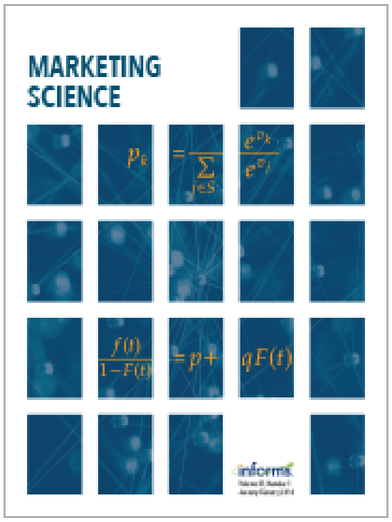


Marketing Science

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

An Empirical Analysis of Consumer Purchase Decisions Under Bucket-Based Price Discrimination

Yacheng Sun, Shibo Li, Baohong Sun

MARKETING SCIENCE
Vol. 34, No. 5, September–October 2015, pp. 646–668
ISSN 0732-2399 (print) | ISSN 1526-548X (online)

http://dx.doi.org/10.1287/mksc.2015.0942

An Empirical Analysis of Consumer Purchase Decisions Under Bucket-Based Price Discrimination

Yacheng Sun
Leeds School of Business, University of Colorado at Boulder, Boulder, Colorado 80021, yacheng.sun@colorado.edu

Shibo Li
Kelley School of Business, Indiana University, Bloomington, Indiana 47405, shili@indiana.edu

Baohong Sun
Cheung Kong Graduate School of Business, New York, New York 10019, bhsun@ckgsb.edu.cn

Bucket-based price discrimination is a unique price format that involves monthly subscription fees and instantaneous quotas (the number of rental products that can be checked out). We propose an empirical model in which consumers make dynamic purchase decisions under consumption uncertainty, accounting for the constraints imposed by the instantaneous quota. Applying the model to an online DVD rental data set, we find that (1) consumers incur a large disutility (∼ \$8) from *stockout* (i.e., unmet consumption needs); (2) such a disutility drives consumers' *overpurchase* of the service quota as a way to avoid potential stockout situations; and (3) the dynamics of overpurchase are driven by the interplay between trends in consumption needs and the magnitude of consumers' plan-switching costs. We run counterfactual exercises to better understand how the instantaneous quota and stockout risk affect consumers' consumption rates, purchase decisions, and firm profitability. We find that the instantaneous quota induces a greater stockout compared with a monthly quota. We further demonstrate that the company should recognize the drivers of the dynamics in overpurchase to balance short- and long-term profitability—for example, by offering targeted discounts to customers with excess overpurchase.

Keywords: bucket-based price discrimination; subscription service; usage uncertainty; stockout; overpurchase; dynamic purchase decisions; optimal pricing; product design
History: Received: March 30, 2012; accepted: June 4, 2015; Preyas Desai served as the editor-in-chief and K. Sudhir served as associate editor for this article.

1. Introduction

The past decade has seen a novel price format gain great popularity in many online rental services. Well known by its tagline "\$$X$ per month for Y," this price format involves a menu of "tiered" service plans. A representative service plan consists of a fixed monthly subscription fee (\$$X$) and a number of rental products that can be checked out at any given point in time (Y). We refer to this price format as *bucket-based price discrimination* (BBPD) for two reasons. First, it is a new second-degree price discrimination mechanism (e.g., Mussa and Rosen 1978, Rochet and Stole 2002), where consumers self-select into service plans vertically differentiated by a quota. Second, the quota represents a fixed consumption capacity that cannot be exceeded during the subscription period, which has been metaphorically referred to as a "bucket" (e.g., Lovelock and Wirtz 2007). Table 1 lists additional popular examples of continuous subscription services using BBPD, along with their offerings. A salient example of BBPD service is Netflix: the consumer starts by paying a fixed monthly fee (e.g., \$15.99). Netflix allows the consumer to switch up her plan on any day, and downgrade her plan only at the end of the month. The consumer then receives a fixed number of DVDs (e.g., three) delivered by the United States Postal Service (USPS). The consumer again uses USPS to return movies in exchange for new movies that are sent to her afterward.

Figure 1 plots the price structure of BBPD, along with other popular nonlinear price formats in the subscription-service industry: flat tariffs and two-part pricing (Train et al. 1987, Danaher 2002, Narayanan et al. 2007), increasing block pricing (Iyengar et al. 2007), and three-part tariffs (Lambrecht et al. 2007).

