Measurement of Consumer Preferences for Bucket Pricing Plans with Different Service Attributes

Christian Schlereth,

Bernd Skiera

Date: 2011-08-31

Status: Accepted for publication by the International Journal of Research in Marketing

Subject Areas: 1.380: Services Marketing, 1.290: Pricing, 1.100: Decision-Making

Method Types: 2.020: Choice Modeling, 2.010: Bayesian Estimation Techniques, 2.200: Survey Research

* Dr. Christian Schlereth, Assistant Professor, Goethe University Frankfurt, Faculty of Business and Economics, Department of Marketing, Grueneburgplatz 1, 60323 Frankfurt am Main, Germany, Phone: ++49-69-798-34638, Fax: ++49-69-79-35001, email: schlereth@wiwi.uni-frankfurt.de

Prof. Dr. Bernd Skiera, Chaired Professor, Goethe University Frankfurt, Faculty of Business and Economics, Department of Marketing, Grueneburgplatz 1, 60323 Frankfurt am Main, Germany, Phone: ++49-69-798-34649, Fax: ++49-69-79-35001, email: skiera@skiera.de

Measurement of Consumer Preferences for Bucket Pricing Plans with Different Service Attributes

Abstract

A bucket pricing plan charges a periodic (usually monthly) fixed price that allows consumers to use the service up to a set allowance. The determination of optimal plans requires knowledge about each consumer's simultaneous decision about service subscription, plan choice, and consumption, which are interrelated and difficult to predict. Beside prices, service attributes also influence the three decisions, but how they do depends on the particular service attribute. This article describes a novel method to predict consumers' reactions to bucket pricing plans with varying service attributes and develops an algorithm to optimize bucket pricing plans. Methodologically, we show that the failure to model the influence of service attributes correctly leads to non-optimal prices and profits that differ by up to 22.75% from the optimal solution. Substantially, we show that bucket pricing plans are approximately as profitable as other nonlinear pricing plans if at least three bucket pricing plans serve to segment the market. Bucket pricing plans therefore present an attractive alternative for service providers to differentiate consumers according to their WTP and consumption.

Keywords: pricing, willingness to pay, discrete choice experiments, Bayesian estimation, bucket pricing, two-part pricing plans, three-part pricing plans

1 Introduction

With a bucket pricing plan, a service provider charges a periodic (usually monthly) fixed price, in exchange for which consumers may use the service without further charges, up to a preset allowance during the period. Due to the huge heterogeneity among consumers, most providers offer more than one bucket pricing plan to differentiate consumers with different demand. Thus, consumers can choose freely among several combinations of fixed prices and allowances. If they use up their allowance, they might change to another plan with a higher allowance or stop using the service for the remaining time of the period. The purchase of single units of the service is not possible, but changes in the plan are allowed (typically once a month) without any additional charges.

Bucket pricing plans appear increasingly in various industry sectors. With the introduction of Apple's iPad in 2010, the mobile broadband service provider AT&T waived its unlimited data plans and began charging $14.99 per month for 250 MB (this translates to 6 cents per MB of allowance) and $25.00 per month for 2 GB (1 cent per MB). In the music industry, eMusic offered three monthly bucket pricing plans in 2007 for downloads: 30 songs per month for $12.99, 50 songs for $16.99, or 75 songs for $20.99. Cloud computing service vendors such as Epidirect also charge a monthly price that depends on the number of transactions (e.g., 2000 transactions for $99.95, 5000 for $199.95). Even car insurance companies such as Allianz offer bucket pricing plans with allowances of 6000 km, 10,000 km, or more kilometers. The health care industry uses bucket pricing too. Evolution to Wellbeing, a personal training provider, charges $160 per month for up to 8 training sessions ($20 per session), $200 monthly for up to 12 ($17 per session), or $220 for up to 24 sessions ($9 per session).

Despite the prevalence of bucket pricing, research on the topic remains scarce (see Table 1). Compared with two-part pricing plans (i.e., a usage-independent fixed price plus a marginal price per unit), bucket pricing plans enables consumers to benefit from a fixed allowance, rather than paying for each unit of consumption. Thus, payment is separated from consumption, which enables consumers to enjoy their consumption more, according to the theory of mental accounting (Prelec and Loewenstein 1998).[1] In addition, the monthly price stays the same and is known at the beginning of each month, which is often preferable for risk-averse consumers (Lambrecht and Skiera 2006). Compared with three-part pricing plans (i.e., a usage-independent fixed price, an allowance, and a marginal price per unit), bucket pricing does not allow users to purchase single units of the service if they exceed the allowance. Most providers that use bucket pricing thus allow consumers to change and even cancel their plan each month, such that consumers face only a short-term commitment. This ability represents a major difference from most three-part pricing plans, which require consumers to retain their selected plan over time, often for as long as 24 months (Lambrecht, Seim, and Skiera 2007). Similar to most three-part pricing plans, unused allowance expires after the end of each period though, so that consumers might feel regret, because they have paid for something that they did not use.

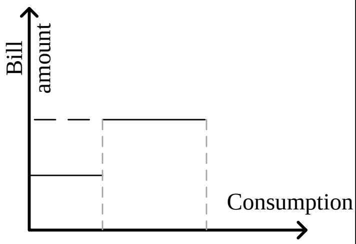
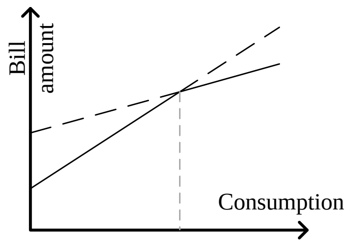
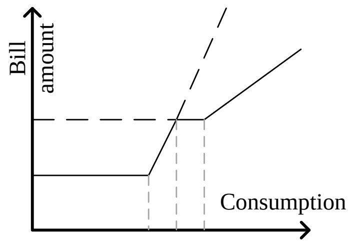
Figure 1 illustrates the differences between bucket pricing plans as well as two-part and three-part pricing plans. The consumer can usually select among several pricing plans. In Figure 1, the dotted line marks the pricing plan, which is more expensive than the other plan for a given consumption. Bucket pricing plans lead to a stepwise function for the bill amount,

---

1 We acknowledge that the situations of payments and hedonic consumption, which are studied in Prelec and Loewenstein (1998), occur only once. In contrast, bucket pricing plans involve reoccurring charges at the beginning of each period. Clear evidence that prepayment is also desired for recurring expenses and that the theory of mental accounting still holds, is yet missing, despite the promising findings of Lambrecht and Skiera (2006).

which only increase if consumers move to the next expensive pricing plan. Two part-pricing plans, in contrast, increase the bill amount with every unit that is consumed. Three-part pricing plans share characteristics of bucket pricing plans (in particular, the stepwise function) and two-part pricing plans (in particular, the increasing part). A bucket pricing plan is a special case of a three-part pricing plan because its marginal price is infinite, i.e. it is so high that no consumer will be willing to pay that price. A two-part pricing plan is also a special case of the three-part pricing plan because its allowance is zero.

FIGURE 1: VISUALISATION OF BUCKET PRICING PLANS OPPOSED TO TWO-PART AND THREE-PART PRICING PLANS

	Bucket Pricing Plans	**Two-Part Pricing Plans**	**Three-Part Pricing Plans**
			
Components	p_j: fixed price q_j: allowance	p_j: fixed price m_j: marginal price per unit	p_j: fixed price q_j: allowance m_j: marginal price for each unit above the allowance
Bill amount	$R_j(p_j) = p_j$	$R_{i,j}(p_j, m_j) = p_j + m_j \cdot n_{i,j}$	$R_{i,j}(p_j, q_j, m_j) = p_j + m_j \cdot \max(0; n_{i,j} - q_j)$
Example	eMusic download plans (2007): • $12.99 for 30 songs • $16.99 for 50 songs • $20.99 for 75 songs	DB Bahn (2011): • 0 € per year and 0% discount • 57 € per year and 25% discount • 230 € per year and 50% discount • 3800 € per year and 100% discount on each train ticket	AT&T iPhone voice plans (2011): • $39.99 for 450 minutes and $0.45 for each additional minute • $59.99 for 900 minutes and $0.40 for each additional minute

$n_{i,j}$ = consumption of a consumer i, when choosing pricing plan j

The determination of optimal bucket pricing plans remains challenging. If service providers offer more than one plan, as in Figure 1, a consumer has to decide simultaneously on subscription, plan choice, and consumption, which are interrelated and difficult to predict. The reason is the interdependency between prices and consumption (Iyengar, Jedidi, and Kohli 2008; Lambrecht, Seim, and Skiera 2007). For example, if AT&T decreases the price of their 250 MB plan, the number of consumers choosing that plan will likely increase. However, each consumer in the 250 GB plan generates less profit and consumers switching from a higher plan will use the service less frequently, which is likely to reduce profit.

The simultaneous prediction of consumers’ decisions on subscription, plan choice, and consumption becomes even more complicated if service providers vary in service attributes to differentiate their service offerings further (Eggers and Sattler 2009). The reason is that different types of service attributes may influence the three decisions in a different way. For example, Apple’s iPad comes in versions with varied storage capacities (e.g., 16, 32, or 64 GB) and access to 3G networks. The differences in storage capacity might have little influence on consumption of 3G network capacity, because most bandwith-extensive downloads move over WiFi networks, and even the smallest storage capacity is still large enough for the bandwith provided by current 3G network operators. Differences in the availability of 3G networks instead, which can vary strongly across operators, might hugely impact the consumption of capacity. Despite the practical importance of this distinction, all previous research applied an “one-size-fits-all approach” (see Section 2.3); most of the studies link service attributes primary on the subscription decision and thereby neglect their potential influence on consumption.

This article aims to develop a novel method to predict consumer's reactions to bucket pricing plans that also differ in their service attributes, as well as to develop an algorithm to optimize a menu of bucket pricing plans. In particular, we attempt to examine if service attributes affect plan choice in relation to the subscription decision or the consumption decision. This difference is important, because the two mechanisms affect willingness to pay and the resulting optimal prices differently. Furthermore, we analyze the profit differences between bucket and other popular pricing plans. Because transaction data rarely are available for such innovative offerings, we use survey data, elicited through discrete choice experiments (e.g., Moore 2004; Wedel et al. 1998).

In Section 2, we outline previous research on nonlinear pricing to develop our model of consumer decision-making for bucket pricing plans in Section 3. We also outline how service attributes might vary in their influence on subscription and consumption decisions. In an empirical study in Section 4, we measure preferences for a music download service that uses bucket pricing and compare the validity of models that predict different influences of service attributes on consumer's reactions to bucket pricing plans. In Section 5, we also compare bucket pricing plans against other popular plans, such as two- and three-part pricing plans. Finally, we conclude with a summary of our findings.

2 Literature Review

Table 1 summarizes previous studies on consumers' reactions to nonlinear pricing plans, which are plans in which the price per unit is not strictly proportional to the number of units purchased. It also outlines how our research differs from previous studies.

