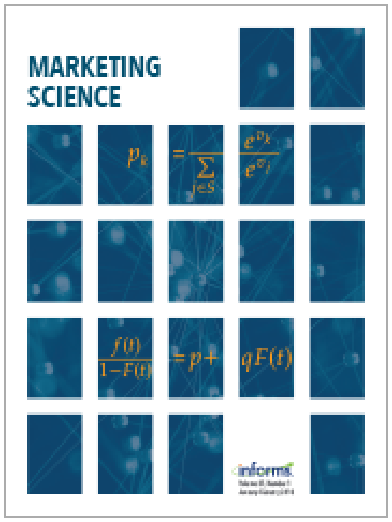


Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

How to Price Discriminate When Tariff Size Matters

Adib Bagh, Hemant K. Bhargava,

MARKETING SCIENCE
Vol. 32, No. 1, January–February 2013, pp. 111–126
ISSN 0732-2399 (print) | ISSN 1526-548X (online)

http://dx.doi.org/10.1287/mksc.1120.0720

How to Price Discriminate When Tariff Size Matters

Adib Bagh
Department of Economics, University of Kentucky, Lexington, Kentucky 40506,
adib.bagh@uky.edu

Hemant K. Bhargava
Graduate School of Management, University of California, Davis, Davis, California 95616,
hemantb@ucdavis.edu

Firms that serve a large market with many diverse consumer types use discriminatory or nonlinear pricing to extract higher revenue, inducing consumers to separate by self-selecting from a large number of tariff options. But the extent of price discrimination must often be tempered by the high costs of devising and managing discriminatory tariffs, including costs of supporting consumers in understanding and making selection from a complex menu of choices. These tariff design trade-offs occur in many industries where firms face many consumer types and each consumer picks the number of units to consume over time. Examples include wireless communication services, other telecom and information technology products, legal plans, fitness clubs, automobile clubs, parking, healthcare plans, and many services and utilities. This paper evaluates alternative ways to price discriminate while accounting for both revenues and tariff management costs. The revenue-maximizing menu of quantity-price bundles can be very (or infinitely) large and hence not practical. Instead, two-part tariffs (2PTs), which charge a fixed entry fee and a per-unit fee, can extract a large fraction of the optimal revenue with a small menu of choices, and they become more attractive once the costs of tariff management are factored in. We show that three-part tariffs (3PTs), which use an additional instrument, the "free allowance," are an even more efficient way to price discriminate. A relatively small menu of 3PTs can be more profitable than a menu of 2PTs of any size. This 3PT menu can be designed with less information about consumer preferences relative to the menu of two-part tariffs, which, in order to segment customers optimally, needs fine-grained information about preferences. Our analysis reveals a counterintuitive insight that more-complex tariffs need not always be more profitable; it matters whether the complexity is from many choices or more pricing instruments.

Key words: price discrimination; multipart tariffs; nonlinear pricing; tariff costs
History: Received: August 22, 2011; accepted: February 29, 2012; Preyas Desai served as the editor-in-chief and Duncan Simester served as associate editor for this article. Published online in *Articles in Advance* July 12, 2012.

> In late 2001, senior executives of the newly formed Virgin Mobile USA grappled with their pricing strategy. Should they offer complex price schedules that increased revenue through price discrimination but also led to high customer acquisition costs? Or should they offer a highly simple plan that sacrificed opportunities for price discrimination but enabled Virgin to acquire customers at lower cost?
>
> (see McGovern 2003)

1. Introduction

This paper evaluates and compares alternative pricing schemes in their ability to price discriminate efficiently, i.e., increase the firm's revenue while holding the administrative costs of implementing the pricing scheme low. Price discrimination is a powerful weapon in the managerial toolkit for raising firms' profits when they face heterogeneous consumers. The firm offers a common nonlinear price schedule (mapping quantity to price) to all consumers who then self-select and separate as they maximize their own surplus (Maskin and Riley 1984, Goldman et al. 1984, Sharkey and Sibley 1993). Firms employ a variety of pricing schemes in order to implement price discrimination (see, e.g., Iyengar and Gupta 2009), many of which are illustrated in Table 1. We refer to these pricing schemes as *tariffs* to represent any nonlinear pricing schedule, following Wilson (1993), who applies it to both regulated and unregulated pricing.

How much should a firm price discriminate? Why not simply pick the revenue-maximizing schedule?[1] There are several reasons to limit the extent of price discrimination. First, the price specification gets more complex or lengthy as the firm continues to increase revenue through more discriminatory pricing. This raises the cost of implementing the pricing scheme while the corresponding revenue gains get increasingly smaller; hence at some point, the higher tariff

[1] Here, and henceforth, *revenue* is defined after accounting for production and distribution costs; tariff administration costs are discussed separately.

Table 1 Examples of Nonlinear Pricing Structures

Tariff structure	Product	Example	Size
Flat rate (unlimited use)	Internet access	SBC: \$17.99/mo.	1
	Music download	Rhapsody: \$9.99/mo.	1
	Online retailing	Amazon Prime shipping: \$79/yr.	1
Two-part tariff	Golf club	Del Monte: \$285/yr. + \$42/day of play	2
Three-part tariff	Data center	RimuHosting: \$20/mo., 30 GB allowance, \$1/GB	3
Menu of two two-part tariffs	International calls (United States to India)	AT&T: \$2.95/mo. plus \$0.39/min. or \$3.99/mo. plus \$0.28/min.	4
Menu of two three-part tariffs	Wireless service	Verizon: \$40/mo., 450 min., \$0.45/min.; or \$80/mo., 1,350 min., \$0.35/min.	6
Menu of quantity-price bundles	Web hosting	GoDaddy: \$3.99/mo. for 10 GB, \$5.99/mo. for 150 GB, or \$7.99/mo. for unlimited space	6
	Freight delivery	FedEx (also DHL, UPS)	> 300
Block-declining tariff (taper)	Online data storage	Amazon S3 storage: 15 c/GB for 0–50 TB, 14 c/GB next 50 TB, 13 c/GB next 400 TB, 12 c/GB thereafter	7

Notes. Tariff size is measured as the number of parameters in the tariff specification. c/GB, cents per gigabyte.

costs offset the revenue gain (Murphy 1977, Dolan 1984, Wilson 1993, Miravete 2007). Second, designing a lengthier tariff generally requires more fine-grained information about customer preferences, raising the firm's information and decision costs (Chen and Iyer 2002, Acquisti and Varian 2005) and infringing on customers' privacy considerations (Odlyzko 2003). Third, consumers are often overwhelmed—and fail to choose anything—when given too many choices (Iyengar and Lepper 2000). This raises the firm's costs of explaining the choices and guiding customers' self-selection toward a purchase. The Harvard Business School case on Virgin Mobile's entry into the U.S. market notes that Virgin Mobile's competitors—who offered fairly complex price discrimination schedules—spent about \$100 per customer on sales staff hired to help customers with the myriad of menu options (McGovern 2003). Such high tariff management costs make it imperative for the firm to price discriminate *efficiently*, i.e., extract more revenue while avoiding a lengthy or complex price specification. In an empirical study of the U.S. cellular telephone industry, Miravete (2007, p. 1) concludes that firms "should only offer [a] few tariff options" because of the costs of tariff management. In a similar vein, Villas-Boas (2004) emphasizes that advertising and communications costs, which increase with the number of offers, exert downward pressure on the number of choices (i.e., menu size) offered to consumers.

Because nonlinear pricing involves indirect segmentation of the market (through self-selection by consumers), the firm must distort the options offered to certain segments in order to ensure that each segment picks the intended option. Different nonlinear pricing schemes vary in the revenue they generate relative to the ideal case of perfect price discrimination. But they also differ on the tariff management costs they impose. Our objective is to compare different pricing schemes on their *efficiency* at price discrimination; i.e., we care about both revenues and tariff management costs. The magnitude of tariff costs relative to revenue depends on the specific context, hence it is not practical to combine costs and revenue into a single objective function. However, tariff *size* (the number of parameters used in specifying the price schedule) serves as a useful proxy for evaluating how efficient a pricing scheme is at price discrimination, because, generally speaking, it is positively correlated with both higher revenue and higher tariff cost.[2]

We write tariff size as $S = n \times m$, where n is the number of parameters in the pricing scheme and m is the menu size or the number of choices in the menu. The rightmost column of Table 1 clarifies these terms. For instance, a single two-part tariff (F, p) has $2 \times 1 = 2$ parameters (F is a fixed entry fee and p is marginal price per unit). A single quantity-price bundle (q, T) also is of size 2 (T is total price for q units). A single three-part tariff has three parameters (F, p, and free allowance Q).[3] An m-item menu of two-part tariffs (or bundles) has tariff size $2m$, and an m-item menu of three-part tariffs has size $3m$. The firm can increase price discrimination by either providing more options (increasing menu size m) or adding more parameters to the pricing scheme (increasing n). Either approach involves an increase

[2] Dolan (1984) suggests tariff complexity as an alternative proxy for tariff efficiency. Tariff complexity is the number of break points in the tariff. For example, a menu of m two-part tariffs has $m-1$ breaks, as does a m-block tariff; a single three-part tariff on the other hand, has one break point. Our results remain valid under this alternative framing.

[3] Some authors refer to three-part tariffs as "fixed-up-to" tariffs, or two-part tariffs with inclusive consumption. Our usage is consistent with more recent marketing publications, including Lambrecht et al. (2007) and Iyengar and Gupta (2009).

in S. For example, starting with a single (q, T) bundle, the firm could either move to a menu of bundles ($S=4$) or offer a three-part tariff ($S=3$).

This paper demonstrates that three-part tariffs (3PTs) are highly efficient at price discriminating a large population of heterogeneous consumers. The argument is as follows. First, ignoring the effect of tariff management costs, the revenue-maximizing scheme is a menu of quantity-price bundles, but the menu is typically very large and often nonenumerable. Second, however, when menu size is restricted to be less than the revenue-maximizing menu size, then two-part tariffs can yield higher revenue than quantity-price bundles, keeping the overall tariff size the same across the two regimes. These two claims can be gleaned from the literature and are covered in §4. Third, we demonstrate that 3PT menus are even more efficient when tariff size is restricted. We show that under fairly plausible conditions, even a single 3PT (i.e., a specification of size 3) can outperform a substantially larger menu of two-part tariffs (2PTs). We also show that, in general, a size-constrained specification composed of 3PTs can produce higher revenue than an equal-size specification composed of 2PT menu or bundles. Another salient point is that this 3PT design can extract higher revenue with less or incomplete information.