It is insightful to compare BBPD and the three-part tariff (3PT), which also incorporates a fixed subscription price and a quota. Both price formats require consumers to advance purchase (Xie and Shugan 2001). Similar to 3PT, BBPD requires the consumer to make plan choices based on expectations of future consumption. These two price formats, however, differ in

Table 1 Examples of BBPD

Industry	Representative service[a]	Monthly price ($)	No. of rentals allowed at one time
Online movie rental	Netflix.com	7.99	1
	Blockbuster Online	11.99	2
Online game rental	Gamefly.com	15.99	3
	Gamerang.com	15.95	1
	Gottaplay.com	22.95	2
	Rentzero.com	29.95	3
	GameMine.com	36.95	4
Book rental	Bookswim.com	23.95	3
	Booksfree.com	29.95	5
	Paperspine.com	35.95	7
	Skoobit.com	59.95	11
Online CD and audio book rental	Audiotogo.com	16.99	1
	Jiggerbug.com	24.99	2
	Audiobooksonline.com	34.99	3
	Kitabe.com	41.99	4
		51.99	5

Note. All plan information was retrieved on November 15, 2013.
[a]Columns 3 and 4 describe the service structure of the business listed first for the category, e.g., Netflix.com in the online movie rental category.

two important ways. First, 3PT gives the consumer the option to incur monetary costs (marginal fees) to cover excessive consumption (Lambrecht et al. 2007), and the quota of 3PT serves as a threshold above which the marginal fee is assessed. By contrast, the quota of BBPD serves as a limit for consumption, and the consumer incurs the costs from unmet consumption (referred to as *stockout*). Second, whereas 3PT (and other nonlinear price formats such as two-part pricing) sets the quota at the monthly level, the quota of BBPD is usually set at the daily level. We refer to such a daily quota as the *instantaneous quota* because it restricts the number of rental products available for instantaneous consumption on any given day. The difference between monthly and instantaneous quotas has subtle, yet important implications for the consumer. Intuitively, a consumer who tends to bunch consumption on specific days is more likely to be restricted by the instantaneous quota, compared with the monthly quota, which gives her more flexibility in matching the quota with her time-varying consumption needs. In other words, the monthly quota discriminates based on the mean consumption needs of the subscription period, and BBPD discriminates based on both the mean and peak consumption needs during the subscription period.

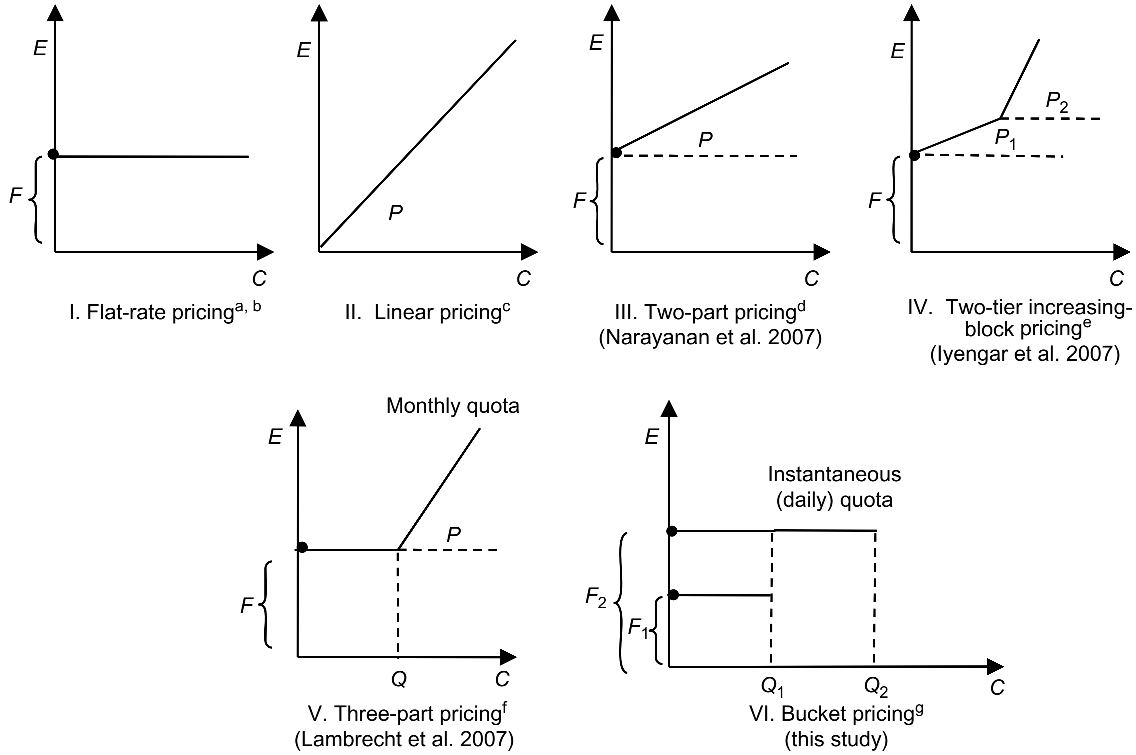
BBPD with instantaneous quota has become a very popular price format for rental services, ranging from movies (Netflix), designer handbags (BagBorroworSteal), toys (BabyPlays), and books (BookSwim). For example, Netflix.com, the market leader of the online movie rental industry, serves 33.1 million subscribers in the United States, and more than 50 million users globally. Despite the increasing popularity of BBPD and its unique structure described above, it has garnered little academic attention. This is in sharp contrast with the recent surge of research on nonlinear pricing in the telecommunications industry. Current research on nonlinear pricing has studied the competitive conditions under which flat-rate pricing is optimal (Hitt and Chen 2005, Essegaier et al. 2002, Oi 1971, Wilson 1993) and has provided empirical evidence of a bias for flat-fee over two-part pricing (Danaher 2002; Hobson and Spady 1988; Kling and van der Ploeg 1990; Kridel et al. 1993; Mitchell and Vogelsang 1991; Train et al. 1989, 1987; Miravete 2002a, b; Lambrecht and Skiera 2006; Narayanan et al. 2007).