TABLE 1: SUMMARY OF STUDIES THAT MEASURE CONSUMERS' REACTIONS TO NONLINEAR PRICING PLANS

		Nonlinear Pricing Plans			Empirical Data Sources			Influence of Service Attributes				Counterfactual Simulations			
General Topic	**Research Work**	**Two-part pricing plans**	**Three-part pricing plans**	**Bucket pricing plans**	**No empirical data**	**Revealed preferences**	**Stated preferences**	**No influence**	**Subscription decision**	**Consumption decision**	**Depends on service attribute**	**Sensitivity analysis**	**Optimization with ≤ 2 optional plans**	**Optimization with > 2 optional plans**	**Comparisons of nonlinear pricing plans**
Demand under different pricing plans	Danaher (2002)	X				X		X				X	X		
Usage uncertainty and/or learning	Narayanan, Chintagunta, and Miravete (2007)	X				X		X				X	X		
	Iyengar, Ansari, and Gupta (2007)		X			X			X			X			
	Lambrecht, Seim, and Skiera (2007)		X			X			X			X			
	Iyengar, Jedidi, and Kohli (2008)		X				X		X			X	X		
	Goettler and Clay (2010)		X			X		X				X	X		
	Ascarza, Lambrecht, and Vilcassi (2010)		(X)			(X)			(X)			(X)			
Comparison across different specification	Iyengar (2010)			(X)			(X)		(X)			(X)	(X)		
	Iyengar and Jedidi (2010)			(X)			(X)			(X)		(X)	(X)		
Comparison across different data collection methods	Schlereth, Skiera, and Wolk (2011)	X					X	X					X		
Optimization of pricing plans	Hui, Yoo, and Tam (2007)		X		X			X				X	X		
	Schlereth, Stepanchuk, and Skiera (2010)	X			X			X						X	X
Service attributes and bucket pricing	**Our paper**			**X**			**X**				**X**			**X**	**X**

Notes: (X) indicates working paper.

2.1 Nonlinear pricing plans

Previous research on nonlinear pricing has its origins in welfare economics and primarily has tackled consumer decision-making for two-part pricing plans (i.e., fixed price and a marginal price per unit; e.g., Hui, Yoo, and Tam 2007; Leland and Meyer 1976; Murphy 1977). The contributions include analytical models that reveal interdependence between prices and consumption. Revealed preference data from field experiments (Danaher 2002) and actual market transactions (Goettler and Clay 2010; Narayanan, Chintagunta, and Miravete 2007) also provide empirical evidence of such interdependence. Transaction data have been used to study three-part pricing plans (e.g. Iyengar, Ansari, and Gupta 2007; Lambrecht, Seim, and Skiera 2007) in studies that also account for consumer’ uncertainty about preferences, quantity in subsequent periods, and learning effects, when making long-term decisions.

Unlike three-part pricing plans, bucket pricing plans include no single-usage price (i.e., no usage is possible beyond the allowance) but allow consumers to switch easily across different plans. With the exception of working papers by Iyengar (2010) and Iyengar and Jedidi (2010), who compare alternative specifications of the utility function and show that differences in internal validity are modest, previous research has ignored such pricing though. Our work differs from theirs, because neither Iyengar (2010) nor Iyengar and Jedidi (2010) consider the potential effects of service attributes on consumers’ choice and consumption decisions or the ability of bucket pricing plans to generate profit, when compared to other nonlinear pricing plans.

2.2 Data sources

Revealed preference data have high external validity, such that they support estimates of consumer learning over time (Iyengar, Ansari, and Gupta 2007; Lambrecht, Seim, and Skiera

2007; Narayanan, Chintagunta, and Miravete 2007). Yet prices in revealed preference data often vary only in a limited range, so estimations of reactions to price changes are difficult. Moreover, revealed preferences are unavailable for companies that enter new markets or sell new services not previously sold in real market conditions (Swait and Andrews 2003; Wertenbroch and Skiera 2002). Stated preference methods can offer assistance in such situations, which is why we focus on them in this study (Eggers and Sattler 2009; Louviere, Hensher, and Swait 2000). Iyengar, Jedidi, and Kohli (2008) also explain how to use discrete choice experiments to estimate demand for three-part pricing plans in a cellular phone service context. This data source is inexpensive and provides good control over the experimental setting to enable tests of consumers' reactions to new attribute ranges and pricing.

2.3 Service attributes

Prior consideration of service attributes in models that capture consumers' responses to nonlinear pricing plans is very limited; most models focus on homogeneous services that do not differ in their attributes (e.g., Hui, Yoo, and Tam 2007; Maskin and Riley 1984; Narayanan, Chintagunta, and Miravete 2007). However, a few recent models have started to account for differences in service attributes, such as brands, rollover minutes, Internet access (Iyengar, Jedidi, and Kohli 2008), service quality, catalogue size (Iyengar 2010), switching costs (Ascarza, Lambrecht, and Vilcassim 2010), and pricing plan types (Iyengar, Ansari, and Gupta 2007; Lambrecht, Seim, and Skiera 2007). These studies consistently assume that every service attribute affects the same behavioral process, such as consumers' usage-independent utility, which then influences the likelihood of subscribing to a service. Only one working paper (Iyengar and Jedidi 2010) anticipates that all service attributes affect the perceived utility of each service unit and thus consumption.

Unlike these prior studies, we account for influences on different behavioral processes and therefore expect that some service attributes affect the consumption decision, whereas others affect the subscription decision. We propose a flexible extension to a bucket pricing model that can distinguish among different influences of the varying types of service attributes. This extension could be included easily into other models of nonlinear pricing. We also outline the extent that the failure to account for differences in service attributes can lead to non-optimal pricing recommendations.

2.4 Counterfactual simulations

Wilson (1993) introduces a rich framework of descriptive statistics to understand consumers' demand behavior, which is rooted in statistics for uniform pricing plans. Yet it cannot inform bucket pricing decisions, because service providers still require a full understanding of the holistic system, including all relevant interactions across optional pricing plans and consumption rates. Counterfactual simulations can complement such understanding and enable an analysis of market conditions with different pricing plans. For example, prior research has studied consumption sensitivity and elasticity when plans change, such as shifts in choice probabilities in response to pricing changes (Iyengar, Jedidi, and Kohli 2008; Lambrecht, Seim, and Skiera 2007) or alterations in pricing plan choices, consumption, and profits if decision-making uncertainty increases (Ascarza, Lambrecht, and Vilcassim 2010; Lambrecht, Seim, and Skiera 2007).

Danaher (2002), Iyengar, Jedidi, and Kohli (2008), and Iyengar (2010) use grid search techniques as a heuristic to determine profit-maximizing prices, though these solution spaces increase exponentially with more pricing plans. This exponential increase might explain why most studies consider only one or two pricing plans. Schlereth, Stepanchuk, and Skiera

(2010) show that simulated annealing reduces this shortcoming and can easily optimize four plans. Additionally, they show that two optional two-part pricing plans frequently are sufficient to skim the market. Yet these authors do not reveal how to optimize bucket or three-part pricing plans. Consequently, we have little knowledge about how pricing plans compare to one another, despite the great importance of such understanding for service providers that must determine which type and number of pricing plans to use. We follow Schlereth, Stepanchuk, and Skiera (2010) and develop a simulated annealing algorithm to optimize bucket pricing and three-part pricing plans. We then extend this application to compare their profitability.

3 Model Development

We develop a model for choice decisions about bucket pricing plans with different service attributes. To do so, we begin with a basic model that considers bucket pricing for homogeneous services that do not differ in their attributes, then specify prior assumptions about the distribution of the parameters in the willingness to pay (WTP) space (also known as the surplus model). Next, we introduce a flexible extension of the model for heterogeneous services that describes how differences in service attributes might influence plan choice through either the subscription or the consumption decision. Finally, we discuss the different implications that arise and outline the application of hierarchical Bayes techniques to estimate parameters.

3.1 Modeling plan choice for bucket pricing

Let J represent the set of bucket pricing plans that can be chosen by all consumers during each period. Each plan $j \in J$ consists of an allowance q_j and a periodic, fixed price p_j. We assume a utility-maximizing consumer i who does not choose more than one plan. This consumer consumes $n_{i,j}(q_j)$ units of the service, depending on his or her preferences and the allowance q_j ($0 \leq n_{i,j}(q_j) \leq q_j$). The amount of consumption $n_{i,j}(q_j)$ in a period can occur all at once or be summed over different consumption phases within that period.

To analyze data with multiple units of a homogeneous service, we need a nonlinear utility specification of the utility function to capture the unique WTP for each quantity unit (Iyengar, Jedidi, and Kohli 2008; Lambrecht, Seim, and Skiera 2007). Let $v_{i,j}(n_{i,j}(q_j),z_{i,j})$ represent the deterministic part of the utility that consumer i obtains by choosing bucket pricing plan j, consuming $n_{i,j}(q_j)$ units of the service, and spending the remaining budget on $z_{i,j}$ units of an unobservable outside good. The outside good provides a utility of $\varpi_i \cdot z_{i,j}$, where ϖ_i represents the price parameter (Sonnier, Ainslie, and Otter 2007). We assume that utility increases with quantity, at a decreasing rate. We choose a quadratic functional form (Iyengar, Jedidi, and Kohli 2008; Lambrecht, Seim, and Skiera 2007; Narayanan, Chintagunta, and Miravete 2007), which results in the following deterministic part of the direct utility function:

$$(1) \qquad v_{i,j}\left(n_{i,j}\left(q_j\right), z_{i,j}\right) = a_{i,j} \cdot n_{i,j}\left(q_j\right) - \frac{b_{i,j}}{2} \cdot n_{i,j}\left(q_j\right)^2 + c_{i,j} + \varpi_i \cdot z_{i,j} \qquad (i \in I,\ j \in J),$$

subject to a budget constraint:

$$(2) \qquad Y_i \geq z_{i,j} \cdot p_z + p_j \qquad (i \in I,\ j \in J),$$

where Y_i represents the budget of consumer i, and p_z is the price of the outside good. Furthermore, the parameter $a_{i,j}$ reflects the increase of utility that accompanies an increase in consumption; $b_{i,j}$ accounts for the decrease in the marginal utility; and $c_{i,j}$ is for the usage-independent utility, that is, the utility for zero quantity. This usage-independent utility equals 0 for many services, but it might be greater than 0 if the service provider includes services that are free of additional charge in a subscription, such as subsidized hardware that comes with telecommunication services or free website space bundled with Internet access.

The index j of the parameters $a_{i,j}$, $b_{i,j}$, and $c_{i,j}$ reflects the differences in the attributes of a pricing plan j, as we describe in more detail subsequently. To ensure a semi-concave function, the parameters $a_{i,j}$, $b_{i,j}$, and $c_{i,j}$ should be greater than or equal to 0. Without loss of generality, we normalize the price of the outside good p_z to 1. Assuming that a consumer exhausts his or her budget, and substituting the rearranged term of the budget constraint in Equation (2), $z_{i,j} = Y_i - p_j$, into Equation (1), we obtain an indirect utility function:

$$(3) \qquad v_{i,j}\left(n_{i,j}\left(q_j\right), p_j\right) = a_{i,j} \cdot n_{i,j}\left(q_j\right) - \frac{b_{i,j}}{2} \cdot n_{i,j}\left(q_j\right)^2 + c_{i,j} + \varpi_i \cdot (Y_i - p_j) \qquad (i \in I,\ j \in J).$$

Thus according to Equation (3), $u_{i,0} = \varpi_i \cdot Y_i$ if consumers do not choose the service, such that they spend all their money on the unobservable outside good. By forming a Lagrange function, we derive the optimal consumption $n^*_{i,j}$ of consumer i for bucket pricing plan j (see the Appendix):

$$(4) \qquad n^*_{i,j}\left(q_j\right) = \begin{cases} q_j & ,\text{if } q_j \le \dfrac{a_{i,j}}{b_{i,j}} \\ \dfrac{a_{i,j}}{b_{i,j}} & ,\text{if } q_j > \dfrac{a_{i,j}}{b_{i,j}} \end{cases} \qquad (i \in I,\ j \in J).$$

This model predicts that consumers either fully use the allowance q_j or leave some of it unused, if his or her saturation level (i.e., $a_{i,j}/b_{i,j}$) is below the allowance. In this case, consuming an additional unit would not provide any additional benefit for the consumer but might cause some disutility (e.g., opportunity cost of time, as implicitly considered herein). Substituting Equation (4) in Equation (3), we can derive the indirect utility function as a function of the allowance q_j. Equation (5) then uses the indicator variable $\mathrm{Ind}_{i,j}$ to distinguish the increase in utility with every unit below the saturation level from the constant part of the utility function for every unit above the saturation level. We assume that the consumer does not consume the $\left(\frac{a_{i,j}}{b_{i,j}}+1\right)$ th unit of the service, which provides no additional utility. Thus inserting Equation (4) in Equation (3) gives:

(5)
$$v_{i,j}(q_j,p_j)=\mathrm{Ind}_{i,j}\cdot\left(a_{i,j}\cdot q_j-\frac{b_{i,j}}{2}\cdot q_j^2\right)+ (1-\mathrm{Ind}_{i,j})\cdot\left(\frac{a_{i,j}^2}{2\cdot b_{i,j}}\right)+c_{i,j}+\varpi_i\cdot(Y_i-p_j) \qquad (i\in I,\ j\in J),$$

where $\mathrm{Ind}_{i,j}=\begin{cases}1, & \text{if } q_j\le\frac{a_{i,j}}{b_{i,j}}\\ 0, & \text{otherwise}\end{cases}$.