We emphasize that our result is not merely that a 3PT menu outperforms an *equal-sized menu* of 2PTs (or an equal-sized menu of bundles): that outcome is trivial and the comparison unfair, because a 3PT subsumes both a 2PT and a bundle (set $Q=0$ to get the 2PT and $p=\infty$ to get the bundle), and the 3PT menu would have more parameters than the 2PT menu. Instead, we focus on a combination of higher revenue, lower tariff size, and less accurate information needed to design it. We caution that the 3PT menu employed in our constructive proofs is not the *optimal* 3PT menu; the latter would, of course, be even more profitable, but it would also require more information about demand preferences.

2. Examples and Intuition

We start with a series of examples to demonstrate the key results and the intuition behind them. Because prior literature already demonstrates that a menu of two-part tariffs is superior to a menu of quantity-price bundles when tariff size matters (see §4), our examples compare two-part tariffs with three-part tariffs having equal or fewer parameters. Example 1 below considers the most simple stylized setting with just two consumer types. We use this simple example only to convey the *intuition* behind the result. Example 2 provides a six-segment setting in which our result is more meaningful: we show that a single 3PT (3 parameters) beats the fully separating optimal design within the space of two-part tariffs (12 parameters); hence it is cheaper to implement and it generates higher revenue.

Example 1. The market has two segments, $i=1, 2$, and the proportion of Segment 1 is $\lambda_1 = 0.8$. The demand curves of consumers in $i = 1, 2$ are, respectively, $D_1(p) = (100 - p)$ and $D_2(p) = 250 - 1.25p$. The optimal menu of two-part tariffs (fully separating, with four parameters) is $(F_1, p_1) = (2{,}093.43, 35.3)$ and $(F_2, p_2) = (10{,}138.4, 0)$ with a revenue of \$5,529.41. The *single* three-part tariff (three parameters), $(F, Q, p) = (4{,}376.96, 64.7, 74.1)$, generates higher revenue: \$5,750.50.

The single three-part tariff produces 4% higher revenue than the optimal menu of two-part tariffs, despite having lower tariff size. The result holds as long as λ_1 is large enough. The intuition is that with a 2PT menu—which relies on self-selection—the firm's design of (F_2, p_2) is constrained by the fear that high-value customers would pick (F_1, p_1) meant for the lower segment. The large λ_1 amplifies this constraint and forces the firm to increase the consumption (and extractable surplus) of the low-value segment rather than distort their allocation in order to extract extra revenue from the high-value customers. In contrast, the single 3PT constructed above allots Segment 1 customers only the free allowance Q; hence the choice of the linear fee parameter p is constrained only by the preference of Segment 2. With this design (in which Segment 1 generates the same revenue as the 2PT menu), the firm extracts incrementally greater revenue from the high-value segment. This intuition is explained in more detail in §6.1.

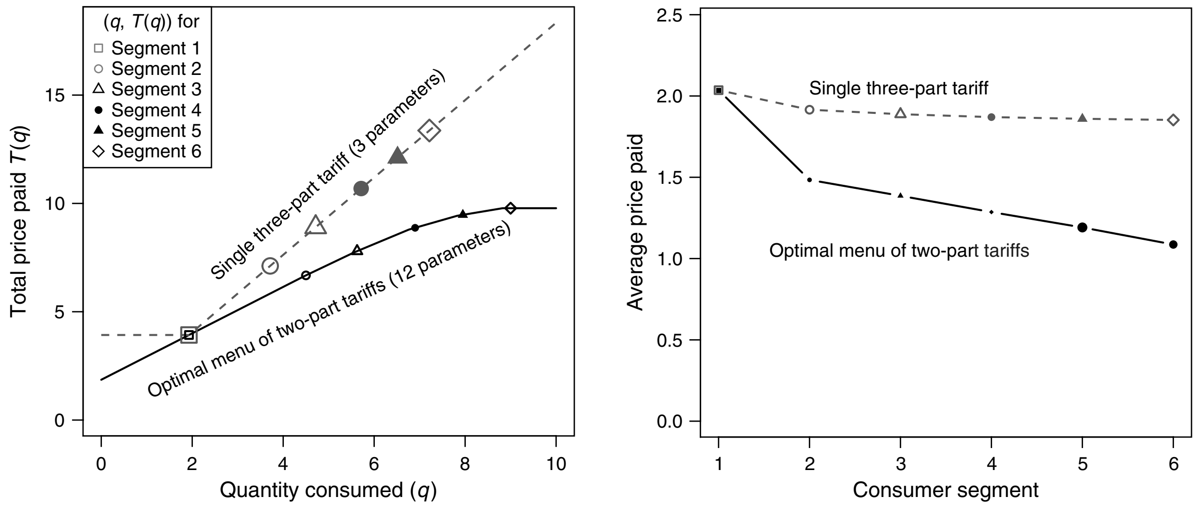
Example 2 (Single Three-Part Tariff Beats Menu of Six Two-Part Tariffs). The market has six segments, with proportion λ_i of the customers and (inverse) demands $v_i(q)$ (see Table 2). On one hand,

Table 2 **Tariff Design and Outcomes for Example 2**

Market description						
Segment i	1	2	3	4	5	6
Proportion λ_i	0.7	0.15	0.08	0.04	0.02	0.01
Inverse demand $v_i(q)$	$3-q$	$5.5-q$	$6.5-q$	$7.5-q$	$8.3-q$	$9-q$
Optimal 2PT menu						
Fixed fee F_i	1.8597	2.1786	2.8739	4.7336	6.6898	9.7786
Variable fee p_i	1.0714	1.0000	0.8750	0.6000	0.3500	0.0000
Quantity consumed q_i	1.9286	4.5000	5.6250	6.9000	7.9500	9.0000
Revenue $\pi_{2,i}$	3.9260	6.6786	7.7958	8.8736	9.4723	9.7786
Single 3PT						
Revenue $\hat{\pi}_{3,i}$	3.9260	7.1148	8.9005	10.6862	12.1148	13.3648

Note. The optimal two-part tariff has tariff size 12, whereas a single three-part tariff, with just three parameters, has a higher revenue and lower tariff cost.

Figure 1 **Revenue from Each Segment Under the Optimal Six-Item Menu of Two-Part Tariffs and a Single Three-Part Tariff**



Note. The left panel shows the total revenue from each segment, and the right panel shows the average price per unit consumed.

we have the optimal menu of two-part tariffs, specified in the table. It provides a unique option for each segment and yields a total revenue $\pi_2^* = \sum_i \lambda_i \pi_{2,i} = 5.0158$. On the other hand, we have a *single* three-part tariff $(\hat{F}, \hat{Q}, \hat{p}) = (3.9260, 1.9286, 1.7857)$, derived from the 2PT menu solution by fixing $\hat{Q} = q_1^*$ and $F = \pi_{2,1}$, and then setting $\hat{p}$ to maximize with respect to the residual demand of Segment 2. The single 3PT has $\hat{\pi}_3 = 5.33089$.

The single 3PT produces 6.2810% higher revenue than the optimal fully separating 2PT menu, and it generates at least as much revenue from *each segment* as does the 2PT menu (see Figure 1). And with only 3 tariff parameters, it has lower tariff management costs than the 2PT menu, which has 6 tariff choices and 12 parameters. Designing this single 3PT requires precise knowledge only of v_1, v_2, and λ_1. Finally, note that both pricing schemes implement quantity discounts (decreasing average price from 2.0357 to 1.8525, as depicted in Figure 1), even though the outlay in the single 3PT (as in any 3PT) is nonconcave.

The density of consumer types in these two examples is monotonically decreasing in their valuations, but this is not necessary for the result to hold. For example, suppose the Segment 1 customers in Example 2 are split into Segments 1a, 1b and 1c (comprising 10%, 40%, and 20% of the market, respectively), with 1a having the lowest valuation and 1c the same as original Segment 1. Then, depending on the valuations of Segments 1a and 1b, the optimal 2PT menu would have six, seven, or eight items. And under analogous conditions, it would still be possible to exceed this revenue with a one-, two-, or three-item menu of three-part tariffs. This idea is covered in §6, and Theorem 6.4 describes when a K-item three-part menu can outperform a fully separating I-item menu of 2PTs (with $3K < 2I$). Finally, Example 5 demonstrates that the 3PT design can extract higher revenue with less (or less accurate) information than needed to design the optimal 2PT menu.

3. Problem Statement

For each pricing scheme τ, we distinguish between (i) the *outlay* it specifies, a mapping from a subset **Q** of possible consumption levels $\mathcal{Q}$ into $\mathcal{R}^+$ (Sharkey and Sibley 1993); and (ii) the *specification* of this outlay, i.e., whether the mapping is announced *explicitly* or *implicitly*. A menu of quantity-price bundles of the form (q_i, T_i) is an explicit specification of the outlay: every quantity-price combination available to the user is stated explicitly.[4] For implicit tariffs, the general formulation is a piecewise linear (or "block") function that specifies marginal usage fees $p_1, p_2, \ldots, p_k$, and $k-1$ break points $q_1, \ldots, q_{k-1}$. Whereas an explicit price specification must explicitly state exactly how many offers the firm intends to make available to consumers, an implicit specification embeds an infinite series of offers in a compact declaration. This distinction is vital to understanding the relative performance of alternative pricing schemes under different circumstances.

The central question examined in this paper is, *which pricing scheme is most efficient at price discrimination when tariff size matters*? Suppose a firm faces a market of heterogeneous consumers indexed by a

[4] Technically, this menu contains implicit offers because one could consume an unspecified quantity between two explicitly specified levels and pay the higher of the two prices. However, this is a red herring, because such a consumption choice would never be optimal for the consumer.

type parameter $i \in \mathcal{I}$ (with distribution G). Let $v_i(q)$ be type i consumer's marginal valuation for the qth unit of consumption (with $q \in \mathcal{Q}$, and $v_i(q)$ decreasing in q). Type i consumer's demand curve $D_i(p)$ is the inverse function of her marginal valuations.[5] Let $\mathcal{D}$ represent the collection of these disaggregated demand functions along with the distribution G. Faced with $\mathcal{D}$, the firm needs to announce a price menu, a mapping $\tau: \mathcal{Q} \mapsto \mathcal{R}^+$ that specifies a tariff $T = \tau(q)$ for each consumption level q. It is aware that given a specification τ, each type i consumer will pick the option that maximizes her net surplus $V_i(q) - \tau(q)$ (if positive), where $V_i(q) = \int_0^q v_i(q)\,dq$ is the total valuation for q units. Let $\Pi(\tau; \mathcal{D})$ denote the firm's net revenues if it chooses a specification τ. Our interest is in the following question: *What specification optimally satisfies the firm's profit from price discrimination subject to a constraint $\bar{S}$ on tariff size*? That is, we solve the problem

$$\tau^* = \arg\max_{\tau} \Pi(\tau; \mathcal{D}) \quad \text{subject to } S(\tau) \le \bar{S},$$

subject to consumers' IC and IR constraints, where IC is the incentive compatibility and IR is the individual rationality.