Two closely related studies are Iyengar (2010) and Schlereth and Skiera (2012). Iyengar (2010) conducted a conjoint study to estimate consumers' willingness to pay for digital songs priced via pay-per-use pricing and bucket pricing (BP). He found that by offering both a BP plan and a pay-per-use plan, the company earned more profit than by offering a pay-per-use plan only. Applying a Bayesian model to the survey data, Schlereth and Skiera (2012) estimated consumers' plan-specific preferences and then used simulations to compare the performance of BP and two pay-per-use (linear) plans: a two-part tariff and 3PT. The authors found that the optimal BP was as profitable as the other nonlinear price formats. Our empirical investigation differs from these two studies in two important ways. First, both papers considered monthly level consumption and ignored the daily level constraint of the instantaneous quota on consumption. Second, both studies used survey data rather than transactional data, and focused on eliciting the consumers' willingness to pay in a *static* context. However, BBPD services (e.g., Netflix) are typically continuous subscription services; thus, it is natural to examine consumers' purchase decisions in a *dynamic* decision framework, which we adopt. As a result, we expect that various factors—such as the dynamics of consumers' consumption needs (e.g., Lambrecht et al. 2007), uncertainty in usage, and switching costs (e.g., Goettler and Clay 2011)—will give rise to interesting dynamics in consumers' plan choice and retention decisions. These dynamics also have important implications for companies wishing to balance short- and long-term profits.

To summarize, the design of BBPD entails a unique set of decision calculus for consumers and strategic implications for the company, leading to several interesting research questions:

- What is the effect of the instantaneous quota on consumers' consumption?
- What drives consumers' purchase decisions in BBPD?
- Do consumers dynamically change their purchase decisions over time?

Figure 1 Comparison Between Bucket Pricing and Other Pricing Formats



[a]In I–VI, the horizontal axis, C, is the usage amount of the service, and the vertical axis, E, denotes the expenditure.
[b]In I: F is the fixed fee.
[c]In II: P is the marginal price.
[d]In III: F is the fixed fee, and P is the marginal price.
[e]In IV: F is the fixed fee, and P_1 (P_2) is the first (second) marginal price.
[f]In V: F is the fixed fee, Q is the "free" allowance, and P is the marginal price.
[g]In VI: Only two plans are shown here: F_1 (F_2) is the price for the first (second) plan, and Q_1 (Q_2) is the quota of the first (second) plan.

• How can the company improve its BBPD design to increase its long-term profitability?

To answer these questions, we propose an empirical model that encapsulates the key aspects describing consumers' dynamic decision process under BBPD: uncertainty about future consumption, the instantaneous quota, the potential risk of stockout, and switching costs. To realistically model the instantaneous quota, we combine a daily level consumption model with a monthly purchase decision model. The model allows consumers to form expectations of their future consumption needs based on idiosyncratic consumption patterns, and recognizes the disutility when consumption needs are capped by the instantaneous quota. We then apply the model to a unique panel data set of consumer purchases and consumption history provided by an online DVD rental service (focal company). The proposed model recognizes the focal company's idiosyncratic plan-switching policy (unlike Netflix, consumers can only make purchase decisions once every month) and explains several empirical regularities emerging from the data. First, on average, consumers substantially overpay for the service quota (referred to as *overpurchase*), resulting in a high effective price per movie rental. Using our model, such overpurchase can be rationalized by the high disutility from unmet consumption needs, or stockout (approximately \$8 per stockout). Therefore, overpurchase can be thought of as "insurance" bought to ensure that future consumption is met. Second, we also find strong evidence for *lock-in*, manifested by high and persistent overpurchase. Many consumers simply do not adjust their plan choices frequently enough, and consequently forgo opportunities for monetary savings.

Third, across consumers, there are interesting differences in the dynamics of overpurchase. Such dynamics can be explained by the interplay between the high switching costs and the systematic change in consumption needs over time: some consumers exhibit a "fatigue" effect, i.e., their consumption needs decrease with accumulated consumption, whereas others exhibit an opposite, "reinforcing" effect. We find that consumers with low switching costs overpurchase earlier and adjust their plan choices more frequently to match the evolution of their expected consumption needs so that their overpurchase diminishes over time. Consumers with higher switching costs realize the possible long-term financial expenditure induced by switching costs, such that they overpurchase less early, but then seldom adjust purchases and end up with more overpurchase. The retention rate for low-switching-cost customers, however, is much higher than that for high-switching-cost customers.