Finally, we transform the indirect utility function in Equation (5) into a surplus function (Sonnier, Ainslie, and Otter 2007) that directly expresses WTP for a service. This transformation is motivated by the belief that consumers typically prefer to think in monetary rather than utility terms, a preference that must be taken into account by appropriate prior distribution specifications in hierarchical Bayesian estimation methods. Similar to classical additive utility models (Sonnier, Ainslie, and Otter 2007), we transfer the surplus model modification to the context of services and determine the parameters $a_{i,j}$, $b_{i,j}$, and $c_{i,j}$ as monetary values (see the Appendix):

(6)
$$a_{i,j}^M=\frac{a_{i,j}}{\varpi_i};\ b_{i,j}^M=\frac{b_{i,j}}{\varpi_i};\ \ c_{i,j}^M=\frac{c_{i,j}}{\varpi_i} \qquad (i\in I,\ j\in J).$$

With this transformation, we can rewrite the indirect utility function in Equation (5) as a surplus model:

$$\text{(7)} \qquad v_{i,j}(q_j, p_j) = \varpi_i \cdot \left(\mathrm{Ind}_{i,j} \cdot \left(a_{i,j}^{M} \cdot q_j - \frac{b_{i,j}^{M}}{2} \cdot q_j^2 \right) + (1 - \mathrm{Ind}_{i,j}) \cdot \left(\frac{a_{i,j}^{M\,2}}{2 \cdot b_{i,j}^{M}} \right) + c_{i,j}^{M} - p_j \right) + \varpi_i \cdot Y_i \qquad (i \in I,\ j \in J).$$

3.2 Service attributes

Few studies include service attributes in models of nonlinear pricing, and when they do, they model all service attributes in an identical way. For example, Iyengar, Jedidi, and Kohli (2008) feature brand names, Internet access, and roll-over minutes in their choice model for three-part pricing plans. It is reasonable to assume that the availability of Internet access from a cell phone provides additional utility but does not alter the number of calls made by a consumer. Like in most other studies, these authors link differences in the utility of service attributes to usage-independent WTP, which we capture with the parameter $c_{i,j}^{M}$. This linkage implies that the attributes have no influence on the consumption decision (i.e., number of units consumed) but make the general offering of the service more or less attractive, which affects the likelihood of subscription.

What is neglected yet in prior literature is, that other attributes might require a different linkage. Consider downloads for digital music and the effects of differences in digital rights management (DRM) restrictions. As an access control technology, DRM frequently is applied by the music industry to prevent the unauthorized copying and distribution of purchased songs. For example, DRM can restrict consumers’ ability to use downloaded songs on different players or computers or the number that songs that may be burned on a CD. These restrictions likely affect consumption, because removing DRM restrictions would increase opportunities for consumers to use their purchased songs (e.g., Sundararajan 2004). However, the utility function only imagines a consumption increase if the differences in these attributes affect the parameter $a_{i,j}^{M}$ of the utility function.

We extend our model to consider heterogeneity in the quality of these services as well. We let x_j be the vector of dummy variables that indicate the service attributes of plan j. We then introduce two binary indicator vectors, l^a and l^c of size |M|, which is the sum of the levels of all the service attributes. The value of an element in the vector l^a (l^c) equals 1 if the dummy variable for the service attribute should be associated with the consumption decision (i.e., parameter a) or the subscription decision (i.e., parameter c), and 0 otherwise. Both vectors are constrained to sum to 1,

$$\text{(8)} \qquad \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = l^a + l^c .$$

We then use these vectors to model the effects of service attributes on the parameters of the utility function:

$$\text{(9)} \qquad \begin{aligned} a_{i,j}^{M} &= a_{i,0}^{M} + \beta_{i}^{M} \cdot \left(x_j \circ l^a\right), \\ b_{i,j}^{M} &= b_{i,0}^{M}, \text{ and} \\ c_{i,j}^{M} &= c_{i,0}^{M} + \beta_{i}^{M} \cdot \left(x_j \circ l^c\right) \end{aligned} \qquad (i \in I,\ j \in J),$$

where $\circ$ indicates the Hadamard Product, that is, the entry-wise multiplication of two vectors. In addition, $a_{i,0}^{M}, b_{i,0}^{M}, \text{and}\ c_{i,0}^{M}$ are the reference levels of the parameters, such that $a_{i,0}^{M}$ describes the individual effect of the increase of utility that accompanies an increase in consumption, $b_{i,0}^{M}$ is the individual decay in the perception as consumption increases, and $c_{i,0}^{M}$ is the individual effect of the general (i.e., usage-independent) perception of the service.

Table 2 illustrates the impacts of these different linkages of service attributes in a numerical example. Case 1 is the base case; Cases 2 and 3 considers an improved service attribute and link the effect of this improvement to either parameter $a^M_{i,j}$ (i.e., consumption decision) or $c^M_{i,j}$ (i.e., subscription decision). The results in Table 2 refer to the corresponding influence on consumer behavior of the two bucket pricing plans (Equation (4)), which we refer to as BP1 and BP2. For this illustration, in BP1 $q_1 = 50$ and $p_1 = \$38.50$, and in BP2 $q_2 = 120$ and $p_2 = \$51.00$. For simplicity, we set the price parameter ϖ_i to 1, so $a^M_{i,j} = a_{i,j}$; $b^M_{i,j} = b_{i,j}$; and $c^M_{i,j} = c_{i,j}$. In the base case, consumption under BP1 is $n_1 = 50$, and that under BP2 is $n_2 = 100$.

TABLE 2: NUMERICAL EXAMPLE OF THE EFFECT OF DIFFERENT LINKS OF SERVICE ATTRIBUTES ON PARAMETERS

	Case 1: No effect		**Case 2: Attribute effects on consumption decision**		**Case 3: Attribute effects on subscription decision**
Parameter a	1.00	→	**1.05**		1.00
Parameter b	0.01		0.01		0.01
Parameter c	1.00		1.00	→	**4.00**
Saturation level and maximum willingness to pay					
Saturation level	100 units		105 units		100 units
MaxWTP	\$51.00		\$56.13		\$54.00
Consumer Surplus					
CS(BP1: 50 units for \$38.50)	\$ 0.00		\$ 2.50		\$ 3.00
CS(BP2: 120 units for \$51.00)	\$ 0.00		\$ 5.13		\$ 3.00
Consumption					
N(BP1: 50 units for \$38.50)	50 units		50 units		50 units
N(BP2: 120 units for \$51.00)	100 units		105 units		100 units
Choice probability					
Prob(BP1: 50 units for \$38.50)	33.33%		6.72%		48.79%
Prob(BP2: 120 units for \$51.00)	33.33%		92.73%		48.79%
Prob(No-choice)	33.33%		0.55%		2.43%

Notes: CS: consumer surplus (see Equation (B.3) in the Appendix B); Saturation level, which is calculated by a/b (see Equation A.3 of the Appendix); MaxWTP: maximum willingness to pay, which is calculated by $a^2/(2\cdot b) + c$ (see Equation B.2 of the Appendix); N: number of units consumed (see Equation (4)); Prob: choice probability (see Equation (10)).

If we link the improvement in the service attribute to parameter a (Case 2), the maximum WTP and the saturation level increase simultaneously. Although consumption under BP1 remains constant in all cases, it increases under BP2 to 105 units. The consumer surplus of BP2 also increases more strongly than in BP1, because of the greater consumption under BP2 (see also Equations (5) and (7)) and the increased utility value of each quantity unit.

In contrast, linking the improved service attribute to parameter c (Case 3) leads to a higher WTP of $3.00 for all units (higher maximum WTP and value of the parameter c, or usage-independent WTP), but it has no influence on the saturation level. The increase of consumer surplus of BP1 ($3.00) is as high as that for BP2. Equations (5) and (7) reveal that the influence on parameter c shifts the utility and WTP functions upward (by $3.00). This shift has no impact on consumption, but it increases proportionally the choice likelihood of each plan and thus the subscription decision.

Overall, linkage through parameter a is appropriate if differences in service attributes affect consumption or the utility provided by consumption. Linkage through parameter c instead is more appropriate if differences in service attributes make the service more attractive but have no effect on consumption.

3.3 Model estimation

For the estimation, we employ hierarchical Bayes, a powerful instrument that delivers parameter estimates for respondents on individual level (Allenby and Rossi 1999). Pricing plans must be aligned carefully with the heterogeneous demand and WTP of consumers if service providers hope to benefit from the potential of second-degree price discrimination. However, observations in discrete choice experiments are scarce, and traditional estimation methods (e.g., maximum likelihood estimator) can barely yield reliable estimates per respondent, which is only possible with advanced statistical techniques.

For this study, respondents make a choice among several offerings with different bucket pricing plans and service attributes, as well as a no-choice option. We assume that consumer i chooses from each choice set s the alternative that yields the highest utility, subject to the budget constraint. To account for any additional factors that influence utility and are known to the consumer but unobservable to the data analyst, we introduce a stochastic component $\varepsilon_{i,j,s}$ (usually labeled the error term) and thus can make statements about the probability $Pr_{i,j,s}$ that consumer i picks plan j in choice set s:

$$(10)\qquad \begin{aligned} Pr_{i,j,s} &= Pr(v_{i,j}(q_j,p_j)+\varepsilon_{i,j,s} > v_{i,j'}(q_{j'},p_{j'})+\varepsilon_{i,j',s}; \forall j \neq j') \\ &= \frac{\exp\left(v_{i,j}(q_j,p_j)\right)}{\exp\left(v_{i,0}\right)+\sum_{j'\in J_s}\exp\left(v_{i,j'}(q_{j'},p_{j'})\right)} \end{aligned} \qquad (i\in I,\ j,j'\in J_s,\ s\in S).$$

The error term $\varepsilon_{i,j,s}$ has a mean of 0 and covariance matrix $\sum$. We assume it is distributed independently and at extreme values (or type I extreme value, Gumbel). Let θ_i summarize all parameters in the utility function of consumer i, that is, $a^M_{i,j}, b^M_{i,j}, \text{and}\ c^M_{i,j}$ and the vector β^M_i, and let $d_{i,j,s}$ represent an indicator variable equal to 1 if consumer i chooses plan j from choice set s, and 0 otherwise. Then matrix θ contains all vectors θ_i, and the following likelihood function of the mixed logit model enables us to estimate the distributions of θ:

$$(11)\qquad L(d\,|\,\theta,\Sigma) = \prod_{i\in I}\prod_{s\in S}\prod_{j\in J_s} Pr_{i,j,s}(d_{i,j,s}=1\,|\,\theta_i,\Sigma) = \prod_{i\in I}\prod_{s\in S}\prod_{j\in J_s}\left(\frac{\exp\left(v_{i,j}(q_j,p_j)\right)}{\exp\left(v_{i,0}\right)+\sum_{j'\in J_s}\exp\left(v_{i,j'}(q_{j'},p_{j'})\right)}\right)^{d_{i,j,s}}.$$

Note that the income term $\varpi_i \cdot Y_i$ is cancelled out in the likelihood function because it has no effect on the differences across the utilities of all alternatives, including the alternative of not using the service at all.