Following prior literature on nonlinear pricing, we choose to compare three major players—quantity-price bundles, two-part tariffs, and three-part tariffs—in identifying the optimal tariff specification under tariff size constraints. The candidacy of the explicit price specification, a menu of quantity-price bundles, is obvious: the profit-maximizing price schedule is a mapping from $\mathcal{Q}$ to $\mathcal{R}^+$; hence the optimal profit $\Pi^*(\mathcal{D})$ can always be achieved by a menu of quantity-price bundles (though this menu may not always be computable). Among the piecewise linear price specifications, every concave outlay[6] can be represented with a menu of two-part tariffs, whereas any nonconcave outlay can be covered with a menu of three-part tariffs (§6.4 of Wilson 1993). Hence our analysis will directly cover these three alternative schemes.

4. Bundles vs. Two-Part Tariffs

A majority of the literature on nonlinear pricing focuses on quantity-price bundles and two-part tariffs (see, e.g., Shy 2008, Iyengar and Gupta 2009). Maskin and Riley (1984) show that when the distribution of types can be approximated with a continuous distribution F that is strictly increasing (positive density everywhere), continuous, and differentiable, the revenue-maximizing menu of bundles almost always implements a concave outlay. The key condition for this result, that the distribution of types has a nondecreasing hazard rate, is satisfied by nearly all continuous distributions and is essentially equivalent to a log-concavity requirement on the distribution function (Bagnoli and Bergstrom 2005). The interest in two-part tariffs stems from the fact that they are nearly optimal (when the market has a large number of types) and widely employed in practice. Faulhaber and Panzar (1977) show that the revenue under a menu of two-part tariffs increases monotonically in menu size and, in the limit, equals the revenue from the fully nonlinear tariff when this outlay is concave. In addition, a menu of two-part tariffs is *nearly* optimal even when the revenue-maximizing outlay is nonconcave (see §6.4 in Wilson 1993). The near-optimality of two-part tariffs also holds in the presence of competition (Armstrong and Vickers 2001). Hence, when tariff costs are inconsequential, then a menu that is optimal in the class of 2PT menus will typically also be optimal or near-optimal in the class of all nonlinear pricing (NLP) mechanisms.

Now consider the role of tariff costs. The profit-maximizing bundle is of infinite size and nonenumerable when the distribution of types is continuous; with a noncontinuous function, it would still need to explicitly specify as many options as the number of consumer types. With a restriction on tariff size, this bundle solution is no longer feasible, and size-constrained two-part tariffs gain an advantage over size-constrained bundles. The argument follows from Wilson's (1993) demonstration that a small menu of two-part tariffs can capture a bulk of the maximum revenue possible from price discrimination under self-selection constraints. Miravete (2007) estimates that a two-part tariff menu with around five items would capture over 90% of the revenue from the fully nonlinear tariff. Moreover, both Wilson (1993) and Miravete (2007) provide efficient procedures to compute the size-constrained optimal two-part tariff menu. In contrast, the size-constrained bundle solution requires a brute-force numerical procedure over a combinatorial search space, and it is likely to be inferior to a two-part tariff menu. This is because a k-item bundle menu makes only k offers to the large consumer population—defeating the key goal of nonlinear pricing—whereas a two-part tariff menu with the same tariff size makes an infinite series of offers, thereby extracting different levels of revenue from

[5] Some authors have considered valuation functions that reflect additional factors such as consumers' uncertainty about their demand (Clay et al. 1992, Lambrecht et al. 2007, Rochet and Stole 2002, Png and Wang 2010) and demand externalities caused by network effects (Oren et al. 1982, Masuda and Whang 2006).

[6] Concavity represents that marginal price is always declining in quantity. An outlay could offer quantity discounts everywhere (i.e., average price is always declining in quantity) or impose quantity premia somewhere. Concavity is not synonymous with quantity discounts. An outlay can be nonconcave yet have quantity discounts everywhere (illustrated in Example 2).

consumer types who pick the same item from the menu. The following example further illustrates that two-part tariffs are superior over bundles at price discriminating between heterogeneous consumers when the size of the menu is constrained.

Example 3. Suppose that there is a large number of agents (say, $I > 10$), with agent i's demand function being $D_i(p) = a_i \cdot (1 - p)$, and the a_i's have a discrete uniform distribution on the interval $(0, 1)$. Suppose that the tariff size is restricted to $S \leq 4$. Then the optimal bundle menu (with two items) is approximately $(q1, T1) = (0.22222, 0.148147)$ and $(q2, T2) = (0.666667, 0.296297)$ with revenue of 0.148148, whereas the optimal menu of two two-part tariffs is approximately $(F_1, p_1) = (0.016, 0.6)$ and $(F_2, p_2) = (0.16, 0.2)$ with a revenue of 0.16.[7]

In Example 3, the tariff size restriction $S \leq 4$ implies that the menu of either bundles or two-part tariffs contains only two items. Because of this, the menu of quantity-price bundles can fetch only two levels of revenue, $T1$ or $T2$. With a two-part tariff, the firm gains additional flexibility because the linear price component allows it to embed many more offers. Despite stating just two tariff options, the firm can extract different levels of revenue from even those consumers who pick the same option from the menu. The example is robust because the same logic applies to alternative demand specifications. Wilson (1993, Example 6.3) demonstrates the argument using demand functions of the form $D(p) = a_i - p$.

We emphasize that two-part tariffs are not always superior to bundles. This is trivially the case when tariff size is irrelevant. For certain products, two-part tariffs may not even be feasible. Contrast, for example, natural gas to bagged charcoal: the former can be sold via a piecewise linear tariff, but the latter requires bundles of specific quantities. In addition, with discrete consumer types, two-part tariffs have the weakness of making *too many offers* (even a single two-part tariff implies a continuous outlay function with an infinite series of offers). This makes it difficult for the firm to entice each consumer to pick the offer that extracts the best revenue from them, and the profit is strictly lower than the bundle profit (Kolay and Shaffer 2003). But the extent of disadvantage reduces as the number of customer types increases. Therefore, two-part tariffs play a prominent role in price discrimination for products that are targeted to large populations. They are particularly attractive when products are consumed over time because they allow the customer to choose consumption quantity (and total price) *after* having chosen the tariff; in contrast, a quantity-price bundle requires customers to make both commitments at the same time. Moreover, compared with linear pricing, the fixed-fee component of 2PTs is an attractive way to offset fixed costs in vertical channels (Raju and Zhang 2005) or the per-customer transaction costs that occur even when customers have little or no use (Sundararajan 2004).

Apart from these theoretical results, the superiority of two-part tariffs in the presence of tariff costs is further validated by firms' choice of pricing schemes in practice. Firms generally do not use quantity-price bundles when either $\mathcal{Q}$ or $\mathcal{I}$ is large. Miravete (2007, p. 1) notes that "fully nonlinear tariffs are rarely implemented as such but rather through a menu of self-selecting tariff options generally consisting of a fixed fee plus a constant charge per unit of usage." Examples include water, gas, or electric utilities; wireless phone service plans; software as a service; legal services; and healthcare services. When bundles must be used (e.g., bags of charcoal), then the firm typically provides only a limited set of options, and the few exceptions that do exist help explain why bundles are not attractive in such settings. For instance, parcel transportation firms such as Federal Express, UPS, and DHL display their price schedules as a menu of quantity-price bundles. Such menus contain over 150 items and run several pages long, even for a specific pair of zones and a specific speed of delivery. The following observation summarizes the findings from existing literature about two-part tariffs versus bundles.

Observation. *Quantity-price bundles are the optimal way to price discriminate when tariff size is inconsequential or the market has a handful of distinct consumer types. Otherwise, two-part tariffs are usually superior to bundles when tariff size is limited.*

5. Three-Part Tariffs vs. Two-Part Tariffs

Given the apparent superiority of two-part tariffs in price discriminating efficiently in the presence of tariff costs, we inquire whether there is an even more parsimonious way to price discriminate. Wilson (1993) indicates an evolution over time from linear to two-part to three-part tariffs (see the historical note at the end of his book). In this section we ask whether three-part tariffs can dominate two-part tariffs after controlling for tariff size.

[7] Computing the exact size-restricted optimal menu (of either two bundles or two 2PTs) poses a combinatorial challenge and is cumbersome. The tight approximation stated above is obtained by approximating the discrete distribution with a continuum of agents that are uniformly distributed on [0, 1] and then applying the technique of (Wilson 1993; see Example 6.2). As the number of agents I increases, these approximations become exact.

We adopt the classical framework of indirect segmentation (Stole 2007), where the monopoly firm cannot separate consumers ex ante but can offer a menu of choices to induce customers to self-select into separate segments. Without loss of generality, we assume that the firm's marginal cost of production, c, equals 0. We model consumer heterogeneity via a single-type parameter $i \in \{1, \dots, I\}$ (a finite, enumerable support, with arbitrary density) that captures differences in marginal valuation. This discrete framing of customer heterogeneity is asymptotically equivalent to a continuous distribution of segments; however, this discrete framing is most useful when the across-types differences are much sharper than within-types differences. Recall that for each $i \in I$, $v_i(q)$ and $V_i(q)$ represent Segment i customer's marginal and total valuations, respectively. Segment i has λ_i proportion of customers. All customers have diminishing marginal value for the service, and we rank the customer segments in increasing order of marginal valuations (so, for each q, a Segment $i+1$ customer has higher marginal valuation than a Segment i customer). We do not impose any additional restrictions on the hazard rate, demand elasticities, or other requirements for concave or unimodal payoff function.[8]

Assumption 1. *For every i, the demand D_i is finite at* 0, *and the marginal valuation function $v_i(q)$ is* (1) *differentiable* (*and, in particular, v_i' is finite*) *and decreasing in q, with* (2) $v_{i+1}(q) > v_i(q)$ *for all q.*

Customer i's demand function can be written as $q_i = D_i(p)$, where $v_i(q_i) = p$. Let $\bar{p}_i$ be Segment i's marginal valuation for the first unit of the product (so that $D(p) = 0$ for $p > \bar{p}_i$). The firm offers a menu of two-part (or three-part) tariffs from which a Segment i customer will choose the tariff and consumption quantity q_i that maximizes her net surplus (if positive, see Appendix A.1 for details). The optimal, fully separating, I-item menu of two-part tariffs solves the following revenue maximization problem (after substituting in all the binding constraints):

$$\max_{i:(F_i, p_i)} \Pi_{2PT} = \sum_{i=1}^{I} \lambda_i (p_i D_i(p_i) + F_i)$$
$$= \sum_{i=1}^{I} \lambda_i \left(p_i D_i(p_i) + \sum_{j=1}^{i} \int_{p_j}^{p_{j-1}} D_j(\tau)\, d\tau \right).$$

[8] Strictly speaking, we assume λ_is such that the optimal 2PT menu is fully separating: there is a distinct tariff for every segment. However, if it were not and the optimal 2PT menu implements some pooling of segments, then we just redefine I as the number of unique items in the menu (fewer than the original I) and redefine the affected λ_is as the sum of consumers that pick the same tariff item i from the optimal menu. At any rate, the ability of a small-sized 3PT menu to outperform the optimal 2PT menu increases when the 2PT menu has some pooling.