Based on the parameter estimates from the empirical model, we use counterfactual exercises to better understand the two key design components of BBPD: the daily level instantaneous quota and the stockout risk it induces. We find that given consumers' high risk aversion to stockout, BBPD allows the focal company to charge high subscription fees and generate a large profit. The focal company would rather not charge a marginal fee and would instead prefer to deny consumers the opportunity to cover stockout with a marginal fee. In addition, we examine how the company can improve its current BBPD design to influence its overpurchase dynamics, retention rate, and overall profit. We find that overpurchase is a double-edged sword: although the company can generate higher profits from locked-in customers, it also causes customers to defect earlier. Consequently, the company can actively fine-tune its marketing mix based on the observed overpurchase and usage. For instance, the company may benefit from offering targeted price discounts to consumers with excessive overpurchase.

2. Background and Data

2.1. Industry Overview

Our empirical investigation focuses on an anonymous online DVD rental company ("the focal company"). The focal company employs the standard DVD rental model, which creatively integrates Internet technology and USPS service. Consumers choose among plans characterized by different combinations of price and quota. They then furnish credit card information so that the company can automatically debit monthly payments from their accounts. Consumers can log onto the company's website to browse movies and create personal queues of movie titles in the order of their viewing preference. Consumers receive the DVDs in the mail and can keep the movies for as long as they like without incurring any late fees. To return the rented DVDs, they mail them back using a postage-paid envelope provided by the company. When the company receives the returned DVDs, it mails the same number of movies to the consumers. The process continues until the subscription is terminated. There is no long-term contract; however, the subscription process is automatically renewed every month unless the consumers change the plan or leave the service. For the company, revenue comes solely from the monthly subscription fees. On the cost side, other than the overhead costs and fees paid to stock the DVDs, the main variable cost for online DVD rental companies is postage; for the focal company, such a cost is $0.90 for two-way shipments. The company provided us with consumer panel data containing a subset of randomly selected registered consumers whose purchase and shipment histories were tracked.

2.2. Overview of the Data

The detailed shipment history for a representative consumer includes the dates when each movie was shipped out to the consumer and was received by the company. The shipping dates, along with the company-estimated one-way shipping time, are used to infer the dates when the consumer received the DVDs, based on the assumption that consumers return the DVDs immediately after watching them (Milkman et al. 2009). Similarly, the receiving dates allowed us to infer the dates of consumption.

The purchase history of a representative consumer includes the dates when the service was initiated (and, possibly, terminated), and dates of her entire history of subsequent plan choices. Two observations can be made regarding the payment histories. First, no customers reinstated service after terminating it. Second, with a few exceptions, all payments were made at the beginning of the payment cycle (month), which is consistent with the company's policy that any plan change does not take effect until the next month. Unfortunately, consumer demographic information was very limited, except whether the consumer resided within the same state as the company.

Table 2 provides the key sample statistics, based on the purchase and consumption data. For example, the average monthly payment was $20.68. Price discounts, averaging $0.50, were offered in approximately 2.57% of cases. The average actual movie consumption per month was 2.71, with a standard deviation of 2.37. During the observation period, consumers stayed with the company for an average of 7.68 months. Finally, 93% of the consumers lived in states different from the focal company, and thus, did not pay sales tax.

Table 3 summarizes the observed plan choices with the focal company. Columns 2 and 3 list the monthly prices and the instantaneous quota of the six service plans offered by the company. Columns 4–7 show the average numbers of movies consumed per month, total revenue, total variable cost, and total profit. We make several observations. When the quota increases, so does the monthly subscription price, although at a slower rate. The dominant purchase share rests

Table 2 Summary Statistics

Variables	Explanation	Mean	Standard deviation
P_{ijt}	Monthly payment including tax	20.68	3.89
$DSCT_{ijt}$	Amount of discount off monthly payment	0.50	2.72
C_{it}	Actual monthly consumption	2.71	2.37
TENURE	Number of months with the company	17.68	6.25
MON	Monday dummy	0.143	0.350
TUE	Tuesday dummy	0.142	0.349
WED	Wednesday dummy	0.142	0.349
THU	Thursday dummy	0.143	0.350
FRI	Friday dummy	0.143	0.350
SAT	Saturday dummy	0.143	0.350
SUN	Sunday dummy	0.142	0.349
D_{TAX_i}	Dummy variable equal to 1 if the consumer resides outside the state and 0 otherwise	0.07	0.25

with the standard plan, followed by premium and lite plans; the elite plan has the lowest purchase share. Thus, plan popularity does not appear to increase with a volume discount. Rather, the fee and quota seem to jointly determine the popularity of a service plan. Intuitively, given the low market shares of high-quota (and more profitable) plans, there is potential for the company to better align the popularity and profitability of its service plans by either making popular plans more profitable or by making profitable plans more popular.