We specify the density of the parameters θ to be normally distributed, with mean $\bar{\theta}$ and covariance matrix Ω, as denoted by $g(\theta_i\,|\,\bar{\theta}, \Omega)$. The conditional posterior on $\theta_i, \forall i$, given $\bar{\theta}$ and Ω, is:

$$(12) \qquad \Lambda(\theta \mid \bar{\theta}, \Omega) \propto \prod_{i \in I} \prod_{s \in S} \prod_{j \in J_s} L(d_{i,j,s} \mid \theta_i, \Sigma) \cdot g(\theta_i \mid \bar{\theta}, \Omega).$$

4 Empirical Study

With our empirical study, we illustrate our model’s ability to reflect consumers' chosen subscriptions, plans, and consumption in a bucket pricing scenario. It also outlines how the linkages of service attributes affect the validity of the results. We use an online survey to study consumers’ preferences for digital music offerings with different bucket pricing plans, which are alligned to the business model, the company eMusic introduced to the European market six months prior to this study.

4.1 Digital music context

New technical possibilities for distributing digital contents and services through the Internet have not been a blessing for all industry sectors. In particular, the music industry has suffered mightily from digital music piracy. After enjoying annual revenue growth rates of 10% or more in the 1990s (Rob and Waldfogel 2006), it experienced a sudden 42% decline in revenues between 2000 and 2008 (from $15 to $8 billion; RIAA 2009). Technologies such as MP3 compression of audio and peer-to-peer platforms enabled consumers to encode purchased CDs and distribute or consume them for free through the Internet. The music industry reacted with both a stick and a carrot (Sinha and Mandel 2008): It discourages illegal download activities through lawsuits but also provides legal alternatives to consume digital music. Such alternatives include services that sell individual songs or bundles on a download-to-own basis (e.g., iTunes) and those that stream music through the Internet, for which customers usually pay a monthly price (e.g., Napster).

These legal alternatives complement traditional CD sales but as yet have not completely compensated for losses (RIAA 2009). They also have triggered a growing debate about the extent to which digital music piracy harms the music industry (Bhattacharjee et al. 2007; Rob and Waldfogel 2006) and the factors that will get people to stop illegally downloading music (Dilmperi, King, and Dennis 2011; Sinha and Mandel 2008). The discussion also features new business models such as advertisement-sponsored streaming services (e.g., the Swedish DRM-based music streaming service Spotify; Papies, Eggers, and Wlömert 2010) and the abandonment of single-song sales in favor of album-only sales (Elberse 2010). To assess new business models, a key metric is the value of digital music to consumers (Papies, Eggers, and Wlömert 2010; Rob and Waldfogel 2006; Sinha and Mandel 2008; Sinha, Machado, and Sellman 2010). However, most prior studies attempt to capture consumers' WTP for just one particular – e.g., a favorite or previously purchased – song or album. Papies, Eggers, and Wlömert (2010) underline the strong need for models that consider changes in consumption by individual consumers to assess price recommendations for new business models in the music industry.

In this context, the business model of eMusic is interesting for two reasons. First, it does not distinguish between songs for which consumers have varying WTP, but treat them as a commodity good, for which consumers have a recurring demand, which can be influenced by pricing. Second, eMusic offers consumers ownership, in the form of DRM-free MP3 downloads, which constituted an innovative service attribute at the time of the study. This strategy conflicts with almost all other services that use DRM restrictions to decrease the incidence of digital piracy (e.g., Sundararajan 2004). However, DRM also restricts consumers' ability to listen to music they have purchased on different technical devices like MP3-player, car stereo, or various PCs. Thus 2007, a month after we completed our study, Apple's chief executive officer Steve Jobs called for the elimination of DRM from legally sold digital songs and introduced DRM-free songs on iTunes for a surchage of 0.30€ per song (i.e., 0.99€ for DRM-restricted and 1.29€ for DRM-free songs).

4.2 Study design

Questionnaire. The questionnaire consists of three parts: self-stated preferences and experiences with legally purchasing digital music, choice sets that ask respondents to select the most attractive alternative, and demographic information items. The bucket pricing plans in the choice sets differ in their DRM restrictions (DRM-free and DRM-restricted),[2] brand names (Musicload, Napster, iTunes, and a fictive name, Loadasong), monthly subscription prices (4.99€, 7.99€, 14.99€, 24.99€, and 34.99€), and allowance levels (5, 10, 20, 40, and 60 songs). Consistent with the eMusic business model, we informed respondents that they could switch their plan once a month. Even though illegal downloading is pervasive for the music industry, we did not explicitly considered it in the choice of attributes and levels (e.g., by offering an illegal option at the price of 0 €). The reasons are two-fold: First, the resulting decision model would also have to account for respondents' perceived probability of being exposed to the risks of a lawsuit and possibly moral consequences. Therefore, illegally downloaded songs may also provide some negative utility, which is difficult to elicit with discrete choice experiments. Second, including an option, which depict an act against the law might cause social desirability bias, which would influence the parameter estimates.

Choice set design. We created two D-optimal versions of the choice designs, each consisting of 15 choice sets, and added two additional choice sets to each version for holdout

2 We informed respondents that they could burn DRM-restricted songs up to five times on a CD and were unrestricted in copying them to a mobile player or playing them on a single PC.

predictions (see the Appendix). Each choice set consists of three bucket pricing plans and a no-choice option. Because illegal downloading has been such a big challenge for the music industry (Rob and Waldfogel 2006), we incorporated music piracy implicitely in the no-choice option. That is, if a respondent prefers illegal sources for digital music, he or she should select the no-choice option, because none of the pricing provides sufficient utility.

Data collection. We conducted an online survey among undergraduate and graduate students of a major European university and received 123 completed questionnaires. We considered students suitable respondents for this study because they represent one of the main target groups for music download providers. Of the respondents, 26.83% stated that they had legally purchased digital music,[3] mainly using iTunes, Musicload, or Napster (64.71%, 41.18%, and 26.47%, respectively; multiple responses were possible).

Estimation. We use the 1845 discrete choice decisions (= 123 × 15) that constitute the data set for the estimation and use Equation (7) as the indirect utility function of choice in the estimation. Our model links DRM restrictions to the parameter $a_{i,j}^{M}$ and the brand name to the usage-independent parameter $c_{i,j}^{M}$ (Model 1). Therefore, Model 1 reflects our expectation that DRM restrictions affect plan choice though the consumption decision, whereas the name of the platform is connected to brand awareness (Agarwal and Rao 1996) and influences plan choice through the subscription decision. In addition, we estimate three possible combinations of attribute linkages to either $a_{i,j}^{M}$ or $c_{i,j}^{M}$ (Models 2–4). For example, linking the brand to parameter $a_{i,j}^{M}$ indicates that the brand influences consumption, perhaps because

3 We also tested a model, where consumers' experience with legal music download platforms entered as an observable covariate for explanation of the parameters, but find no improvement in model fit (see the Appendix).

brands indicate differences in platform usability or the availability of a recommender system (Senecal and Nantel 2004). The model with the best fit should reveal the best match of attribute linkages to respondents' decision-making.

The base levels of the parameters $a^{M}_{i,0}, b^{M}_{i,0}$, $c^{M}_{i,0}$, and ϖ_{i}, as well as the parameter for the DRM restrictions, are reparametrizied (e.g., $a^{M}_{i,0} = \exp(a^{M}_{i,0}{}')$) to ensure positive values and the desired concave form of the utility function. Dummy variables for the service attributes in the matrix x_j are effect coded. Estimations rely on hierarchical Bayes techniques and use standard diffuse priors. The reported results come from 20,000 iterations that we retain after discarding the initial 40,000 iterations ($\hat{=}$ 60,000 iterations total). We assess convergence according to the trace plot of the likelihood and parameters.

4.3 Results

We evaluate the validity of the four models by comparing commonly applied measures. Using the harmonic mean estimator (Newton and Raftery 1994), we compute the log marginal density (Iyengar, Jedidi, and Kohli 2008; Newton and Raftery 1994; Sonnier, Ainslie, and Otter 2007). We also calculate the internal hit rates (HR) and mean absolute deviations (MAD; Brazell et al. 2006) for the 15 choice sets, as well as the predictive HR and MAD in the two holdouts, as we show in Table 3.

All models predict consumers’ choices well, but Model 1 performs best. It has the highest internal log marginal density and HR, as well as the lowest MAD. The differences in the log marginal density are larger than the critical value of 10 (Kass and Raftery 1995) and indicate the strong superiority of Model 1. Thus brand names increase the likelihood of service subscription, but DRM restrictions influence consumption.

TABLE 3: INTERNAL AND PREDICTIVE VALIDITY

	Internal Validity			Predictive Validity	
	LMD	HR	MAD	HR	MAD
Model 1: DRM restrictions linked to $a^M_{i,j}$, brand name to $c^M_{i,j}$	**-571.99**	**89.70%**	**0.13**	**81.30%**	**0.21**
Model 2: All attributes linked to $a^M_{i,j}$	-626.86	86.61%	0.16	73.58%	0.29
Model 3: All attributes linked to $c^M_{i,j}$	-660.01	85.92%	0.14	78.46%	0.29
Model 4: DRM restrictions linked to $c^M_{i,j}$, brand name to $a^M_{i,j}$	-710.25	83.09%	0.19	70.33%	0.31

Notes: LMD: log marginal density; HR: hit rate; MAD: mean absolute deviations; DRM: digital rights management

In Table 4, we report the posterior mean, median, and diagonal of the covariance matrix of the best fitting model. Posterior standard deviations appear in parentheses. The mean value of the parameter $a^M_{i,0}$, which drives the increase of the utility function, is 0.61. As we expected, this value is higher than that of the parameter $b^M_{i,0}$, which equals 0.18 and is responsible for the decrease in marginal utility. The mean value of the parameter $c^M_{i,0}$ describes the usage-independent WTP and is considerably lower, with a value of 0.03. According to Table 4, consumers are willing to pay an average of 0.40€ (median 0.18€) more for each DRM-free song. This result has high face validity and matches iTunes' price premium of 0.30€. Nearly all parameters display the expected signs and are of reasonable size. The estimates for brand names are small, which indicates that they play a negligible role in consumers' decision-making. Perhaps respondents consider songs commodity goods and perceive only minor differences in the quality or comfort provided by different music download platforms.