The optimal tariff design follows directly from the first-order conditions for the optimal $p_i's$ and then substituting the optimal $p_i's$ into the $F_i's$.

Lemma 1. *Under Assumption* 1, *the optimal*, *fully separating menu of two-part tariffs* $[(F_i^*, p_i^*)]$ *satisfies*

$$p_i^* = \left(\frac{1 - \sum_{j=1}^{i} \lambda_j}{\lambda_i} \right) \frac{(D_{i+1}(p_i^*) - D_i(p_i^*))}{-D_i'(p_i^*)}, \tag{1}$$

$$F_i^* = \int_{p_1^*}^{\bar{p}_1} D_1(\tau)\, d\tau + \int_{p_2^*}^{p_1^*} D_2(\tau)\, d\tau + \cdots + \int_{p_i^*}^{p_{i-1}^*} D_i(\tau)\, d\tau, \tag{2}$$

with $p_1^* > p_2^* > \cdots > p_I^* = 0$.

Note that $p_I^* = 0$ because we assumed that $c = 0$. Lemma 1 can be deployed to produce exact solutions for the optimal 2PT menu after a specific demand function is identified (the solution for linear demand is illustrated in §6.3).

6. Results

Before diving into the results, we provide a map of the results and their positioning with respect to the literature and design of price discrimination schemes in practice. Our first result, Theorem 1, formalizes the conclusion of Example 2 by deriving a sufficient condition for the existence of a single three-part tariff that generates higher revenue than a 2PT menu of any size. Section 6.3 illustrates the general results for the special case that each individual customer has linear demand. Theorem 2 in §6.4 provides the general result: conditions for a small-K three-part tariff to surpass the optimal 2PT menu. Finally, §6.5 explains why the design of the superior 3PT menu is less sensitive to information error and hence imposes lower information requirements and tariff design costs. All proofs are in Appendix A.2.

From §4, we recall that two-part tariffs are optimal or near-optimal when the distribution of consumer types I is (i) continuous and (ii) satisfies certain smoothness assumptions (essentially, log-concavity) that lead to the optimal schedule being concave. When the optimal schedule is not fully concave, then 3PTs have an inherent advantage. The striking feature of Theorem 1 is that a *single* 3PT (or in its general counterpart, Theorem 2, a small-K 3PT menu) outperforms the large and possibly infinite 2PT menu. The superiority of 3PT does not require imposition of quantity premia, despite nonconcavity in the outlay (see Example 2). Our results also apply when conditions (i) or (ii) fail. Log-concavity, which implies that preferences not be too diverse (Cowan 2007), is violated when adjacent types are distinct, leaving points of zero mass in the density function. One example is when the market has distinct segments such as home and business users, students and professors and labs and corporate users, or professional and amateur users. In such markets, cross-segment differences may be much sharper than within-segment variations (Wedel and Kamakura 2000), leading to multiple modes in the (continuous) demand distribution (Johnson and Myatt 2003). Then the apparatus of Maskin and Riley (1984) becomes inapplicable; however, the (more general) formulation of §5 still applies, and the optimal 2PT menu is still given by Lemma 1. The 3PT design outperforms this 2PT menu while keeping tariff size small. Finally, when the size of the menu is constrained, then our results are meaningful even if the unconstrained-optimal schedule is concave everywhere.

6.1. Single Three-Part Tariff Outperforms Optimal 2PT Menu

We develop the result by constructing a particular single three-part tariff $(\hat{F}, \hat{Q}, \hat{p})$ using the optimal 2PT menu. Set the free allowance $\hat{Q}$ to q_1^*, the optimal consumption level of Segment 1 under the optimal menu of two-part tariffs. Similarly, set $\hat{F}$ to the total amount paid by Segment 1 (sum of the fixed fee and usage fee). Therefore, by construction, the single 3PT and the 2PT menu both extract identical profit from Segment 1 customers. Now, set the third parameter $\hat{p}$ in the 3PT to optimize the profit from Segment 2 customers, $p(D_2(p) - \hat{Q})$.

Quasi-optimal three-part tariff:

$$\left\{ \begin{array}{l} \hat{Q} = q_1^* = D_1(p_1^*) \\ \hat{F} = F_1^* + p_1^* \hat{Q} \\ \hat{p} = \dfrac{(D_2(\hat{p}) - \hat{Q})}{-D_2'(\hat{p})} \end{array} \right\}. \tag{3}$$

Let $\hat{\Pi}_{3PT}$ denote the profit from this design. Note that this is not an optimal tariff even within the subspace of single three-part tariffs. In fact, because $\hat{p}$ is optimized using Segment 2 demand only (rather than the types higher than 2), it is equivalent to solving a problem where types $3, \ldots, I$ collapse into type 2. Therefore this quasi-optimal tariff can be viewed as a "least optimistic" 3PT design. Yet even this suboptimal single three-part tariff can outperform the optimal menu of two-part tariffs.

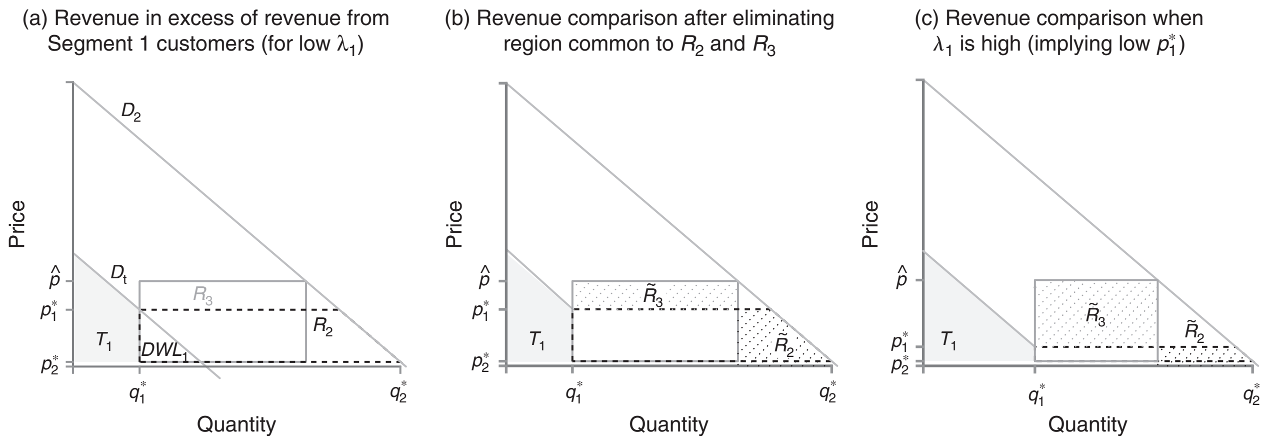
Theorem 1 (Single Three-Part Tariff Outperforms the Optimal 2PT Menu). *Suppose there are I customer segments, with demand functions $D_i(p)$ and weights λ_i, such that the optimal menu of two-part tariffs is fully separating and has I items of the form (F_i^*, p_i^*). Then there always exists a threshold $\hat{\lambda}_1 \in (0, 1)$ such that if $\lambda_1 \in (\hat{\lambda}_1, 1)$, then the quasi-optimal single three-part tariff defined in Equation* (3) *generates strictly higher revenue than the optimal* 2PT *menu.*

The result is striking because the two-part tariff design carries a tariff size $2I$ while the three-part tariff generates higher revenue with only three parameters (recall Example 1, where $2I = 12$). Clearly, the latter becomes even more advantageous once tariff management costs are factored in. The intuition behind the result is that when the market is tilted toward the low-value segment, the 2PT design must set a relatively low p_1^* in order to minimize the deadweight loss for this segment. Yet this low p_1^* reduces the firm's price discrimination ability by making it harder to induce higher-type customers to pick a different item from the menu (i.e., successively, an item with higher F and lower p). In contrast, the three-part tariff design decouples the unit variable fee for the low-value segment and the remaining higher-value segments. It serves as a quantity-price bundle for the low-value segment (i.e., they consume no more than their allowance Q), whereas the higher segments consume fewer units than under the 2PT menu (because $\hat{p} > p_2^* = 0$). Thus, the single three-part tariff functions as a hybrid that combines the best features of linear pricing and quantity-price bundle.

To see the logic and intuition behind the result more clearly, consider the special case of $I = 2$. In Figure 2, $T1$ represents the revenue realized from Segment 1 via the (F_1^*, p_1^*) from the optimal 2PT menu, and q_1^* is the corresponding consumption level. In the single 3PT, we set $\hat{F}_1$ to $T1$ and $\hat{Q} = q_1^*$. Now, we just need to compare the two pricing schemes on the revenue they generate from Segment 2. For the 2PT menu, each Segment 2 customer generates an additional revenue R_2 (i.e., each Segment 2 customer pays $T1 + R_2$). With the quasi-optimal 3PT, each Segment 2 customer generates a revenue R_3 in excess of $T1$. Figure 2(a) depicts these two regions. Note that the difference in profits $\hat{\Pi}_{3PT} - \Pi_{2PT}$ is the difference $R_3 - R_2$. Canceling the common region in R_2 and R_3, we can write $\hat{\Pi}_{3PT} - \Pi_{2PT} = \tilde{R}_3 - \tilde{R}_2$, as depicted in Figure 2(b). In panel 2(a), the triangular area marked DWL_1 is the deadweight loss for Segment 1 customers. When λ_1 is large, the optimal 2PT menu has a smaller p_1^* in order to reduce this deadweight loss (see Figure 2(c)). However, the reduction in p_1^* limits the firm's ability to price discriminate and extract excess revenue from Segment 2 customers, thereby reducing the region $\tilde{R}_2$. Because of this, the single three-part tariff starts dominating the optimal 2PT menu for large enough λ_1.