A further look into the plan switching in the data shows four patterns. First, in the majority of subscription periods, consumers chose to stay with the status-quo plan choices. The average number of switches (including both switching among plans and dropping out) for a consumer during her entire tenure was 1.24. There was also significant heterogeneity in switching frequencies across consumers: whereas 81.4% of all consumers only switched once, 4.8% of them switched more than three times. Plan stickiness was highest for the standard plan and lowest for the elite plan. Second, most switches involved a move to adjacent plans. For the standard plan and up, more consumers switched down to lower-level plans; however, consumers starting with the economy and lite plans were more likely to switch up to standard or beyond. Third, consumers were more likely to switch down than up. Fourth, the attrition rate was higher for consumers with the advantage and elite plans.

2.3. Initial Insights Into Consumers' Consumption

We first examine the patterns of consumption observed in the data. Consumers' realized consumption is discrete in nature, for which a Poisson model is a good choice. However, we find that there are an excessive number of zero consumption occasions: there are a large percentage (89%) of occasions where the consumer held at least two movies, but had zero consumption. Apparently, this excessive zero consumption suggests that the zero-inflated Poisson (ZIP) model is a more appropriate modeling choice than a simple Poisson model. A further check shows that, consistent with the ZIP model, the variance of daily level consumption (0.137) is significantly larger than the mean (0.084).

It is also important to account for potential censoring, since on any given day, the number of movies that can be watched cannot exceed the number of available movies and the instantaneous quota. To check whether consumers' daily level consumption is indeed censored by the number of movies available from above, we compute the average observed consumption for each possible number of available movies. Because it is possible that consumers are more likely to watch movies during the weekends, this computation is separately conducted for weekdays (Monday–Thursday) and weekends (Friday–Sunday). Confirming prior expectations, we find that conditional on either a weekend or weekday, the average consumption increases with the number of movies available: the average daily consumption rate is 0.075, 0.112, and 0.157 when there are 1, 2, and 3 movies available, respectively. Furthermore, conditional on the number of available movies, the average consumption

Table 3 Prices and Quotas of Alternative Plans and Profit Contribution

Plans	Price ($)	DVD quota	Average actual consumption[b]	Total revenue ($)	Total variable costs[c] ($)	Total profits ($)
Economy[a]	9.95	1	1.26	33,352	8,446	24,906
Lite	12.95	1	1.53	65,475	15,471	50,004
Standard	19.95	2	2.82	1,980,935	560,024	1,420,911
Premium	27.95	3	4.31	186,147	57,408	128,739
Advantage	37.95	5	6.29	140,339	46,520	93,819
Elite	57.95	7	8.02	51,402	14,226	37,176

[a]The total monthly consumption on the economy plan is limited to two.
[b]Average actual consumption is the total number of DVDs shipped to the consumer each month, adjusted by the DVDs not shipped back at the end of that month and the DVDs that the customer held over from the previous month.
[c]Variable cost is approximated as the sum of the postage cost, or $0.45 for one-way delivery and an estimated $1.10 for overhead costs.

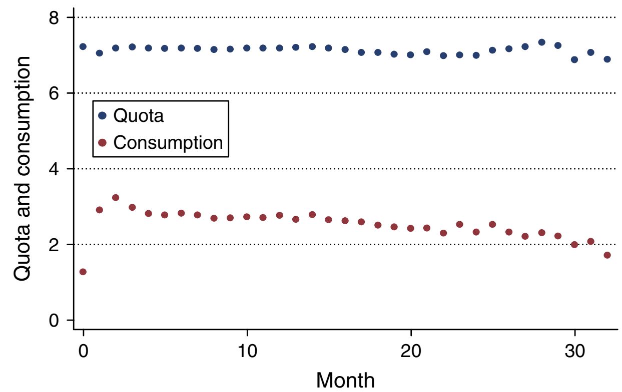
rate is larger on weekends, compared to weekdays. For example, when the number of available movies is two, the average consumption rate for weekdays (weekends) is 0.094 (0.160). To summarize, the consumption data support a zero-inflated Poisson modeling with censoring (at the available movie inventory). We formally develop the consumption model in §3.1.

