θ_j^c by $\hat{\theta}_j^c = \hat{\theta}_j(q_j^1)$ and $u_0(\underline{q}_j)$ by $\widehat{u}_0(q_j^1)$. These estimates will allow me to obtain an estimate of $v_j(\cdot)$ following (12).

Following Proposition 4, the fixed cost parameters κ_0, Δ_1 and Δ_2 are estimated by

$$\hat{\kappa}_0 = \frac{\hat{\gamma}}{T_0'(\underline{q}_0)} T_0(\underline{q}_0) - \hat{\gamma}\underline{q}_0,$$
$$\hat{\Delta}_j = \Big[\frac{\hat{\gamma}}{T_0'(\underline{q}_0)} T_0(\underline{q}_0) - \hat{\gamma}\underline{q}_0\Big] - \Big[T_j(\underline{q}_j) - T_j(0) - \gamma\underline{q}_j\Big] + \hat{\theta}_0(\underline{q}_j)\hat{u}_0(\underline{q}_j)\frac{T_j'(\underline{q}_j) - \hat{\gamma}}{T_0'(\underline{q}_j)}, \quad \text{for } j = 1, 2,$$

where $\underline{q}_0$ and $\underline{q}_j$ can be replaced by their estimated counterparts, which are q_0^1 and q_j^1. The bounds for κ_1 and κ_2 are directly obtained from the data as the monthly fees for the internet-only users.

Estimation of $f_0(\cdot)$, $f_1(\cdot)$ and $f_2(\cdot)$

The previous two steps provide estimates of the pseudo types $\{\hat{\theta}_0^1, \hat{\theta}_0^2, \ldots, \hat{\theta}_0^{N_0^*}, \hat{\theta}_1^1, \hat{\theta}_1^2, \ldots, \hat{\theta}_1^{N_1^*}, \hat{\theta}_2^1, \hat{\theta}_2^2, \ldots, \hat{\theta}_2^{N_2^*}\}$, where $\hat{\theta}_j^i = \hat{\theta}_j(q_j^i)$ for $i = 1, 2, \ldots, N_j^*$ and $j = 0, 1, 2$. I could use standard kernel estimators to estimate $f_0^*(\cdot)$, $f_1^*(\cdot)$ and $f_2^*(\cdot)$ using these pseudo values. From the model of Section 3, the conditional density of types should satisfy the hazard rate property given by Assumption 2-(v). I then propose a regression spline estimator that allows me to impose the monotonicity restriction on $H(\cdot|j)$ for $j = 0, 1, 2$. In addition, I remark that $H(\theta|j)$ is bounded by θ since $[1 - F_j(\theta)]/f_j(\theta) \geq 0$. This represents a bound restriction that will also be imposed in the estimator. Specifically, I estimate $f_j^*(\cdot)$ under the restrictions that $H_j(\cdot) \equiv \cdot - \frac{1-F_j(\cdot)}{f_j(\cdot)}$ is increasing on $[\theta_j^c, 1]$, and $H_j(\theta) \leq \theta$ for all $\theta \in [\theta_j^c, 1]$.

Let the hazard function be $h_j(\theta) = f_j^*(\theta)/[1 - F_j^*(\theta)] = f_j(\theta)/[1 - F_j(\theta)] = 1/[\theta - H_j(\theta)]$ for $j = 0, 1, 2$. Using splines to approximate $H_j(\cdot)$, I can use the well-known expression

$$f_j^*(\theta) = h_j(\theta) \exp\Big[- \int_{\theta_j^c}^{\theta} h_j(x) dx\Big],$$

to obtain a maximum likelihood estimate for $f_j^*(\cdot)$. However, the direct implementation of the usual quadratic splines can ensure the monotonicity restriction but may violate the bound restriction. Instead, I define quadratic splines imposing the former, and then transform the coordinates to ensure the latter. In particular, I define the knots $\theta_j^c = \vartheta_j^0 < \vartheta_j^1 < \cdots < \vartheta_j^{k_j} < \vartheta_j^{k_j+1} = 1$. For any $\theta \in [\theta_j^c, 1]$, let $\psi(\theta; \delta_j) \equiv \sum_{l=1}^{k_j+3} \delta_j^l s_j^l(\theta)$, where the $s_j^l(\cdot)$s are quadratic basis functions satisfying $\psi(\cdot; \delta_j)$ positive and increasing if and only if the coefficients δ_j are positive. I then define

$$H_j(\theta; \delta_j) = \theta \frac{\psi(\theta; \delta_j)}{1 + \psi(\theta; \delta_j)}.$$

One can show that $H_j(\cdot; \delta_j)$ is positive, increasing and bounded by the 45 degree line if the coefficients δ_j^l are non-negative. Therefore, the hazard function can be expressed as $h_j(\theta; \delta_j) = [1 + \psi(\theta; \delta_j)]/\theta$, from which I can construct the log-likelihood of the pseudo sample as

$$l_j(\delta_j) = \sum_{i=1}^{N_j^*} \Big\{ \log \Big[\frac{1 + \psi(\hat{\theta}_j^i; \delta_j)}{\hat{\theta}_j^i} \Big] - \int_{\hat{\theta}_j^c}^{\hat{\theta}_j^i} \frac{1 + \psi(x; \delta_j)}{x} dx \Big\}.$$

Finally, I estimate $f_j^*(\cdot)$ by

$$\hat{f}_j^*(\theta) = \hat{h}_j(\theta) \exp[- \int_{\theta_j^c}^{\theta} \hat{h}_j(x) dx],$$

for $\theta \in [\theta_j^c, 1]$, where $\hat{h}_j(\theta) = [1 + \psi(\theta; \hat{\delta}_j)]/\theta$ and $\hat{\delta}_j$ maximizes $l_j(\delta_j)$.