Finally, we note that the incremental gain from using a single three-part tariff is not monotonically increasing in λ_1. That is, if $\Pi_{2PT}(\lambda_1)$ and $\hat{\Pi}_{3PT}(\lambda_1)$ denote the profits from optimal 2PT and the quasi-optimal 3PT, respectively, at a given λ_1, then

Figure 2 Revenues from Each Segment 2 Customer, in Excess of Revenue from Segment 1 Customers, for the Optimal 2PT Menu (Region R_2) and the Quasi-Optimal Single 3PT (Region R_3)



$\Delta(\lambda_1) = \Pi_{3PT}(\lambda_1) - \Pi_{2PT}(\lambda_1)$ is not monotone in λ_1. This can be understood by examining the behavior of $\Delta(\lambda_1)$ starting at $\lambda_1 = 0$. For very small values of λ_1, the low type is ignored, and hence $\Delta(\lambda_1) = 0$ as both the two-part tariff and three-part tariff extract full surplus from Segment 2. As λ_1 increases, both types get served, and the 2PT menu discriminates better than the single 3PT; hence Δ initially becomes negative. But as λ_1 continues to increase, the logic of Theorem 1 kicks in: both types continue to be served, and Δ turns positive as the single quasi-optimal 3PT dominates. But for very high λ_1, the market collapses into a single customer type, and Δ becomes 0 (take the extreme case $\lambda_1 = 1$). Therefore, Theorem 1 is most useful when λ_1 is larger than $\hat{\lambda}$ but not too close to 1.

6.2. Determining When to Employ Three-Part Tariffs

Whereas Theorem 1 merely asserts an existence condition for the superiority of a single three-part tariff, the actual values of the threshold $\hat{\lambda}_1$ can easily be calculated for any given demand functions. For instance, if the demand functions of individual consumers are of the form $D_i(p) = a_i - b_i(p)$, then an upper bound for $\hat{\lambda}_1$ is 0.75 (see Corollary 1). That is, $\hat{\lambda}_1 = 0.75$ is sufficient (but not necessary) for a single three-part tariff to dominate. The parameter conditions for the result cover many practical applications. The upper bound of 0.75 simply implies that there are three times as many "low-value" consumers as there are higher-value consumers. This distribution occurs in several markets, for instance in information technology (IT), telecommunications, healthcare, and other services where a small segment of "heavy" users has vastly different consumption patterns than the remaining majority of "light" users. For IT services, the ratio of light to heavy users routinely exceeds 4:1 (20% of users account for over 80% of usage).

We can think of a few examples where the firm employs only a single three-part tariff even though market conditions suggest intensive price discrimination (because the firm faces large numbers of customers who are substantially heterogeneous in their demand preferences). For example, the University of California (UC) offers a single legal plan (to tens of thousands of employees) that bundles a free allowance of eight hours' legal consultation into the annual premium, with additional hours billed at a specified rate. The pricing strategy fits our findings because (based on discussions with the benefits managers at UC) most employees have very low consumption while a small fraction of heavy users accounts for the majority of consumption. Similarly, many fitness clubs include one or two sessions with a physical trainer in the initiation fee and charge (about \$60/session) for additional sessions, consumed by only 10%–20% of members. Another example is roadside assistance services where the annual membership fee includes a towing allowance up to some distance (e.g., in Kentucky, AAA covers 100 miles for the \$37 annual fee and charges \$3 per mile above that). Arguably, a majority of members (both in expectation and ex post) utilize only this free allowance, and a small number goes over.[9] Other examples that fit the pattern (though we do not have empirical data to confirm the claim) include website design companies that offer up to N pages with the service plan and charge a fee for additional pages and Web hosting companies that offer similar plans using some computational metric such as the number of gigabytes transferred, the number of FTP accounts, or the number of individual email accounts.

[9] We thank one of the reviewers for pointing out this example and for inspiring us to verify consumption patterns with the University of California and fitness clubs.

6.3. Illustration: Linear Demands

The computations described thus far can be made more precise given a specific functional form of the demand functions for individual consumers. We demonstrate using demand functions of the form $D_i = a_i - b_i p$, where $b_i \geq b_{i+1}$ and $a_{i+1}/b_{i+1} \geq a_i/b_i$ for all i. This specification implies that for each Segment i, the marginal valuations $(v_i(q) = (a_i - q)/b_i)$ are decreasing at a constant rate $1/b_i$, and the total valuation $V(q)$ is quadratic in q. This functional form is commonly employed to model the demand preferences when consumers demand multiple units of a product (Lambrecht and Skiera 2006). Moreover, it is not very restrictive because it applies to individual demand rather than market demand, and the aggregate market demand faced by the firm remains nonlinear.

Corollary 1. *For demand functions* $D_i(p) = a_i - b_i p$,

$$p_1^* = \frac{(1-\lambda_1)(a_2 - a_1)}{\lambda_1 b_1 + (1-\lambda_1)(b_2 - b_1)}, \quad \hat{p} = \frac{a_2 - a_1 + b_1 p_1^*}{2b_2},$$

and the quasi-optimal three-part tariff $(\hat{F}, \hat{Q}, \hat{p})$ *generates strictly higher revenue than a* 2PT *menu of any size whenever* $\lambda_1 \in (\hat{\lambda}_1, 1)$ (*where* $\hat{\lambda}_1 = (3b_2)/(b_1 + 3b_2)$) *or, equivalently,* $\lambda_1/(1-\lambda_1) > (3b_2)/b_1$ *and* $\lambda_1 < 1$.

For the special case where $b_1 = b_2$, this yields a threshold $\hat{\lambda}_1 = 0.75$ as stated above. Moreover, once the firm knows the parameter values of the demand function, it can use a more refined computational procedure to compute a tighter threshold $\hat{\lambda}_1$. For example, this procedure would yield $\hat{\lambda}_1 = \sqrt{0.5} \approx 0.707$ for the case where $b_i = 1$ for all i. Finally, if $\lambda_1 > \hat{\lambda}_1$ does not hold, then it is still possible that, for some $1 < K < I$, a menu of K 3PTs (size $3K$) generates more profits that the optimal menu with m_2 2PTs (size $2m_2$), even when $3K \leq 2m_2$. This is demonstrated in the next section.

6.4. *K*-Item Menu of Three-Part Tariffs

Theorem 1 states an extreme result, that a single three-part tariff can outperform an arbitrarily large optimal 2PT menu when there is a sufficiently high concentration of low-type (Segment 1) customers. The threshold $\hat{\lambda}_1$ depends on the demand functions D_i and the weights λ_i. However, if the demand distribution is such that λ_1 is not larger than the threshold $\hat{\lambda}_1$, then a single three-part tariff may fail to beat the optimal 2PT menu. Might there still be some value in the logic of price discriminating via a small menu of three-part tariffs?

We identify conditions under which a K-item menu of three-part tariffs (with small K such that $K < (2m_2)/3$, where m_2 is the number of options in the optimal 2PT menu) produces higher revenue than a 2PT menu of any size. As a first step, we extend the logic of Theorem 1 to $K = 2$, defining a two-item, three-part tariff as follows. The first item in this solution is a "package deal" (a quantity-price bundle) identical to the first item in the 2PT menu: $\hat{Q}_1 = D_1(p_1^*)$ and $\hat{F}_1 = F_1^* + p_1^* \hat{Q}_1$, with $\hat{p}_1$ set sufficiently high to deter additional consumption by Segment 1. The second item is also derived from the second item in the 2PT menu—$\hat{Q}_2 = D_2(p_2^*)$ and $\hat{F}_2 = F_2^* + p_2^* \hat{Q}_2$—and the $\hat{p}_2$ in this offer is chosen to optimize the 3PT residual profit from Segment $K+1 = 3$. This item serves as a package deal for Segment $K = 2$, whereas customers in Segments $3, \ldots, I$ consume additional units at $\hat{p}_2$. This two-item 3PT menu beats the optimal 2PT menu so long as $\lambda_2/(1-\lambda_1)$ is large enough. If this fails, repeat the computation for $K = 3$. The formal result is stated by relating the threshold level at the Kth type to the discretized hazard rate $HR(K)$.[10]

Theorem 2 (K-Item Three-Part Tariff Outperforms the Optimal 2PT Menu). *In the setting of Theorem* 1, *there exists for every* $K = 1, \ldots, I-1$ *a corresponding threshold* $\tau_K < 1$ *such that a* K*-item menu of three-part tariffs outperforms the optimal* 2PT *menu if* $HR(K)$ *is finite and* $HR(K) > \tau_K$. *The* 3PT *menu has* $\hat{Q}_i = D_i(p_i^*)$ *and* $\hat{F}_i = \pi_{2,i}^*$ *for all* $i \leq K$ *and* $\hat{p}_K = \arg\max_p (p \cdot (D_{k+1}(p) - \hat{Q}_K)$ (*while the remaining* $\hat{p}_i$*s can be set sufficiently high*).

An example of price discrimination via a short menu of three-part tariffs is the wireless telecommunications industry. Most cell phone service providers in the U.S. market their wireless voice service to tens of millions of consumers with a handful of three-part tariffs. To determine such a menu, a firm would look for the smallest K that satisfies Theorem 2. The result demonstrates that a short 3PT menu can dominate two-part tariffs on both revenue and tariff cost even when the hazard rate satisfies the usual "monotone increasing" property (whereupon the optimal schedule implies a concave tariff outlay and quantity discounts everywhere). The exact form of the threshold (for each K) can be specified algebraically and more tightly for any given functional form of D_i's. For instance, if we assume the linear demand specification used in §6.3, then a K-item 3PT menu generates more revenue than the optimal 2PT menu if $HR(K)$ is finite and

$$HR(K) > \tau_K = \frac{3b_K}{b_K + 3b_{K+1}}. \quad (4)$$

Note that the above condition involves the distribution of types and is distinct from the monotonicity of

[10] For demand massed at points $z_1 < \cdots < z_n$, the discrete hazard rate at $z_i \leq z_n$ is $HR(z_i) = \Pr(Z = z_i)/(1 - \sum_{j<i} \Pr(Z = z_j))$ (Barlow et al. 1963, Shaked et al. 1995).

the hazard rate (a distribution with an increasing hazard rate can be skewed to the left or to the right; for example, consider the Burr distribution with various parameters). The condition is independent of whether the optimal NLP mechanism offers quantity premia or quantity discounts.