2.4. Initial Insights Into Consumers' Purchase Decisions

To understand the possible drivers of consumers' purchase decisions, we look into the systematic patterns in consumers' dynamic purchase histories. Three interesting empirical regularities in consumers' purchase decisions emerge. First, consumers' consumption needs are not always fully covered by purchased consumption capacity: consumers hit their instantaneous quota for a nontrivial percentage ($\sim 4\%$) of all occasions, indicating a possible risk of stockouts (i.e., unmet consumption needs) at the daily level. Second, we observe significant and persistent *overpurchase:* across all plans, the average actual consumption rates are less than half of the purchased consumption capacities. Consumers thus overpay a significant amount (an average effective price of $6–$8 per movie rental), compared with the "ideal" situation where the plan is used more fully. Third, there is strong evidence for consumer lock-in with their current plan choices. Figure 2 plots the dynamics of average purchased consumption capacity and average movie consumption by month. Consistent with the overpurchase discussed previously, purchased consumption capacity is always higher than actual consumption. Furthermore, whereas there is a notable decline in the average consumption rate, the average amount of the purchased quota remains relatively steady. This finding implies that consumers are not fully responsive in adjusting their plan choices based on their reduced consumption. Even though some consumers adjust their purchases to avoid excessive overpurchase, many of them stay with their current plans and increasingly overpay for the service. Interestingly, 6.4% of consumers who eventually discontinued the service had *zero* consumption for at least one month before discontinuing the service. Presumably, such lock-in situations can be attributed to these consumers' switching costs, which have been found to be economically significant for subscription services (Goettler and Clay 2011).

In the context of the continuous subscription industry, a forward-looking framework is suitable for studying the dynamics in purchase choices that arise because of switching costs (Goettler and Clay 2011), and the dynamics in consumption needs (Lambrecht et al. 2007). To test whether consumers' decisions are affected by future state variables, we run the following reduced-form regression:

Figure 2 (Color online) Evolution of Purchased and Realized Consumption Over Tenure



$$\begin{aligned}\log(OVP_{im}) &= \phi_{0i}+\phi_1\cdot\log(OVP_{i,m-1})+\phi_2\cdot Consump_{i,m+1}\\ &\quad+\phi_3 Nswitch_i+\phi_4\cdot Consump_{i,m+1}\cdot Nswitch_i\\ &\quad+\phi_5 PeakConsump_{i,m+1}+\phi_6\cdot NumPeakConsump_{i,m+1}\\ &\quad+\varepsilon_{im},\end{aligned}$$

where the dependent variable, $\log(OVP_{im})$, is the log of the overpurchase by consumer i in month m, and $\log(OVP_{i,m-1})$ is the log of the amount of overpurchase in the preceding month, which we include to capture the possible inertia in the overpurchase. The variable $Consump_{i,m+1}$ is the total amount of consumption in the next month, and is a reduced-form approximation of future consumption needs (assuming rational expectations by the consumers). The main distinction between a static and dynamic model is whether consumers' purchase decisions are influenced by future state variables. $Nswitch_i$ is the total number of switches and is a proxy for the (lack of) switching costs. $PeakConsump_{i,m+1}$ is the maximal realized consumption in the next month, and $NumPeakConsump_{i,m+1}$ is the number of days on which maximal consumption occurred.

We find that ϕ_1 is positive ($\phi_1 = 0.144$, $t = 12.2$), indicating strong inertia toward overpurchase. The second coefficient ϕ_2 is positive ($\phi_2 = 0.006$, $t = 3.65$): when future consumption needs are high, forward-looking consumers are more likely to purchase higher plans in the current period to avoid future stockout. ϕ_3 is negative ($\phi_3 = -0.225$, $t = -23.7$), suggesting that consumers with higher switching costs on average pay more for the quota. An intuitive and rational explanation is that whereas these consumers do incur a higher disutility from payment, they are able to avoid the mental costs associated with frequent switching. ϕ_4 is negative ($\phi_4 = -0.059$, $t = -27.4$), indicating that consumers with both higher switching costs and high consumption needs are more likely to buy higher plans to cover future consumption. Both ϕ_5 and ϕ_6 are positive and significant ($\phi_5 = 0.0008$, $t = 2.05$, $\phi_6 = 0.328$, $t = 12.7$), again implying that consumers consider potential stockout risk factors in the future when making their purchase decisions—consistent with forward-looking consumers.

To sum, we found in the data evidence of stockout, overpurchase, (imperfect) lock-in, and consumer forward looking, which motivates a purchase utility model that incorporates the instantaneous quota, usage uncertainty, switching costs, and consumer forward looking. This model, presented in §3.2, enables us to rationalize seemingly suboptimal consumer choices and explain the systematic dynamics of overpurchase.

3. The Model

In this section, we first present the ZIP model for the consumer's daily consumption needs and demonstrate how a representative consumer forms her expected consumption and stockout, accounting for the role of the daily level instantaneous quota. We then present the model for the consumer's monthly purchase utility.

3.1. The ZIP Model for Consumption Needs

We model the consumer's realized consumption needs as a ZIP process with censoring based on three considerations. First, a Poisson model is a natural choice for modeling discrete consumption outcomes, which includes both zero and positive integers (e.g., two movies). Second, based on previous research in consumer-packaged goods (e.g., Erdem et al. 2003), it is important to account for occasions when consumers may have zero consumption needs for the service (e.g., the consumer may be too busy to watch any movies). Adding the zero-inflation part is also consistent with the fact that realized consumption is zero for many days, even when the consumer has movies available for consumption. Finally, the censoring is consistent with the fact that the number of movies available restricts consumption.