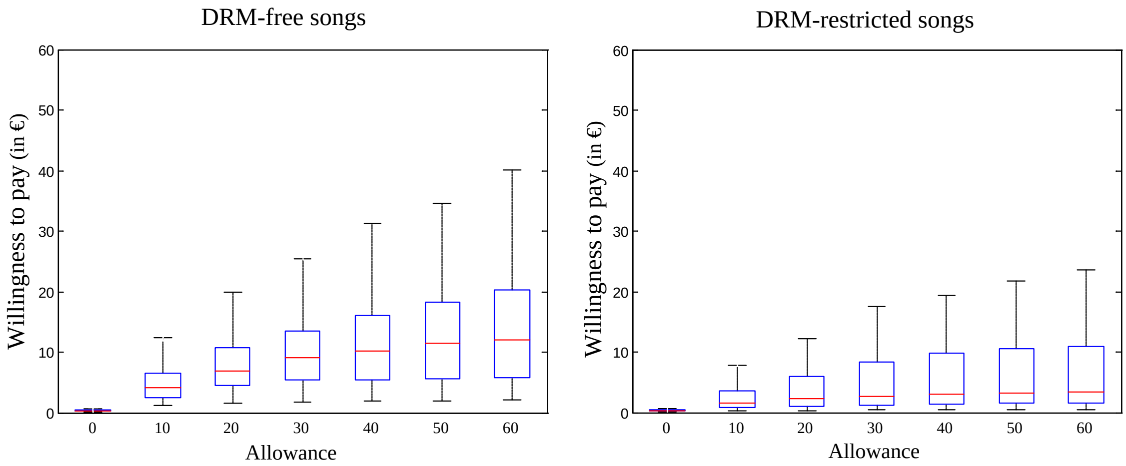
TABLE 4: PARAMETER ESTIMATES

Attribute	Level	Mean	Median	Heterogeneity
Parameters of willingness-to-pay function	$a^M_{i,0}$	0.61	0.41	0.59
		(0.07)	(0.04)	(0.27)
	$b^M_{i,0}$	0.18	0.03	0.30
		(0.25)	(0.00)	(1.93)
	$c^M_{i,0}$	0.03	0.02	0.08
		(0.04)	(0.00)	(0.26)
Price parameter	ϖ_i	1.57	1.81	0.74
		(0.10)	(0.09)	(0.11)
Brand name	Musicload	0.23	0.19	0.37
		(0.22)	(0.23)	(0.20)
	Napster	0.41	0.34	0.36
		(0.21)	(0.22)	(0.18)
	iTunes	-0.25	-0.21	0.42
		(0.27)	(0.28)	(0.19)
	Base: Loadasong	-0.38	-0.43	0.48
		(0.31)	(0.32)	(0.21)
DRM restrictions	DRM-free	0.20	0.09	0.31
		(0.04)	(0.01)	(0.28)
	Base: DRM-restricted	-0.20	-0.09	0.31
		(0.04)	(0.01)	(0.28)

We calculate WTP for various allowances and show its median and 5%, 25%, 75%, and 95% quantiles in the boxplots in Figure 2. The left- and right-hand sides present the willingness to pay estimates for DRM-free and DRM-restricted songs, respectively. We used iTunes as the brand name, but the differences are similar for the other brand names.

On average, respondents are willing to pay 4.94€ for 10 DRM-free songs but 46.03% less (i.e., 2.67€) if they are DRM-restricted. For 60 songs, the WTP difference increases to 49.54% (7.59€ = 15.31€ for DRM-free – 7.72€ for DRM-restricted songs). Furthermore, the median WTP continues to increase for 50 to 60 DRM-free songs; for 30 songs or more, it stays nearly constant with that for DRM-restricted songs. This finding supports the claim that offering DRM-free songs increases consumption.

FIGURE 2: WILLINGNESS TO PAY FOR DRM-FREE AND DRM-RESTRICTED SONGS



5 Counterfactual Simulations

Our discrete choice model attempts to capture decision-making processes for bucket pricing plans and emphasizes the importance of choosing the correct linkage of service attributes to the parameters of the utility function. However, service providers are more interested in recommendations of optimal (profit-maximizing) prices and allowances in bucket pricing plans (e.g., Danaher 2002; Iyengar, Jedidi, and Kohli 2008). Therefore, in a counterfactual simulation, we first analyze the extent to which price recommendations vary across service attribute linkages. Considering the increasing popularity of bucket pricing, in a second counterfactual simulation we tackle the question of how their profitabilty compares with that of other pricing plans (e.g., pay-per-use, two-part pricing, and three-part pricing).

5.1 Comparison of pricing recommendations across different service attribute linkages

To quantify the extent to which managerial implications vary when different types of behavioral processes are associated with each service attribute, we compare the recommendations of optimal (profit-maximizing) bucket pricing plans, obtained from the four models in our empirical study. The individual draws of the Metropolis Hasting step in the hierarchical Bayesian estimation indicate respondents' simultaneous decisions about subscriptions, plan choices, and consumption; these predictions then provide a means to estimate the profits of various menus of bucket pricing plans denoted by J:[4]

$$(13) \qquad \pi\left(p_J, q_J\right) = \sum_{j \in J} \sum_{i \in I} \left[\left(p_j - k_v \cdot n^*_{i,j}(q_j)\right)\right] \cdot \Pr_{i,j}\left(p_j, q_j\right) \rightarrow \max .$$

Profit equals the sum of the profit contributions of all consumers, which consists of two components: the probability of a consumer choosing plan j, and the margin, or the monthly price p_j minus the product of variable costs k_v and consumption $n^*_{i,j}(q_j)$, as specified in Equation (4). Although not considered here, it is straightforward to incorporate other types of costs, such as fixed costs for each subscriber served.

From the MCMC sampler, we consider a subsample of the posterior distributions of the individual parameters to determine optimal pricing recommendations. We use a subsample rather than posterior means to describe respondents according to the distribution of their individual parameters, which can be asymmetric. This subsample helps us calculate profits for a given set of bucket pricing plans J. The challenge then is to determine the optimal bucket pricing plans that maximize profit, because the objective function in Equation (13) is nonlinear and nonconvex with many local maxima. We implemented and extensively tested various heuristic search methods; simulated annealing provides the best search performance (see also Schlereth, Stepanchuk, and Skiera 2010). It randomly accepts solutions with a decreasing objective functional value and can thus overcome local maxima. To our

4 We depict the full model of the profit maximization problem in the Appendix.

knowledge, ours is the first study to determine optimal bucket and three-part pricing plans while investigating more than just one or two options.

We apply simulated annealing to menus of one, two, and three bucket pricing plans, using the individual parameter distributions of each model, which results in 12 optimization runs (= 3 menus × 4 models). We neglect competition to illustrate more clearly the effects of linking attributes to either consumption or subscription decisions. It is straightforward to account for static competition[5] by including the pricing schemes of the most important competitors. We use the fictive brand name “Loadasong”; this scenario reflects the situation when eMusic entered the European market and was not as popular or familiar as iTunes, Musicload, or Napster. We assume that variable costs, such as those for licensing, taxes, and technical infrastructure, are 0.22€, which account for approximately 80% of the price per song in eMusic’s largest plan. We use the parameter estimates of the best fitting Model 1 to simulate the effects of the optimal bucket pricing plans on profits across all models. The differences in Table 5, compared against the optimal profit in Model 1, indicate the degree to which the recommendations and strategies for segmenting the market vary, if one uses the wrong model to obtain price recommendations.

According to Table 5, applying bucket pricing recommendations from Models 2–4 decreases profits for the three bucket pricing plans from 8.96% to 22.75% (see last column). The recommendations of Model 2 yield the smallest deviations, because the DRM restrictions influence the parameter $a_{i,j}^{M}$. The greatest deviation appears in Model 4, which links DRM

5 The price histories of eMusic and market leaders like iTunes indicate no observable dependencies in how prices are set. For other service fields, it might be possible to include Bertrand competition or Stackelberg leader–follower price competition (Kadiyali, Sudhir, and Rao 2001) in the simulation.

restrictions on parameter $c_{i,j}^{M}$ and brand on $a_{i,j}^{M}$ (i.e., the reverse of Model 1). With fewer plans, this difference increases up to 30.76%.

TABLE 5: DEVIATION IN PROFIT AND OPTIMAL BUCKET PRICING PLANS

Profit	**One Bucket Pricing Plan**		**Two Bucket Pricing Plans**		**Three Bucket Pricing Plans**	
Model 1: DRM restrictions linked to $a_{i,j}^{M}$, brand name to $c_{i,j}^{M}$	3,764.69 €		5,088.50 €		5,490.75 €	
Model 2: All attributes linked to $a_{i,j}^{M}$	-5.29%		-1.53%		-8.96%	
Model 3: All attributes linked to $c_{i,j}^{M}$	-21.61%		-18.43%		-13.97%	
Model 4: DRM restrictions linked to $c_{i,j}^{M}$, brand nameto $a_{i,j}^{M}$	-7.45%		-30.76%		-22.75%	
Optimal Prices	**Monthly Price**	**Allowance**	**Monthly Price**	**Allowance**	**Monthly Price**	**Allowance**
Model 1: DRM restrictions linked to $a_{i,j}^{M}$, brand name to $c_{i,j}^{M}$	18.57 €	48	13.25 €	26	6.00 €	8
			49.39 €	80	15.28 €	31
					47.82 €	80
Model 2: All attributes linked to $a_{i,j}^{M}$	14.74 €	30	12.66 €	24	4.32 €	3
			43.75 €	80	12.33 €	23
					43.98 €	80
Model 3: All attributes linked to $c_{i,j}^{M}$	8.77 €	20	6.74 €	11	5.15 €	5
			59.70 €	80	9.91 €	23
					53.52 €	80
Model 4: DRM restrictions linked to $c_{i,j}^{M}$, brand nameto $a_{i,j}^{M}$	12.64 €	27	5.72 €	10	4.94 €	2
			72.41 €	80	12.35 €	26
					73.45 €	80

The recommended prices and allowances in the bucket pricing plans thus indicate differences for segmenting the market. Model 1 suggests three bucket pricing plans with allowances of 8, 31, and 80 songs.[6] The prices per song then range between 0.75€ (= 6.00€/8) and 0.60€ (= 47.82€/80), lower than the prices currently charged by iTunes or Musicload (i.e., 0.99–1.39€ per song). Moreover, eMusic's monthly bucket pricing plans (i.e., as of 2007 in Europe, 30

6 We set the upper bound of the allowance to 80 songs and thus account for the range of allowances (i.e., 5 – 60 songs) in the empirical study.

songs for 12.99€, 50 for 16.99€, or 75 for 20.99€) provide songs at substantially lower prices than its competitors. Thus, the results have high face validity.

In contrast with this pricing plan, the results of Model 1 suggest targeting consumers with demand of less than 30 songs and charging higher prices per song for consumers with higher demand quantities. Both suggestions predict actual changes in eMusic's subscription plan in the years 2008–2010; it reduced the number of songs in its lowest plan and increased the average price per song in all bucket pricing plans (i.e., in 2010, 24 songs for 11.99€ per month, 35 for 15.99€, and 50 for 20.79€).

The model that deviates most in terms of profit (i.e., Model 4) recommends skimming usage-independent WTP, as reflected by the parameter $c_{i,0}^{M}$. Therefore, for the three bucket pricing plans, the model recommends offering very few songs for a high average price (i.e., 4.94€ for two songs, or 2.47 € per song). In addition, the price for 80 songs is unreasonably high and exceeds 73€, or 0.92 € per song. Thus, we conclude a lack of face validity for the recommendations of Model 4.

If the menu consists of just one bucket pricing plan, Models 1 and 2 recommend offering 30–48 songs for about 14.00–18.57€, whereas Models 3 and 4 recommend fewer songs for a lower price. Both Models 3 and 4 include DRM restrictions in the usage-independent WTP, so their price recommendations target consumers with lower demand. Similarly, when offering two bucket pricing plans, the recommendations in Models 3 and 4 differ from those in Models 1 and 2, because they emphasize the outer limits of the price range. Therefore, Models 3 and 4 recommend one plan with few songs (about 10) and one plan with 80 songs for a high price (i.e., at least 59.70€). In contrast, the models that link the service characteristic to the parameter $a^{M}_{i,j}$ suggest targeting consumers whose demand is approximately 25 songs and asking a lower price for 80 songs.