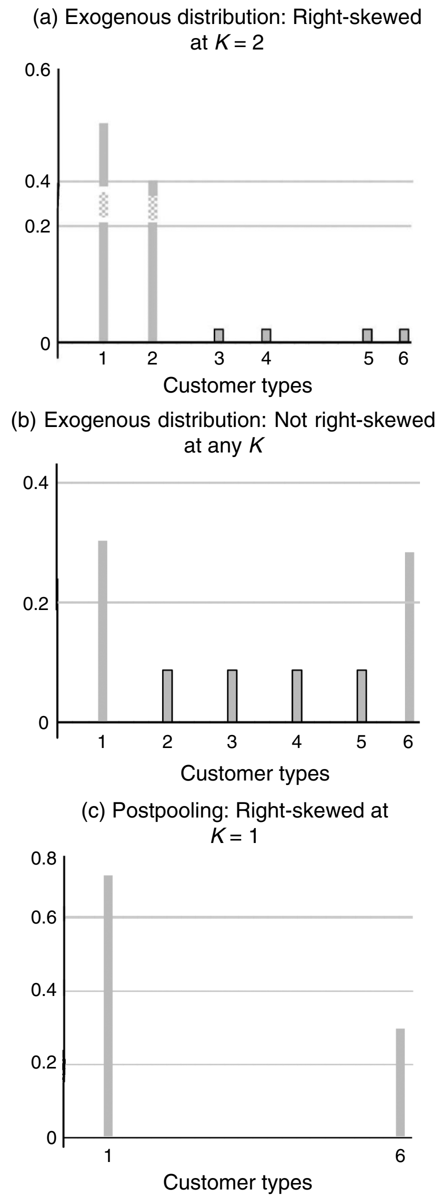
To illustrate the mechanics of Theorem 2, suppose $I = 4$, with $(\lambda_1, \ldots, \lambda_4) = (0.35, 0.45, 0.15, 0.05)$, $v_1(q) = 4 - q$, $v_2(q) = 5 - q$, $v_3(q) = 6 - q$, and $v_4(q) = 7 - q$. The optimal 2PT menu is fully separating, with revenue of 8.66. To design the 3PT menu, we first test Equation (4) at $K = 1$: this test fails (because it requires $0.35 > 3/4$), and it may be verified that the quasi-optimal *single* three-part tariff generates lower revenue (8.17). Next we test the condition at $K = 2$; i.e., $0.45/0.55 > 3/4$. The test succeeds; the two-item 3PT menu produces higher profit.

Theorem 2 can be interpreted as stating that *three-part tariffs are a more efficient way to price discriminate when consumption is dominated by a small group of heavy users*. As to the applicability of the condition, we note that (1) it needs to hold only for some small K (rather than just $K = 1$), and (2) it can be satisfied either by the exogenous distribution of segments or by the "postpooling" distribution created after customers pick their surplus-maximizing tariff from the 2PT menu. These points are graphically demonstrated in Figure 3 and illustrated with the following example.

Example 4. Suppose that there are six types of customers with demand functions $D_i(p) = a_i - p$, with $a_{i+j} = a_i + j\gamma$:

1. If $\lambda_1 = 0.5$ and $\lambda_2 = 0.4$, then the condition fails at $K = 1$ but is satisfied at $K = 2$ (see Figure 3(a)). A two-item 3PT menu outperforms the six-item optimal 2PT menu.
2. If $\lambda_1 = 0.31$, $\lambda_6 = 0.29$ (and the remaining segments are each 10% of the market), the demand distribution is *not* "sufficiently right-skewed" (see Figure 3(b)), and the condition fails at every $K < I$. However, because the optimal menu of two-part tariffs pools segments 1 through 5, we compute the post-pooling λs (thus the revised $\lambda_1 = 0.71$; see Figure 3(c)) and find that the condition is satisfied at $K = 1$. Thus, a single 3PT outperforms the optimal 2PT menu.

Figure 3 Distribution of Customer Types



Notes. The exogenous distribution of segments is not right-skewed. However, the optimal 2PT menu in Example 4 pools types 1–5, and the resulting distribution (of customers that pick the same item from the menu) is skewed to the right.

6.5. Sensitivity to Information Error

Theorems 1 and 2 have demonstrated that when the firm faces a large heterogeneous market, a single three-part tariff (or, more generally, a relatively small menu of such tariffs) produces higher revenue than a menu of two-part tariffs of any size. There is a second benefit: designing the three-part tariff requires precise information about segments $1, \ldots, k+1$ only, less than what is needed to design the optimal 2PT menu. It does not require information about higher-value customer segments, making it unnecessary for the firm to undertake costly information-gathering actions that might infringe on the privacy preferences of its highest-value customers (which it would, when designing the optimal 2PT menu). We start with an example and then provide a formal proof.

Example 5 (Information Requirements). Assume that $I = 3$, with $D_i(p) = a_i - p$ and $\lambda_1 = 0.8$, $\lambda_2 = 0.18$, and $\lambda_3 = 0.02$. The firm knows, correctly, that $a_1 = 1$ and $a_2 = 2$, and it believes that $a_3 = 3$. With this information, the optimal 2PT menu is $(F_1^*, p_1^*) = (0.46875, 0.25)$, $(F_2^*, p_2^*) = (0.533951, 0.111111)$, and $(F_3^*, p_3^*) = (0.861111, 0)$. Using (F_1^*, p_1^*), the corresponding quasi-optimal single 3PT is $(\hat{F} = 0.46875, \hat{Q} = 0.75, \hat{p} = 0.625)$.

Denote the true (or realized) value of a_3 as a_R, and consider the ex post performance of the two tariff designs for various values of $a_R \in [a_2 = 2, 4]$. Let $\Pi_{2PT}(a_R)$ and $\Pi_{3PT}(a_R)$ denote, respectively, the 2PT menu and single 3PT ex post profits when a_R is realized:

1. *2PT menu*: For $a_R \in [3, 4]$, customer 3 will pay F_3^* (because $p_3^* = 0$) and consume a_R, yielding a constant $\Pi_{2PT}(a_R) = 0.526111$. Incorrect information hurts the firm because it should have set a higher F_3^*. For $a_R \in [2, 3]$, customer 3 will choose to buy (F_2^*, p_2^*) and will pay $F_2^* + p_2^*(a_R - p_2^*)$. The quantity $\Pi_{2PT}(a_R)$ increases monotonically, and incorrect information still hurts the firm because it should have set a lower F_3^* and induced this customer to separate.
2. *Single 3PT*: Recall that this is a least optimistic design obtained by collapsing all higher types into type 2 (hence equivalent to assuming $a_3 = 2$). Therefore, the 3PT design will be identical for all values of a_R, and the ex post profit under incorrect information (even though it increases with a_R) will be identical to the perfect-information profit. Hence incorrect information about Segment 3 does not hurt the firm in this case.

We write the percentage superiority of the 3PT over the 2PT menu as a function of the true realization a_R:

$$\mathrm{PctSup}(a_R; a_3) = \frac{\Pi_{3PT}(a_R) - \Pi_{2PT}(a_R)}{\Pi_{2PT}(a_R)} = \begin{cases} \dfrac{0.0026 + 0.0103a_R}{0.5219 + 0.0125a_R} & \text{if } a_R \in [2, 3], \\[2ex] \dfrac{-0.0042 + 0.0125a_R}{0.5219 + 0.0125a_R} & \text{if } a_R \in [3, 4]. \end{cases}$$

Computing the respective ex post profits, we see that as a_R increases from 2 to 4, $\mathrm{PctSup}(a_R; a_3)$ increases in an almost linear fashion from 6.2% to approximately 9% (there is a kink at $a_R = 3$). This means that, with probability 1, the single 3PT will outperform the 2PT menu for any possible realization of a_3. The advantage of a single 3PT will be more pronounced if obtaining the information pertaining to the believed value of a_3 (which is required for designing the 2PT menu) involves nonzero costs. The 3PT design requires knowing only that a_3 is in a certain interval.

Now we prove the information claim formally, first for $k = 1$ and then for higher k. To prove for $k = 1$, we must show that the firm can both *determine* the desirability of a single 3PT and *design* it with less information than needed to compute the 2PT menu. First, consider the determination of desirability, i.e., the threshold $\hat{\lambda}_1$ in Theorem 1. The result is easily seen for the special case where the individual demand functions are linear (Corollary 1) because $\hat{\lambda}_1 = (3b_2)/(b_1 + 3b_2)$ is independent of preferences of type 3 and higher. The claim for the general demand case is proved in the following result.

Corollary 2. *The threshold that ensures superiority of the quasi-optimal three-part tariff* ($\hat{\lambda}_1$ *of Theorem* 1) *depends only on the demands of Segment* 1 *and* 2 *customers.*

Next, consider the design of the single 3PT. From Equation (3), this design only depends on the 2PT design for Segment 1 (i.e., F_1, p_1) and the demand function for Segment 2. From Lemma 1, F_1 and p_1, in turn, also depend only on the demand parameters for Segments 1 and 2. The rest of the model parameters $(D_3, \ldots, D_I, \lambda_2, \ldots, \lambda_I)$ are not needed. We note that one could increase revenue further by fully optimizing the single three-part tariff, but that would require precise information about the demand preferences and distribution of all the segments. Hence we can propose the following managerial heuristic.

When the firm is generally aware that the market is dominated by the low-type segments (i.e., high enough λ_1 as required by Theorem 1), then it is better to (a) invest in learning accurately the preferences of Segments 1 and 2 and then construct a single three-part tariff as specified above, rather than (b) incur greater investments to learn the preferences of all consumer segments and design an optimal 2PT menu.

Finally, we note that the claim about information requirements also applies when the firm must employ a small K-item 3PT. To determine whether Equation (4) is satisfied at some specific k, the firm needs detailed information only for types $1, \ldots, k$. Often, firms can easily identify factors that might signal the need for a short menu of three-part tariffs rather than relying on self-selection from a very large menu. This approach is particularly useful when the consumption pattern (the Ds in our model) can be predicted by demographic factors (such as ethnicity, race, gender, income level) for which the aggregate distribution (the λs in our model) is known at the local or regional level but which cannot be used as the basis for price discrimination.

7. Concluding Remarks

This paper has examined the firm's choice of a pricing scheme when tariff management costs are significant enough to impose a constraint on tariff size. A quantity-price bundle works well when the market has a small number of distinct consumer segments and when tariff size is inconsequential. Otherwise, two-part tariffs have previously been considered to be highly effective and more appropriate than bundles. We show that when it is costly to devise and manage discriminatory tariffs, firms that face a large market with many diverse consumer types are better off using three-part tariffs, a pricing scheme that price discriminates more efficiently. We derived this result using a formulation of discrete consumer types, which can approximate a continuous distribution when the number of types is arbitrarily large.

We designed a quasi-optimal menu of three-part tariffs that outperforms the optimal menu of two-part tariffs (and hence outperforms a 2PT menu of any size) when the distribution of types is sufficiently skewed to the right (more customers in the low-type than in the high-type segments). This condition is observed in many markets—for instance, when a majority of customers are light users whereas a small fraction are heavy users. When a single three-part tariff does not dominate, we show that a short menu of three-part tariffs will generally still be more efficient at price discrimination than a much larger menu of two-part tariffs. The novel insight of our analysis is that the goals of price discrimination can be better achieved by employing additional pricing instruments (i.e., increasing the number of parameters in the tariff) than by inducing more fine-grained self-selection (i.e., increasing the number of items in the menu). For instance, in the last decade, telecommunications firms have introduced many instruments such as rollover minutes, contract periods, off-peak discounts, and friends and family specials.