In the absence of restrictions involving the instantaneous quota, the daily expected consumption needs, c_{it}^*, is assumed to follow a ZIP distribution with two parameters, $c_{it}^* \sim \text{ZIPoisson}(\pi_{it}, \lambda_{it})$. The first parameter, π_{it}, captures consumer i''s decision about whether to consume on day t. Specifically, π_{it} is the probability that the consumer does *not* have a need for movie consumption, such that

$$\pi_{it} = \text{Logit}(\theta_{0i} + \theta_{1i} \cdot Weekend_t + \theta_{2i} \cdot Ngenre_{it}). \quad (1)$$

In Equation (1), θ_{0i} is the baseline tendency for consumer i to have zero consumption needs. The second term, $Weekend_t$, is dummy-coded to 0 if day t is Monday through Thursday, and 1 if day t is Friday through Sunday. Thus, θ_{1i} measures the difference in consumer i's tendency to watch movies on a weekday compared with the weekend. We expect θ_{1i} to be negative for consumers who are more likely to watch movies on weekends, and the magnitude of θ_{1i} to be smaller for consumers who have flexibility in watching movies on weekdays. The final term, $Ngenre_{it}$, is the cumulative number of genres consumer i has watched up to time t. We include this term to account for the possibility that more variety-seeking consumers may be more likely to experience a consumption need.

The second parameter of the ZIP model, λ_{it}, is the mean of the Poisson model. It captures factors affecting the consumer's decision about *how many* movies to watch

$$\lambda_{it} = \alpha_{0i} + \alpha_{1i}\, Weekend_t + \sum_{k=1}^{3} \alpha_{1+k,\,i}\, PastConsump_{it,\,k} + \alpha_{5i}\, AccConsump_{it}. \quad (2)$$

In Equation (2), α_{0i} is the baseline rate of consumption needs. $Weekend_t$ captures the persistent difference between the consumption needs on weekends versus weekdays, conditional on having a consumption need. To account for the persistent consumption pattern, we include realized consumption on the same weekday (e.g., Thursdays when t is a Thursday) in the past three weeks, denoted by $PastConsump_{it,k}$ $(k = 1, 2, 3)$. The coefficient $\alpha_{1+k,\,i}$ captures potential habit persistence in consumption needs (e.g., Chaloupka 1988, Mullahy 1986).[1] The term $AccConsump_{it}$ measures accumulated movie consumption for consumer i up to time t. The effect of accumulated consumption can be either a positive, "reinforcing" effect (consumers become more addicted to movie consumption) or a negative, "fatigue" effect. The fatigue effect can occur, given the limited inventory of the company's content-edited movies. Including accumulated consumption allows for the possibility that it becomes more difficult to find movies suitable for consumption as accumulated consumption increases. A priori, it is unclear which effect is stronger, and α_{5i} should be interpreted

[1] There exists empirical evidence that consumers may have time-inconsistent preferences and exercise self-control (e.g., Wertenbroch 1998), which is certainly possible for movie consumption (Milkman et al. 2009, Read et al. 1999). Including past consumption into the consumption needs equation also accounts for the possibility that consumers exhibit self-control, such that they reduce current consumption needs when past consumption has been excessive (e.g., Jain 2012). Thus, the coefficients should be interpreted as the net effect of habit formation and self-control. Broadly speaking, the degree of self-control likely depends on the type of rental products; for example, one may expect self-control to be more serious for businesses specializing in video game rentals (e.g., Gamefly), compared with those that rent books (e.g., BookSwim).

as the net effect of accumulated consumption on consumers' consumption needs.[2]

3.2. The Purchase Utility Model

Informed by the focal company's policy that plan switching can only occur at the beginning of each month,[3] we model consumers' purchase decisions at the monthly level. Note that other BBPD services may use different plan-switching policies. For example, the most recent policy of Netflix allows its customers to upgrade on any given day and downgrade at the beginning of the next month. Nevertheless, our modeling framework can be readily adapted to accommodate alternative types of plan-switching policies without affecting the way we model the instantaneous quota and stockout risk. In §4.5, we investigate the effects of Netflix's plan-switching policy with a counterfactual exercise.

The company offers J plans ("buckets"). Each plan j is defined by a fixed and prepaid monthly fee P_j and an *instantaneous quota* IQT_j, which stipulates the number of "outstanding" movies allowed at one time. There are I consumers who make plan-choice decisions among the J plans at the beginning of each month $m = 1, \ldots, M$. Consumers can also choose the outside option $j = 0$, thereby discontinuing the service.

Because of the temporal separation between payment and consumption, consumers must form expectations about their future daily consumption needs and then compare them with the quota of each plan, so that the expected consumption realization and expected consumption stockout can be determined. We propose a simple model to approximate the expected daily consumption needs and stockout at the beginning of each payment cycle. Next, we consider how the available quota restricts consumers' consumption decisions at the daily level (compared with the monthly level), a fact that has not been addressed by existing research (Iyengar 2010, Schlereth and Skiera 2012).