5.2 Comparison with other pricing plans

Bucket pricing plans differ from pay-per-use, two-part, and three-part pricing plans in that they do not contain a marginal price (see also Figure 1). A two-part pricing plan for music downloads is a plan, which charges a monthly fixed fee plus a marginal price per song, which is below the marginal price per song of common pay-per-use operators like iTunes. A three-part pricing plan offers an allowance in exchange for the fixed fee, however differs from bucket pricing plans in that consumers can purchase single songs for a certain marginal price if they exceed that allowance. Thus, the bucket pricing plans are less flexible, but they encourage consumers to use a preset number of units.

In a second counterfactual simulation, we therefore analyze the extent to which profits differ across the alternative plans. Furthermore, service providers frequently must choose a particular number of plans, which involves a trade-off between offering more plans to attain better market segmentation (Maskin and Riley 1984; Murphy 1977) and offering fewer plans to avoid confusing the consumers and minimize administrative costs for managing the plans (Hui, Yoo, and Tam 2007). Extant recommendations only apply to two-part pricing plans (e.g., Murphy 1977; Schlereth, Stepanchuk, and Skiera 2010).

Using the subsample from the posterior distribution of Model 1, we determine profit-maximizing plans, analogous to the first counterfactual simulation. We vary the type and number of pricing plan, such that they range between one and four, except for the pay-per-use plan, which can feature only one. We also vary the variable costs (i.e., 0.02–0.42€ per song, in two increments of 0.20€), because previous research shows strong dependence for some pricing plans (i.e., flat rates) between variable costs and profitability (e.g., Schlereth, Stepanchuk, and Skiera 2010).[7] For each of the 39 conditions (= 3 types of nonlinear pricing plans × 4 plans × 3 variable costs + 3 pay-per-use plans), we determine the optimal pricing plans and report the results in Table 6.

The second counterfactual simulation requires some important assumptions. First, we assume that the individual parameters $a_{i,0}^{M}, b_{i,0}^{M}, \text{and } c_{i,0}^{M}$ remain the same for different types of pricing plans and neglect potential pricing plan effects (e.g., Lambrecht and Skiera 2006). Second, we ignore the length of the commitment to a pricing plan, which tends to be longer for two- and three-part than for bucket pricing plans. Third, we neglect the strategic role of the type of pricing plans in oligopolies. Yang and Ye (2008) argue that instead of just offering cheaper prices, service providers should employ pricing plans strategically to differentiate themselves from competitors. We would welcome research that eliminates at least some of these assumptions.

In Table 6, we provide the optimal profits relative to the profit of four optimal bucket pricing plans. The latter are always lower than the profits from the four two-part or three-part pricing plans. However, the differences range between 0.65% and 2.86%, which we consider modest. Variable costs have no observable impact on differences in profit.

---

7 We do not incorporate a flat-rate plan because our range of allowances with up to 60 songs is too narrow to study it as well.

TABLE 6: COMPARISON OF OPTIMAL PROFITS OF ALTERNATIVE PRICING PLANS, RELATIVE TO FOUR BUCKET PRICING PLANS

Variable Costs	Pricing Plan	1 Plan	2 Plans	3 Plans	4 Plans
Low (i.e., 0.02 €)	**Bucket Pricing**	90.42%	96.34%	99.21%	100.00%
	Pay-Per-Use Plan	86.61%		-	
	Two-Part Pricing	98.32%	99.94%	100.51%	100.65%
	Three-Part Pricing	98.35%	100.33%	100.62%	100.89%
Medium (i.e., 0.22 €)	**Bucket Pricing**	66.58%	89.97%	97.14%	100.00%
	Pay-Per-Use Plan	89.11%			
	Two-Part Pricing	98.91%	100.60%	100.60%	100.95%
	Three-Part Pricing	98.91%	100.66%	100.66%	100.95%
High (i.e., 0.42 €)	**Bucket Pricing**	55.15%	88.55%	95.70%	100.00%
	Pay-Per-Use Plan	86.57%			
	Two-Part Pricing	100.87%	101.28%	101.62%	102.76%
	Three-Part Pricing	100.87%	101.88%	101.64%	102.86%
Average (over the three variable costs)	**Bucket Pricing**	70.72%	91.62%	97.35%	100.00%
	Pay-Per-Use Plan	87.43%			
	Two-Part Pricing	99.37%	100.61%	100.91%	101.46%
	Three-Part Pricing	99.38%	100.96%	100.97%	101.37%

Bucket pricing plans substantially outperform pay-per-use plans by at least 10.89% (i.e., = 100% – 89.11%). Unexpectedly, we find only marginal differences in profit between two-part and three-part pricing plans. The largest difference occurs for high variable costs and only is up to 2.86% (102.86% – 100.00%), which contradicts the common belief that a more flexible pricing plan (i.e., the possibility in three-part pricing plans to purchase single songs, after exceeding the allowances) generates additional and substantial profits. Bucket pricing plans seem to have the same ability to differentiate consumers' WTP and consumption-pattern like two- and three-part pricing plans. We note that reasons for this surprising result are not completely clear to us.

The choice of a specific type of pricing plan has important implications for the service provider's decision about how many plans to employ. In the case of two-part pricing plans, the results in Table 6 confirm previous findings; two optional two-part pricing plans are sufficient to realize most of the profit. We find similar results for three-part pricing plans.

However, for bucket pricing plans, offering two, three, or four plans, instead of just one, increases in profit by 20.90% (i.e., 91.62% – 70.72%, see Table 6), 26.63%, and 29.28%, respectively (values are averaged over the three variable costs). Thus service providers benefit even more from the introduction of additional bucket pricing plans, assuming variable costs are high.

The finding that a service provider is best advised to use a high number of bucket pricing plans might actually provide an explanation for their surprising profitability. Consumers, who choose the same plan, have an incentive to consume the whole allowance or to stop the usage if their saturation level is lower than the allowance (which in case of a low allowance applies to only few consumers). Therefore, differentiation occurs mainly through the optional plans among which consumers can choose. In contrast for two-part pricing plans, consumers already differentiate themselves according to their usage, because consumption depends on their heterogeneous preferences, which might differ among consumers (see Schlereth, Stepanchuk, and Skiera 2010). While increasing the number of two-part pricing plans enables service providers to better differentiate among consumers, because of the already differentiated market, when compared to bucket pricing plans, the profit potential of both pricing plans eventually seems to be about the same. For three-part pricing plans, we expect the explanation to be somewhere in-between, given that three part pricing plans consist of a combination of allowance and marginal price after exceeding the allowance.

In summary, service providers can freely choose among multiple two-part, three-part, and bucket pricing plans, because their differences in profit are small. They just need to realize that they likely will need to offer more bucket pricing plans, especially if variable costs are high. Therefore, digital music providers such as eMusic could benefit from offering even more bucket pricing plans than the three it offered during 2007–2010.

6 Conclusions

Learning about the heterogeneous demand of consumers and using that information to define optimal pricing plans is an important challenge. Modern companies increasingly adopt bucket pricing, an alternative to flat-rate, pay-per-use, two-part, and three-part pricing plans. So in this work, we develop a discrete choice model of consumers' preferences for bucket pricing plans that accounts for the interdependencies between consumers' decisions about their subscription, plan, and consumption. These preferences in turn provide a means to derive bucket pricing recommendations that differ across consumers.

We propose a model to capture consumers’ decisions and provide a flexible extension of our model that allows for different considerations of service attributes in the utility function. For example, eMusic traditionally has offered DRM-free songs and thereby built a unique brand image. These attributes potentially affect consumers' subscriptions, plan choices, and consumption decisions, but the specific nature of the effects depends strictly on the attributes. An analyst might have an intuitive understanding of the influence of some attributes, but without utter certainty about the most appropriate linkage that reflects consumers’ common decision-making processes, analysts must undertake a comparison of the fit of alternative modeling approaches. In the case of digital music for example, DRM restrictions influence consumption decisions, but the choice of service provider only influences the decision about service subscriptions. Our results also show that a well-specified linkage of attributes has a strong impact on bucket pricing recommendations, so failures to account for such linkages correctly can cause possible profit losses of up to 22.75%.

This study has also several important implications for marketers. We show that profits under bucket pricing plans are almost the same as those under two-part and three-part pricing plans but substantially higher than those under pay-per-use plans. Unlike two-part and three-part pricing plans, it is beneficial to increase the number of bucket pricing plans, which substantially increases profit (in our study, by on average up to 29.28%). The greater number of bucket pricing plans enables service providers to differentiate better among heterogeneous consumer demands, because the charges for consumption reflect specific allowances. We therefore conclude that bucket pricing plans present an attractive alternative for service providers.

Our study also has some limitations that warrant attention. For example, we did not incorporate a potential learning effect of consumers over time, despite the likely influence on consumers' plan choices and consumption (e.g., Iyengar, Ansari, and Gupta 2007; Narayanan, Chintagunta, and Miravete 2007). However, when using discrete choice experiments, these effects are not observable, because respondents do not receive feedback about their decisions and therefore cannot mentally process the outcome of previous decision-making for future decisions. Further research therefore might combine discrete choice experiment data and transaction data to obtain the advantages of both data sources (e.g., Swait and Andrews 2003). We consider the application of this approach to nonlinear pricing valuable, because such a model could capture such (e.g., learning) effects and strategically control for different pricing offers, which are not observed in real markets. Another valuable extension of our model is that of quality differentiated subscription plans. We acknowledge that service providers can differentiate themselves from their competitors not only through their prices but also through their choices of service attributes. For example, the Swedish streaming service Spotify offers a free streaming service, which has restrictions such as a lower music quality and radio-like advertisements, as well as a premium subscription for a monthly fee with no restrictions. How to link the preferences to differentiated pricing plans is another research topic of high practical importance.

Acknowledgments

The authors gratefully acknowledge the help of Johanna Burkhardt during the data collection and excellent suggestions of Thomas Otter. In addition, Editor Marnik Dekimpe, the area editor, and the two anonymous referees provided many very helpful and excellent suggestions. Large parts of this manuscript were written during stays of the authors at the Centre for the Study of Choice (CenSoC) at the University of Technology in Sydney (Australia), for which we thank Jordan Louviere for his support and valuable comments. This research was funded by the German Ministry for Education and Research (BMBF – FKZ: 01 IA 08003C), by the German Research Foundation (DFG – GZ: SCHL 1942/1-1), and by the E‐Finance Lab Frankfurt.