Our results apply in many business settings that employ three-part tariffs. In addition to the examples discussed in §6.2 (e.g., legal plans, fitness clubs, automobile clubs), applications include telecommunications and IT industries (see, e.g., Mitchell and Vogelsang 1991, Masuda and Whang 2006), parking, healthcare plans, and many services and utilities (Brown and Sibley 1986). For instance, phone companies charge a fixed fee for the line and additional service fees, credit card firms charge merchants (and consumers) an annual fee plus per-transaction fees, bars and nightclubs have cover and per-drink fees, and managed hosting plans charge monthly fee plus fees linked to quantity. Multipart tariffs also play a pricing and coordination role in intra- and interbusiness contracts such as manufacturer–retailer contracts, which often feature advance commitment plus an option to buy additional units (Kuksov and Pazgal 2007, Cachon 2004), and bonus-based compensation packages for top executives or sales personnel (Chen 2005).

Our results should be useful to marketing managers concerned about the cost of price discrimination and about the information necessary to price discriminate. Finally, our findings are also relevant when firms' fears about tariff costs and complexity lead them to offer a uniform price to all customers (either a single per-unit fee or a single fee for unlimited consumption). In such cases, the firm sacrifices profits by opting for too much simplicity. Instead, a single three-part tariff could raise profit substantially without compromising simplicity. Moreover, because the fixed-fee component of a 3PT would deter the lowest-value segments, the firm could extend the pricing scheme to a small menu consisting of one three-part tariff (with fairly low F and Q) and an unlimited consumption price with high F. The evolution of Virgin Mobile's pricing plans is a good illustration. Concerns about tariff simplicity and costs motivated a launch with extremely simple and uniform pricing, but Virgin now employs a small menu of three-part tariffs to get the dual benefits of simplicity and higher revenue.

There are multiple reasons why firms prefer three-part tariffs, and the advantage of 3PTs over 2PTs can be amplified under any of these characteristics. 3PTs are preferable when (1) in the presence of segment development and data collection costs that increase with the length of the menu, (2) there are negative externalities as a result of congestion effects (Masuda and Whang 2006), (3) there are some time-inconsistent (naïve) consumers with uncertain demand who are unaware of the fact that they are time inconsistent (Della Vigna and Malmendier 2006, Eliaz and Spiegler 2006), and (4) consumers exhibit a bias toward tariffs with high allowance, preferring tariffs with high $\hat{Q}$ independent of their consumption patterns (Lambrecht and Skiera 2006). And finally, 3PTs can be used to target overconfident consumers who underestimate the variance of their future demands (Grubb 2009).

Appendix

A.1. Customer Choice Under a Multipart Tariff

Given a multipart tariff, a segment i customer's optimal consumption quantity q_i depends only on D_i and p. Specifically, $q_i = D_i(p)$ because this consumption level produces maximum surplus:

$$S_i(p) = \int_0^{D_i(p)} (v(q) - p)\, dq = \int_p^{\bar{p}_i} D_i(\tau)\, d\tau. \tag{5}$$

The remaining components of the tariff will influence which, if any, tariff to accept. Since a two-part tariff (F, p) implies an outlay $T(q) = F + pq$, the customer will accept the tariff if and only if $S_i(p) \geq F$. Similarly, the customer would accept a three-part tariff (F, Q, p)—which implies an outlay $T(q) = F + p(q-Q)^+$ but also provides Q "free" units—if and only if $S_i(p) + p \cdot \min\{D_i(p), Q\} \geq F + p \cdot \max\{D_i(p) - Q, 0\}$. Finally, given a menu of choices, the customer will accept the item that maximizes $S_i(p) - F$.

A.2. Formal Proofs of Results

Proof of Lemma 1. The seller's problem is

$$\max_{(F_1, p_1), \ldots, (F_I, p_I)} \sum_{i=1}^{I} \lambda_i (F_i + p_i D_i(p_i))$$

$$\text{subject to } F_1 \le \int_{p_1}^{v_1(0)} D_1(\tau)\, d\tau,$$

$$F_2 \le F_1 + \int_{p_2}^{p_1} D_2(\tau)\, d\tau,$$

$$\vdots$$

$$F_I \le F_{I-1} + \int_{p_I}^{p_{I-1}} D_I(\tau)\, d\tau.$$

The above constraints are all binding and can be substituted into the objective function. Then the problem becomes

$$\max_{p_1, \ldots, p_I} \lambda_1 \left[p_1 D_1(p_1) + \int_{p_1}^{a_1} D_1(\tau)\, d\tau \right] + \lambda_2 \left[p_2 D_2(p_2) + \int_{p_1}^{a_1} D_1(\tau)\, d\tau + \int_{p_2}^{p_1} D_2(\tau)\, d\tau \right] + \cdots .$$

The first-order condition (FOC) for p_1 yields

$$p_1 = \frac{\lambda_2 + \cdots + \lambda_I}{\lambda_1} \; \frac{D_2(p_1) - D_1(p_1)}{-D_1'(p_1)}.$$

Similarly, for $1 \le i \le I$, the FOCs yield

$$p_i = \frac{\lambda_{i+1} + \cdots + \lambda_I}{\lambda_i} \; \frac{D_{i+1}(p_i) - D_i(p_i)}{-D_i'(p_i)}. \quad \square$$

Proof of Theorem 1. Both the optimal 2PT menu and the quasi-optimal single 3PT of §6.1 extract the same surplus from Segment 1 customers. Therefore, we only need to show that the superior single 3PT of extracts more surplus than the optimal 2PT from customers in Segments $2, \ldots, I$. Let $\pi^*_{2,i}$ be the surplus extracted from a consumer of type i using the optimal 2PT menu. The total profit under the optimal 2PT is then given by

$$\Pi_{2PT} = \sum_{i=1}^{I} \lambda_i \cdot \pi^*_{2,i}. \tag{6}$$

For $i = 2, \ldots, I$, the Segment i profit in excess of the profit from the *adjacent* segment $i-1$ is

$$\begin{aligned}(\pi^*_{2,i} - \pi^*_{2,i-1}) &= p^*_i (D_i(p^*_i) - D_{i-1}(p^*_{i-1})) + \int_{p^*_i}^{p^*_{i-1}} (D_i(\tau) - D_{i-1}(\tau))\, d\tau \\ &\le p^*_{i-1}(D_i(p^*_i) - D_{i-1}(p^*_{i-1})).\end{aligned} \tag{7}$$

This follows from the fact that the area representing the excess profit is a parallelogram contained in a rectangle with height p^*_{i-1} and base $D_i(p^*_i) - D_{i-1}(p^*_{i-1})$. Now, let $\tilde{\pi}_{2PT,i} = (\pi^*_{2,i} - \pi^*_{2,1})$ denote the *residual* profit from Segment i, defined as the Segment i profit in excess of the profit from Segment 1. Therefore, we can establish an upper bound on this residual as follows:

$$\begin{aligned}\tilde{\pi}_{2PT,i} &= \sum_{j=2}^{i} (\pi^*_{2,j} - \pi^*_{2,j-1}) \\ &\le \sum_{j=2}^{i} p^*_{j-1}(D_j(p^*_j) - D_{j-1}(p^*_{j-1})) \\ &\le p^*_1 \sum_{j=2}^{i} (D_j(p^*_j) - D_{j-1}(p^*_{j-1})) \\ &= p^*_1 (D_i(p^*_i) - D_1(p^*_1)).\end{aligned} \tag{8}$$

For comparison, under the single 3PT, the residual profit from Segment i customer is

$$\tilde{\pi}_{3PT,i} = \hat{p}(D_i(\hat{p}) - \hat{Q}). \tag{9}$$

With the above construction of the residual profits under the two pricing schemes, and as a result of the direction of the inequality in Equation (8),

$$\hat{\Pi}_{3PT} - \Pi_{2PT} \ge \sum_{i=2}^{I} \lambda_i (\tilde{\pi}_{3PT,i} - \tilde{\pi}_{2PT,i}). \tag{10}$$

Now consider what happens when λ_1 is relatively large (close to 1)—equivalently, when $\lambda_1 / (\sum_{i=2}^{I} \lambda_i)$ is large. From Equation (1) it can be inferred that p^*_1 gets very small (since v_i' is finite, D_i' is never 0, and the other term—the ratio $(D_2(p) - D_1(p)) / -D_i'(p)$—in Equation (1) is bounded for all p). Therefore, at the extreme, as $\lambda_1 \to 1$, we know that each $\tilde{\pi}_{2PT,i} \to 0$. And, in general, the residual profits $\tilde{\pi}_{2PT,i}$ in the 2PT menu become very small as λ_1 gets large. By contrast, each of the residuals in the quasi-optimal 3PT is bounded from below: the right-hand side of Equation (9) remains strictly positive because $\hat{p}$ remains bounded from below by $\arg\max_p \{pD_2(p) - D_1(0)\}$, which is strictly positive, and by Assumption 1, implying $D_i(\hat{p}) > \hat{Q}$ for $i > 1$. Putting everything together leads to the conclusion that as λ_1 gets large, the quasi-optimal 3PT generates higher revenue than the optimal menu of two-part tariffs. $\square$

Proof of Theorem 2. The proof works along the lines of the proof for Theorem 1. For the 2PT menu, the revenue from Segment i ($\Pi^*_{2,i}$) and the total revenue Π_{2PT} is as in Equation (6). Now, for Segments $K+1, \ldots, I$, define the "excess revenue over Segment K" for Segment $i = K+1, \ldots, I$, analogous to Equation (7) but replacing 2 with $K+1$. For the 3PT, replace $\hat{p}$ and $\hat{Q}$ with $\hat{p}_K$ and $\hat{Q}_K$. Now, corresponding to Equations (8) and (9), we get the inequalities

$$\tilde{\pi}_{2PT,i} \le p^*_K (D_i(p^*_i) - D_K(p^*_K)),$$
$$\tilde{\pi}_{3PT,i} = \hat{p}_K (D_i(\hat{p}_K) - \hat{Q}_K).$$

From Lemma 1, p^*_K becomes arbitrarily small as the ratio $(\sum_{i=K+1}^{I} \lambda_i)/\lambda_K$ decreases, or equivalently, if $(\sum_{i=K}^{I} \lambda_i) < (1 + z_K)\lambda_K$ (for some $z_K > 0$). Rearranging this term yields the condition $HR(K) > \tau_K$, where $\tau_K = 1/(1 + z_K) < 1$. $\square$

Proof of Corollary 1. Let the optimal 2PT be $(F_1^*, p_1^*), \ldots, (F_n^*, p_n^*)$, and let $\hat{Q} = D_1(p_1^*)$. Equation (18) can be reduced to

$$\hat{p} = \frac{D_2(\hat{p}) - \hat{Q}}{b_2}. \tag{11}$$

Now using Equations (8) and (9) (and the notation introduced in the proof of Theorem 1), we have

$$\begin{aligned}\tilde{\pi}_{3PT,2} - \tilde{\pi}_{2PT,2} &> (\hat{p} - p_1^*)[D_2(\hat{p}) - \hat{Q}] - p_1^*[D_2(p_2^*) - D_2(\hat{p})] \\ &= (\hat{p} - p_1^*)[D_2(\hat{p}) - \hat{Q}] - p_1^* \cdot b_2 \cdot \hat{p} \\ &= (\hat{p} - p_1^*)[D_2(\hat{p}) - \hat{Q}] - p_1^*[D_2(\hat{p}) - \hat{Q}], \end{aligned} \tag{12}$$

where the second equality holds because $b_2 \hat{p} = D_2(\hat{p}) - \hat{Q}$ by Equation (11). For $i > 2$,

$$\begin{aligned}&\hat{p}[D_i(\hat{p}) - D_{i-1}(\hat{p})] - p_i^*[D_i(p_i^*) - D_{i-1}(p_i^*)] \\ &\quad > (\hat{p} - p_i^*)[D_i(\hat{p}) - D_{i-1}(\hat{p})],\end{aligned}$$

where the second inequality holds because $D_{i+1} - D_i$ is nondecreasing in p (recall $b_i \geq b_{i+1}$) and $[D_i(\hat{p}) - D_{i-1}(\hat{p})] > [D_i(p_i^*) - D_{i-1}(p_i^*)]$.