APPENDIX

A Proof of optimal consumption

We derive optimal consumption $n_{i,j}^{*}$ by forming a Lagrange function for the problem outlined in Equation (3), subject to $0 \le n_{i,j}(q_j) \le q_j$:

$$L\left(n_{i,j}\left(q_j\right),\lambda\right)=v_{i,j}\left(n_{i,j}\left(q_j\right),p_j\right)-\lambda\cdot\left(q_j-n_{i,j}\left(q_j\right)\right)\rightarrow \max \qquad (i\in I,\ j\in J). \tag{A.1}$$

Differentiating the Lagrange function in Equation (A.1) yields:

$$\frac{\partial L\left(n_{i,j}\left(q_j\right),\lambda\right)}{\partial n_{i,j}\left(q_j\right)}=a_{i,j}-b_{i,j}\cdot n_{i,j}\left(q_j\right)+\lambda \qquad (i\in I,\ j\in J). \tag{A.2}$$

We obtain two optimal solutions, based on the Kuhn Tucker conditions $\frac{\partial L\left(n_{i,j}\left(q_j\right),\lambda\right)}{\partial n_{i,j}\left(q_j\right)}=0$,

$\lambda \geq 0$, and $\lambda \cdot \left(q_j - n_{i,j}\left(q_j\right)\right)=0$. A consumer uses the service up to the saturation level if this saturation level is smaller than the allowance; otherwise, the consumer completely exhausts the available allowance:

$$\text{(A.3)} \qquad n^*_{i,j}\left(q_j\right)=\begin{cases} q_j & ,\text{if } q_j \leq \frac{a_{i,j}}{b_{i,j}} \\ \frac{a_{i,j}}{b_{i,j}} & ,\text{if } q_j > \frac{a_{i,j}}{b_{i,j}} \end{cases} \qquad (i\in I,\ j\in J).$$

B *Estimation space transformation*

We transform the indirect utility function in Equation (5) into the surplus function. Thus, we move the preference space, also known as the utility model, to the WTP space, which is also known as the surplus model (Sonnier, Ainslie, and Otter 2007). For this purpose, we define $WTP_{i,j}(q_j)$ as the price for an allowance q_j, for which the consumer is indifferent between purchasing and not purchasing ($v_{i,0}=v_{i,j}[q_j,WTP_{i,j}(q_j)]$; e.g., Moorthy, Ratchford, and Talukdar 1997). Thus, we can write:

$$\text{(B.1)} \qquad \begin{aligned} \varpi_i \cdot Y_i = {} & Ind_{i,j} \cdot \left(a_{i,j} \cdot q_j - \frac{b_{i,j}}{2} \cdot q_j^2 \right) + \\ & (1-Ind_{i,j}) \cdot \left(\frac{a_{i,j}^2}{2 \cdot b_{i,j}} \right) + c_{i,j} + \varpi_i \cdot (Y_i - WTP_{i,j}(q_j)) \end{aligned} \qquad (i\in I,\ j\in J).$$

Transforming Equation (B.1) so that $WTP_{i,j}(q_j)$ appears on the right-hand side yields:

$$\text{(B.2)} \qquad WTP_{i,j}(q_j)=\frac{1}{\varpi_i}\left(Ind_{i,j} \cdot \left(a_{i,j} \cdot q_j - \frac{b_{i,j}}{2} \cdot q_j^2 \right) + (1-Ind_{i,j}) \cdot \left(\frac{a_{i,j}^2}{2 \cdot b_{i,j}} \right) + c_{i,j} \right) \qquad (i\in I,\ j\in J),$$

and the consumer surplus, given the monthly package price p_j then becomes:

$$\text{(B.3)} \qquad CS_{i,j}(q_j, p_j) = WTP_{i,j}(q_j) - p_j \qquad (i \in I,\ j \in J).$$

If we compare Equations (5) and (B.2), we find that the relationship between the parameters that express preferences and the monetary parameters (indicated by the subscript M) yields statements about WTP:

$$\text{(B.4)} \qquad a^M_{i,j} = \frac{a_{i,j}}{\varpi_i};\ b^M_{i,j} = \frac{b_{i,j}}{\varpi_i};\ \ c^M_{i,j} = \frac{c_{i,j}}{\varpi_i} \qquad (i \in I,\ j \in J),$$

which enables us to transform Equation (5) into Equation (7).

Equations (5) and (7) are behaviorally equivalent, and their usage theoretically leads to the same parameter estimates. This equivalence occurs if we use estimation methods that do not incorporate prior information, such as the maximum likelihood estimator. However, when incorporating prior information, as required by hierarchical Bayes methods, the distributions of these priors typically are not equivalently included in both estimation spaces. In particular, the commonly applied standard diffuse priors of monetary parameters derived from the utility model are normal and usually divided by normal or lognormal distributions. These priors may emphasize the tails of the distribution and place greater weight on outliers, which may result in unreasonably high parameter estimates. Sonnier, Ainslie, and Otter (2007) thus argue that in the surplus model, the prior distribution for WTP expressions is just normal or lognormal. This model then yields more face-valid estimates than the utility model, especially if data are scarce, as is the case in most discrete choice experiments. We account for their considerations by using Equation (7) as an indirect utility function of choice in our empirical estimation.

C Choice design

In the empirical study, two designs each consisted of 15 choice sets for the estimation and 2 additional holdouts, which were the same for both designs. Thus, respondents answered 17 choice sets in total.[8] With the exception of the brand attribute, the attributes in the discrete choice experiment contain levels with associated benefits or costs that monotonically increase or decrease. Therefore, it is possible for choice sets to include bucket pricing plans that strongly dominate other bucket pricing plans in the same choice set. A dominant plan has more of every benefit attribute (allowance and DRM restrictions) and less of the monthly package price attribute than any other plan. When such plans are present, the decision is trivial and thus such a choice set provides limited information.

There are several possibilities for creating the choice design, including the methods suggested by Street, Burgess, and Louviere (2005) or the use of software (e.g., NGene or Sawtooth) to maximize D-efficiency. To make the choices more realistic, we excluded two unlikely bucket pricing plans (i.e., 5 songs for 34.99€ and 60 songs for 3.99€) and generated a D-efficient starting design. We avoid the dominance of any alternative by creating a set of additional 80 candidates. Next, we employed the analog to Strategy 5 from Street, Burgess, and Louviere (2005), which is an iterative approach, and carefully tested each choice set for dominating or dominated occurrences. We replaced each such occurrence with five candidates and selected the one that maximized the resulting D-optimal design efficiency. The final design appears in Table A1. Column 1 features the allowance level, column 2 the

8 The topic of how many choice sets to use per person prompts little consensus. On the one hand, researchers want to collect as much information as possible, but on the other hand, they should avoid straining respondents' cognitive effort to prevent answers with little value or subject to high error variance. Marketing research frequently employs 16–32 choice sets per respondent (e.g., Hensher, Stopher, and Louviere 2001; Iyengar, Jedidi, and Kohli 2008; Parker and Schrift 2011), and Hensher, Stopher, and Louviere (2001) recommend around 16 treatments.

price, column 3 the brand, and column 4 indicates DRM restrictions. Holdouts are indicated by “H” in the left-side column. Using online software provided by Street and Burgess (http://crsu.science.uts.edu.au/choice/choice.html), we find that each of the designs obtains an efficiency of at least 76% compared with the optimal design (which does not account for dominant alternatives) or 86% when combined.

TABLE C.1: CHOICE DESIGN

	Version 1												Version 2											
	Pricing Plan 1				Pricing Plan 2				Pricing Plan 3				Pricing Plan 1				Pricing Plan 2				Pricing Plan 3			
	3	2	3	1	2	4	0	0	0	0	2	0	3	1	1	0	0	0	0	1	4	4	2	1
	1	0	1	1	4	2	2	1	2	3	3	0	1	0	3	1	4	2	1	1	2	4	2	0
	0	0	3	0	1	3	1	1	3	4	2	1	1	0	3	0	3	3	1	1	2	0	0	1
	3	4	2	1	1	3	1	0	0	0	0	0	3	3	0	0	0	2	1	0	4	4	2	1
	0	1	0	0	4	2	3	0	0	0	2	1	3	4	3	0	4	3	2	1	2	0	1	0
	4	4	0	0	3	3	2	1	0	1	1	1	0	0	2	0	2	3	1	1	1	0	3	1
	2	1	0	1	2	4	3	0	1	0	1	0	0	2	0	0	1	1	2	1	3	3	1	0
	3	4	0	1	2	1	2	0	0	0	1	1	0	0	2	1	4	3	0	1	0	2	3	0
H	4	4	0	0	2	0	1	0	3	1	3	1	4	4	0	0	2	0	1	0	3	1	3	1
	1	1	2	0	3	3	0	1	0	0	1	0	0	1	2	1	1	2	3	1	4	3	0	0
	2	0	1	1	1	2	0	0	3	4	3	0	0	0	3	1	3	1	2	0	4	4	1	1
H	3	4	1	0	1	3	2	1	0	1	0	0	3	4	1	0	1	3	2	1	0	1	0	0
	4	4	3	1	2	2	1	0	1	0	2	1	0	0	3	0	1	1	1	1	4	2	2	0
	3	1	3	1	0	0	2	0	4	2	0	1	3	3	2	0	4	2	3	1	0	1	1	0
	2	3	2	1	3	4	3	1	1	2	1	0	2	1	1	1	1	0	0	1	0	2	3	0
	2	4	3	0	0	2	2	0	4	3	1	1	2	2	1	1	1	0	3	0	4	4	2	0
	0	1	3	1	2	2	1	0	4	3	0	1	4	3	0	0	1	2	2	0	3	1	1	1

Holdouts are indicated by “H” in the left-side column.

D Influence of experience with the download platform on decision-making

In our study, 26.83% of the respondents reported that they already had purchased songs from legal music download platforms. Their experience might cause their decision process to differ from that of the remaining 73.17% of respondents without any experience. Our hierarchical Bayes estimation in the empirical study assumes only one common normal distribution on the population level for both types of respondents, but if the groups behave differently, it seems reasonable to shrink their individual estimates not to one common population mean $\bar{\theta}$ but to a conditional mean $\bar{\theta} \cdot z_i$, given experience with online music downloads.

To test the potential gain in model fit, we extend the Bayesian algorithm and account for differences pertaining to whether a respondent already has legally purchased music using an effect-coded covariate variable. The basic implementation for linear utility functions has been described by Lenk et al. (1996), Rossi, McCulloch, and Allenby (1996), and Renken (1997). The major difference between the models with and without covariates appears in the upper level of the Bayes sampler. The hierarchical Bayes model with upper-level covariates assumes that respondents' partworths relate to the covariates through a multivariate regression model, of the following form:

(D.1) $\theta_i = \bar{\theta} \cdot z_i + \xi_i$, where $\xi_i \sim \text{Normal}(0, \Omega)$ $(i \in I)$.

If we assume that individual preferences are modeled by m partworths and n covariates (including a constant), then $\bar{\theta}$ is a $m \times n$ matrix of regression parameters, z_i is a vector of n elements, and ξ_i is a vector of random error terms. The partworths are drawn from a normal distribution with means $\bar{\theta} \cdot z_i$.

TABLE D.1: INTERNAL AND PREDICTIVE VALIDITY (COVARIATE MODEL)

	Internal Validity			Predictive Validity	
	LMD	HR	MAD	HR	MAD
Model 1: DRM restrictions linked to $a^M_{i,j}$, brand name to $c^M_{i,j}$	**-563.99**	**90.62%**	**0.12**	**82.52%**	**0.20**
Model 2: All attributes linked to $a^M_{i,j}$	-632.99	84.23%	0.18	72.76%	0.29
Model 3: All attributes linked to $c^M_{i,j}$	-671.54	85.76%	0.14	80.89%	0.28
Model 4: DRM restrictions linked to $c^M_{i,j}$, brand name to $a^M_{i,j}$	-767.24	80.27%	0.22	70.33%	0.30

Notes: LMD: log marginal density; HR: hit rate; MAD: mean absolute deviations; DRM: digital rights management.

Comparing the results of Table A2 with those of Table 3, we find only marginal gains in internal validity. For example, the log marginal density increases for the first model from −597.24 to −587.64. However, the predictive validity decreases slightly. The changes in internal and predictive validity are small and inconsistent for the other models. Orme and Howell (2009) observe similar results that indicate only modest or no improvements when they include covariates in their models. In summary, we find no significant influence of experience with legal purchases of music downloads on any of the parameters. Therefore, experience provides no additional information to improve parameter estimates or predictions.

E Profit maximization problem

We adapt the model provided by Schlereth, Stepanchuk, and Skiera (2010) to formulate the profit maximization problem for a given number of bucket pricing plans:

(E.1) $$\pi\left(p_J, q_J\right) = \sum_{j \in J} \sum_{i \in I} \left[\left(p_j - k_v \cdot n^*_{i,j}(q_j) \right) \right] \cdot \Pr_{i,j}\left(p_j, q_j\right) \rightarrow \max ,$$

(E.2) $$q_j \le q_{j'} ,$$

(E.3) $$p_j \le p_{j'} , \text{ and}$$

(E.4) $$p_j \in \mathbb{R}_0^+ ; q_j \in \mathbb{N}_0 .$$

The objective function in Equation (E.1) searches for the best set of prices and allowances (pj, qj), which maximizes the profit, calculated as the monthly price minus variable costs times individual consumption, as in Equation (4). Furthermore, we specify pj to be positive and continuous and qj to be positive but discrete.