The above inequalities imply that if $\hat{p} > 2p_1^*$, then $\tilde{\pi}_{3PT,i} > \tilde{\pi}_{2PT,i}$ for all $i \geq 2$, and therefore

$$\hat{p} > 2p_1^* \quad \Rightarrow \quad \hat{\Pi}_{3PT} > \Pi_{2PT}. \tag{13}$$

More specifically, for linear demand functions satisfying the assumptions in §6.3, we have

$$\hat{p} = \frac{a_2 - a_1 + b_1 p_1^*}{2b_2}, \tag{14}$$

$$p_1^* = \frac{(1 - \lambda_1)(a_2 - a_1)}{\lambda_1 b_1 + (1 - \lambda_1)(b_2 - b_1)}. \tag{15}$$

Substituting Equation (15) into Equation (14) and simplifying, we obtain

$$\hat{p} = \frac{(a_2 - a_1)}{2b_2} \left(\frac{\lambda_1 b_1 + (1 - \lambda_1) b_2}{\lambda_1 b_1 + (1 - \lambda_1)(b_2 - b_1)} \right). \tag{16}$$

Now using Equations (16) and (15), the first inequality in (13) becomes simply

$$\lambda_1 > \frac{3b_2}{b_1 + 3b_2}, \tag{17}$$

and whenever the above inequality holds, a single 3PT outperforms any menu of 2PTs. □

Proof of Corollary 2. We will show that $\hat{\lambda}_1$ can be computed by solving an inequality of the form $f(g(\lambda_1)) > h(g(\lambda_1))$, where g, f (given in Equations (18) and (19)), and h are functions of λ_1 whose forms depends only on D_1 and D_2.

Using an argument similar to the one used in the proof of Corollary 1 (Equations (12) and (13)), there is a function h that depends only on D_1 and D_2 such that if $\hat{p} > h(p_1^*)$, then $\tilde{\pi}_{3PT,2} - \tilde{\pi}_{2PT,2} > 0$. Moreover, for $i > 2$, we have $\tilde{\pi}_{3PT,i} - \tilde{\pi}_{2PT,i} > 0$. Therefore, if $\hat{p} > h(p_1^*)$, then a single 3PT beats the optimal 2PT.

Now, consider the design of the 3PT. The $\hat{F}$ and $\hat{Q}$ components of the quasi-optimal 3PT depend only on the p_1^* component of the 2PT menu (see Equation (3)). From Lemma 1, we apply Equation (1) to $i = 1$, which yields

$$p_1^* = Sol \cdot \left[p = \left(\frac{1 - \lambda_1}{\lambda_1} \right) \frac{1}{-D_1'(p)} (D_2(p) - D_1(p)) \right] = g(\lambda_1), \tag{18}$$

where g is some function that depends on D_1 and D_2 only, not on the demand specification for the remaining Segments $i = 3, \ldots, I$. Knowledge of these parameters immediately yields $\hat{F}$ and $\hat{Q}$. To compute the $\hat{p}$ component of the single 3PT, we use Equation (3) to get

$$\hat{p} = Sol \cdot \left[p = \frac{D_2(p) - D_1(p_1^*)}{-D_2'(p)} \right] = f(p_1^*) \tag{19}$$

for some function f. Therefore $\hat{p}$, like p_1^*, depends only on D_1, D_2, and λ_1, and the firm need not know the detailed distribution or preferences of Segments $3, \ldots, I$. □

References

Acquisti A, Varian HR (2005) Conditioning prices on purchase history. *Marketing Sci.* 24(3):367–381.
Armstrong M, Vickers J (2001) Competitive price discrimination. *RAND J. Econom.* 32(4):579–605.
Bagnoli M, Bergstrom T (2005) Log-concave probability and its applications. *Econom. Theory* 26(2):445–469.
Barlow RE, Marshall AW, Proschan F (1963) Properties of probability distributions with monotone hazard rate. *Ann. Math. Statist.* 34(2):375–389.
Brown SJ, Sibley DS (1986) *The Theory of Public Utility Pricing* (Cambridge University Press, New York).
Cachon GP (2004) The allocation of inventory risk in a supply chain: Push, pull, and advance-purchase discount contracts. *Management Sci.* 50(2):222–238.
Chen F (2005) Salesforce incentives, market information, and production inventory planning. *Management Sci.* 51(1):60–75.
Chen Y, Iyer G (2002) Consumer addressability and customized pricing. *Marketing Sci.* 21(2):197–208.
Clay KB, Sibley DS, Srinagesh P (1992) Ex post vs. ex ante pricing: Optional calling plans and tapered tariffs. *J. Regulatory Econom.* 4(2):115–138.
Cowan S (2007) The welfare effects of third-degree price discrimination with nonlinear demand functions. *RAND J. Econom.* 38(2):419–428.
DellaVigna S, Malmendier UM (2006) Paying not to go to the gym? *Amer. Econom. Rev.* 96(3):694–719.
Dolan RJ (1984) Quantity discounts: Managerial issues and research opportunities. *Marketing Sci.* 6(1):1–22.
Eliaz K, Spiegler R (2006) Contracting with diversely naïve agents. *Rev. Econom. Stud.* 73(3):689–714.
Faulhaber GR, Panzar JC (1977) Optimal two-part tariffs with self-selection. Bell Laboratories Economic Discussion Paper 74, Bell Laboratories, Murray Hill, NJ.
Goldman MB, Leland HE, Sibley DS (1984) Optimal nonuniform prices. *Rev. Econom. Stud.* 51(2):305–319.
Grubb MD (2009) Selling to overconfident consumers. *Amer. Econom. Rev.* 99(5):1770–1807.
Iyengar R, Gupta S (2009) Nonlinear pricing. Rao V, ed. *Handbook of Pricing Research in Marketing* (Edward Elgar Publishing, Northampton, MA), 355–383.
Iyengar SS, Lepper MR (2000) When choice is demotivating: Can one desire too much of a good thing? *J. Personality Soc. Psych.* 79(6):995–1006.
Johnson JP, Myatt DP (2003) Multiproduct quality competition: Fighting brands and product line pruning. *Amer. Econom. Rev.* 93(3):748–774.

Kolay S, Shaffer G (2003) Bundling and menus of two-part tariffs. *J. Indust. Econom.* 51(3):383–403.

Kuksov D, Pazgal A (2007) The effects of costs and competition on slotting allowances. *Marketing Sci.* 26(2):259–267.

Lambrecht A, Skiera B (2006) Paying too much and being happy about it: Existence, causes, and consequences of tariff-choice biases. *J. Marketing Res.* 43(2):212–223.

Lambrecht A, Seim K, Skiera B (2007) Does uncertainty matter? Consumer behavior under three-part tariffs. *Marketing Sci.* 26(5):698–710.

Maskin E, Riley J (1984) Monopoly with incomplete information. *RAND J. Econom.* 15(2):171–194.

Masuda Y, Whang S (2006) On the optimality of fixed-up-to tariff for telecommunications service. *Inform. Systems Res.* 17(3):247–253.

McGovern G (2003) Virgin Mobile USA: Pricing for the very first time. HBS Case 9-504-028, Harvard Business School, Boston.

Miravete EJ (2007) The limited gains from complex tariffs. Working paper, Department of Economics, University of Texas at Austin, Austin.

Mitchell BM, Vogelsang I (1991) *Telecommunications Pricing: Theory and Practice* (Cambridge University Press, Cambridge, UK).

Murphy MM (1977) Price discrimination, market separation, and the multi-part tariff. *Econom. Inquiry* 15(4):587–599.

Odlyzko A (2003) Privacy, economics, and price discrimination on the internet. *ICEC '03: Proc. 5th Internat. Conf. Electronic Commerce* (ACM, New York), 355–366.

Oren SS, Smith SA, Wilson RB (1982) Nonlinear pricing in markets with interdependent demand. *Marketing Sci.* 1(3):287–313.

Png IPL, Wang H (2010) Buyer uncertainty and two-part pricing: Theory and applications. *Management Sci.* 56(2):334–342.

Raju J, Zhang ZJ (2005) Channel coordination in the presence of a dominant retailer. *Marketing Sci.* 24(2):254–262.

Rochet JC, Stole LA (2002) Nonlinear pricing with random participation. *Rev. Econom. Stud.* 69(1):277–311.

Shaked M, Shanthikumar JG, Valdez-Torres JB (1995) Discrete hazard rate functions. *Comput. Oper. Res.* 22(4):391–402.

Sharkey WW, Sibley DS (1993) Optimal non-linear pricing with regulatory preference over customer type. *J. Public Econom.* 50(2):197–229.

Shy O (2008) *How to Price: A Guide to Pricing Techniques and Yield Management* (Cambridge University Press, New York).

Stole LA (2007) Price discrimination and competition. Armstrong M, Porter RH, eds. *Handbook of Industrial Organization*, Vol. 3 (Elsevier, Amsterdam), 2221–2299.

Sundararajan A (2004) Nonlinear pricing of information goods. *Management Sci.* 50(12):1660–1673.

Villas-Boas JM (2004) Communication strategies and product line design. *Marketing Sci.* 23(3):304–316.

Wedel M, Kamakura W (2000) *Market Segmentation: Conceptual and Methodological Foundations* (Kluwer Academic Publishers, Dordrecht, The Netherlands).

Wilson RB (1993) *Nonlinear Pricing* (Oxford University Press, New York).