Assuming the parameters do not change for other types of pricing plans, it is straightforward to formulate the profit maximization problem for three-part pricing plans (see

Equations (E.5)–(E.6)).[9] Each three-part pricing plan consists of a monthly fixed price p_j, an allowance q_j, and a marginal price m_j. The objective function is then

(E.5)

$$\pi\left(p_J, q_J, m_J\right) = \sum_{j \in J} \sum_{i \in I} \left[\left(p_j + m_j \cdot \max(0; n^*_{i,j}(q_j, m_j) - q_j) - k_v \cdot n^*_{i,j}(q_j, m_j) \right) \right] \cdot \Pr_{i,j}\left(p_j, q_j, m_j\right)$$
$$\rightarrow \max,$$

with consumption analogous to that in Iyengar, Jedidi, and Kohli (2008) or Lambrecht, Seim, and Skiera (2007), namely,

(E.6)

$$n^*_{i,j}\left(q_j, m_j\right) = \underbrace{\begin{cases} \dfrac{a_{i,j} - \varpi_i \cdot m_j}{b_{i,j}} & \text{, if } q_j < \dfrac{a_{i,j} - \varpi_i \cdot m_j}{b_{i,j}} \\ \dfrac{a_{i,j}}{b_{i,j}} & \text{, if } q_j > \dfrac{a_{i,j}}{b_{i,j}} \\ q_j & \text{, if otherwise} \end{cases}}_{\text{using the preference model}} = \underbrace{\begin{cases} \dfrac{a^M_{i,j} - m_j}{b^M_{i,j}} & \text{, if } q_j < \dfrac{a^M_{i,j} - m_j}{b^M_{i,j}} \\ \dfrac{a^M_{i,j}}{b^M_{i,j}} & \text{, if } q_j > \dfrac{a^M_{i,j}}{b^M_{i,j}} \\ q_j & \text{, if otherwise} \end{cases}}_{\text{using the surplus model}} .$$

9 This very strong assumption should be abolished in future research. Lambrecht and Skiera (2006) analyze why flat rates are perceived differently from pay-per-use plans; Ascarza, Lambrecht, and Vilcassim (2010) incorporate differences between two-part and three-part pricing plans in their choice model. A potential extension to our study would be to present respondents with multiple types of pricing plans and capture the differences in decision-making with additional parameters.

References

Agarwal, M. K., & Rao, V. R. (1996). An Empirical Comparison of Consumer-Based Measures of Brand Equity. *Marketing Letters,* 7(3), 237-247.

Allenby, G. M., & Rossi, P. E. (1999). Marketing Models of Consumer Heterogeneity. *Journal of Econometrics*, 89(1/2), 57-78.

Ascarza, E., Lambrecht, A., & Vilcassim, N. (2010). When Talk is "Free": An Analysis of Subscriber Behavior Under Two- and Three-Part Tariffs. *London Business School* Working Paper.

Bhattacharjee, S., Gopal, R. D., Lertwachara, K., Marsden, J. R., & Telang, R. (2007). The Effect of Digital Sharing Technologies on Music Markets: A Survival Analysis of Albums on Ranking Charts. *Management Science,* 53(9), 1359-1374.

Brazell, J. D., Diener, C. G., Karniouchina, E., Moore, W. L., Séverin, V., & Uldry, P.-F. (2006). The No-Choice Option and Dual Response Choice Designs. *Marketing Letters,* 17(4), 255-268.

Danaher, P. J. (2002). Optimal Pricing of New Subscription Services: Analysis of a Market Experiment. *Marketing Science,* 21(2), 119-138.

Dilmperi, A., King, T., & Dennis, C. (2011). Pirates of the Web: The Curse of Illegal Downloading. *Journal of Retailing and Consumer Services,* 18(2), 132-140.

Eggers, F., & Sattler, H. (2009). Hybrid Individualized Two-Level Choice-based Conjoint (HIT-CBC): A New Method for Measuring Preference Structures with Many Attribute Levels. *International Journal of Research in Marketing,* 26(2), 108-118.

Elberse, A. (2010). Bye Bye Bundles: The Unbundling of Music in Digital Channels. *Journal of Marketing,* 74(3), 107-123.

Goettler, R. L., & Clay, K. (2010). Tariff Choice with Consumer Learning and Switching Costs. *Journal of Marketing Research,* forthcoming.

Hensher, D. A., Stopher, P. R., & Louviere, J. J. (2001). An Exploratory Analysis of the Effect of Numbers of Choice Sets in Designed Choice Experiments: an Airline Choice Application. *Journal of Air Transport Management*, 7(6), 373-379.

Hui, W., Yoo, B., & Tam, K. Y. (2007). The Optimal Number of Versions: Why Does Goldilocks Pricing Work for Information Goods? *Journal of Management Information Systems,* 24(3), 167–191.

Iyengar, R. (2010). Pricing for a Menu of Service Plans – An Application to the Digital Music Industry. Working Paper.

Iyengar, R., Ansari, A., & Gupta, S. (2007). A Model of Consumer Learning for Service Quality and Usage. *Journal of Marketing Research*, 44(4), 529-544.

Iyengar, R., & Jedidi, K. (2010). A Conjoint Model for Quantity Discounts. Working Paper.

Iyengar, R., Jedidi, K., & Kohli, R. (2008). A Conjoint Approach to Multi-Part Pricing. *Journal of Marketing Research,* 45(2), 195-210.

Kadiyali, V., Sudhir, K., & Rao, V. R. (2001). Structural Analysis of Competitive Behavior: New Empirical Industrial Organization Methods in Marketing. *International Journal of Research in Marketing*, 18(1-2), 161-186.

Kass, R. E., & Raftery, A. E. (1995). Bayes Factors and Model Uncertainty. *Journal of the American Statistical Association,* 90(430), 773-795.

Lambrecht, A., Seim, K., & Skiera, B. (2007). Does Uncertainty Matter? Consumer Behavior under Three-Part Tariffs. *Marketing Science,* 26(5), 698-710.

Lambrecht, A., & Skiera, B. (2006). Paying Too Much and Being Happy About it: Existence, Causes and Consequences of Tariff-Choice Biases. *Journal of Marketing Research,* 43(2), 212-223.

Leland, H., & Meyer, R. (1976). Monopoly Pricing Structures with Imperfect Discrimination. *Bell Journal of Economics,* 7(2), 449-462.

Lenk, P. J., DeSarbo, W. S., Green, P. E., & Young, M. R. (1996). Hierarchical Bayes Conjoint Analysis: Recovery of Partworth Heterogeneity from Reduced Experimental Designs. *Marketing Science,* 15(2), 173-191.

Louviere, J. J., Hensher, D. A., & Swait, J. D. (2000). Stated Choice Methods: Analysis and Application. Cambridge: Cambridge University Press.

Maskin, E., & Riley, J. (1984). Monopoly with Incomplete Information. *Rand Journal of Economics*, 15(2), 171-196.

Moore, W. L. (2004). A Cross-Validity Comparison of Rating-Based and Choice-Based Conjoint Analysis Models. *International Journal of Research in Marketing,* 21(3), 299-312.

Moorthy, S., Ratchford, B. T., & Talukdar, D. (1997). Consumer Information Search Revised: Theory and Empirical Analysis. *Journal of Consumer Research,* 23(4), 263-276.

Murphy, M. M. (1977). Price Discrimination, Market Separation, and the Multi-Part Tariff. *Economic Inquiry,* 15(4), 587-599.

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

Newton, M. A., & Raftery, A. E. (1994). Approximate Bayesian Inference with the Weighted Likelihood Bootstrap. *Journal of the Royal Statistical Society,* 56(1), 3-48.

Orme, B., & Howell, J. (2009). Application of Covariates Within Sawtooth Software's CBC/HB Program: Theory and Practical Example. Sawtooth Software Research Paper Series.

Papies, D., Eggers, F., & Wlömert, N. (2010). Music for Free? How Free Ad-Funded Downloads Affect Consumer Choice. *Journal of the Academy of Marketing Science,* forthcoming.

Parker, J. R., & Schrift, R. Y. (2011). Rejectable Choice Sets: How Seemingly Irrelevant No-Choice Options Affect Consumer Decision Processes. *Journal of Marketing Research,* forthcoming.

Prelec, D., & Loewenstein, G. (1998). The Red and the Black: Mental Accouting of Savings and Debt. *Marketing Science,* 17(1), 4-28.

Renken, T. (1997). Disaggregate Discrete Choice. *Marketing Research,* 9(1), 18-22.

RIAA (2009). 2008 Year-End Shipment Statistics. Report.

Rob, R., & Waldfogel, J. (2006). Piracy on the High C's: Music Downloading, Sales Displacement, and Social Welfare in a Sample of College Students. *Journal of Law and Economics Bulletin,* 49(1), 29-62.

Rossi, P. E., McCulloch, R. E., & Allenby, G. M. (1996). The Value of Purchase History Data in Target Marketing. *Marketing Science,* 15(4), 321-340.

Schlereth, C., Skiera, B., & Wolk, A. (2011). Measuring Consumers' Preferences for Metered Pricing of Services. *Journal of Service Research,* forthcoming.

Schlereth, C., Stepanchuk, T., & Skiera, B. (2010). Optimization and Analysis of the Profitability of Tariff Structures with Two-Part Tariffs. *European Journal of Operational Research,* 206(3), 691-701.

Senecal, S., & Nantel, J. (2004). The Influence of Online Product Recommendations on Consumers' Online Choices. *Journal of Retailing,* 80(2), 159-169.
Sinha, R. K., Machado, F. S., & Sellman, C. (2010). Don't Think Twice, It's All Right: Music Piracy and Pricing in a DRM-Free Environment. *Journal of Marketing,* 74(2), 40-54.
Sinha, R. K., & Mandel, N. (2008). Preventing Digital Music Piracy: The Carrot or the Stick? *Journal of Marketing,* 72(1), 1-15.
Sonnier, G., Ainslie, A., & Otter, T. (2007). Heterogeneity Distributions of Willingness-to-Pay in Choice Models. *Quantitative Marketing and Economics,* 5(3), 313-331.
Street, D. J., Burgess, L., & Louviere, J. J. (2005). Quick and Easy Choice Sets: Constructing Optimal and Nearly Optimal Stated Choice Experiments. *International Journal of Research in Marketing,* 22(4), 459-470.
Sundararajan, A. (2004). Managing Digital Piracy: Pricing and Protection. *Information Systems Research,* 15(3), 287-308.
Swait, J., & Andrews, R. L. (2003). Enriching Scanner Panel Models with Choice Experiments. *Marketing Science,* 22(4), 442-460.
Wedel, M., Vriens, M., Bijmolt, T. H. A., Krijnen, W., & Leeflang, P. S. H. (1998). Assessing the Effects of Abstract Attributes and Brand Familiarity in Conjoint Choice Experiments. *International Journal of Research in Marketing,* 15(1), 71-78.
Wertenbroch, K., & Skiera, B. (2002). Measuring Consumer Willingness to Pay at the Point of Purchase. *Journal of Marketing Research,* 39(2), 228-241.
Wilson, R. (1993). Nonlinear Pricing. New York et al.: Oxford University Press.
Yang, H., & Ye, L. (2008). Nonlinear Pricing, Market Coverage, and Competition. *Theoretical Economics,* 3(1), 123-153.