[2] We conducted various robustness checks before settling on Equation (2). Specifically, we tested for the possible inclusion of the quadratic form of accumulated consumption, which we found to be statistically insignificant ($p > 0.40$). We tested different orders of $PastConsump_{it,k}$ and found that the model fit significantly increased, with up to a third lagged consumption; improvements from higher orders of $PastConsump_{it,k}$ were negligible. We also checked for multicollinearity between accumulated consumption and lagged weekly consumption. These correlations were all less than 0.16. Furthermore, following Menard (2002), we found that the variance inflation factors were all smaller than 2. Thus, multicollinearity was not an issue. Finally, we included days to the next payment date to test whether there was a payment effect on consumption, and found no evidence.

[3] This policy is confirmed by observed plan decisions in the data: with only a few exceptions, all payments occurred at regular monthly intervals.

3.2.1. Instantaneous Quota and Available Quota. A key characteristic of the BBPD model is the instantaneous quota (also known as "max-outs"), which specifies the number of DVDs in the mailing process on any given day. It is important to note that because of the nontrivial mailing time, the real consumption constraint facing consumers is the number of movies immediately available (referred to as the available quota). We briefly discuss the relationship between the instantaneous and available quota as follows. Let IQT_{ijt} be the instantaneous quota of plan j chosen by consumer i, and let A_{ijt} be the movies *available* to her on day t. The instantaneous quota is the total number of movies in the mailing process, which consists of A_{ijt} and the movies in transit that are not available for immediate consumption, denoted as T_{ijt}

$$\begin{gathered} IQT_{ijt} = A_{ijt} + T_{ijt}, \\ A_{ijt}, T_{ijt} \geq 0. \end{gathered} \tag{3}$$

Equation (3) implies that A_{ijt} is nonnegative and will never exceed IQT_{ijt}. Conditional on the same IQT_{ijt}, A_{ijt} is larger (smaller) if the number of movies in the mailing process (T_{ijt}) is smaller (larger). The detailed shipping dates allow us to compute A_{ijt} for each day the consumer stays with the company. Furthermore, we can impute A_{ijt} for any hypothetical plan j based on the consumption needs and the exogenous shipping policy of the company, i.e., after receiving the returned DVDs, the company sends the consumer the same number of DVDs. Figures A.1 and A.2 in the appendix provide a simple illustration of the relationship between IQT_{ijt} and A_{ijt}.

3.2.2. Expected Consumption. The consumer's consumption utility comes from the realized part of expected consumption, which is determined by the chosen plan and movie availability. We model the consumption outcome at the daily level to be consistent with the design of the instantaneous quota. We define c_{ijt}, the realized daily consumption, given the latent consumption need c^*_{it} and plan choice j

$$c_{ijt} = \begin{cases} 0, & \text{if } c^*_{it} \leq 0, \\ c^*_{it}, & \text{if } A_{ijt} > c^*_{it} > 0, \\ A_{ijt}, & \text{if } c^*_{it} \geq A_{ijt}. \end{cases} \tag{4}$$

Equation (4) states that realized consumption is left censored at 0 (realized consumption must be nonnegative) and right censored at A_{ijt} (realized consumption cannot exceed the expected available movies).

Combining the definition of realized consumption, conditional on plan choice j (Equation (4)) and the distributional assumption made regarding consumption decisions, we can write the expected consumption, $E[c_{ijt}]$

$$E[c_{ijt}] = (1 - \pi_{it}) \cdot \left[\sum_{k=0}^{A_{ijt}-1} \frac{\lambda_{it}^k e^{-\lambda_{it}} (k - A_{ijt})}{k!} + A_{ijt} \right]. \tag{5}$$

Using Equation (5), it is straightforward to show that expected consumption increases with the probability of having a positive consumption need $(1-\pi_{it})$, the magnitude of consumption needs (λ_{it}), and the number of DVDs available for consumption (A_{ijt}).

3.2.3. Expected Stockout. Note that when a consumer's consumption needs exceed the number of movies available, she experiences stockout. The stockout cost can be thought of as the consumer's disutility from her consumption "fall[ing] short of the desired amount" (Erdem et al. 2003, p. 16). Broadly speaking, the magnitude of such costs should depend on the specific service category—it is low if the service is nonessential and there are many substitutes, and high otherwise. A priori, we expect the stockout cost to be large for the focal service (content-edited movies) because watching such movies can be an inexpensive, but not easily substitutable pastime for the entire family; such costs should be especially large for consumers with a strong preference for "movie night" (watching multiple movies on one occasion) because the disutility from not being able to do so is high.

Unlike 3PT, in which consumers are not constrained in their consumption, consumers of a BBPD service must evaluate the chance of stockout situations associated with each plan j. The expected value of the stockout plan j, $E[so_{ijt}]$, is

$$E[so_{ijt}] = (1-\pi_{it})\cdot\left[\sum_{k=A_{ijt}+1}^{\infty}\frac{\lambda_{it}^{k}e^{-\lambda_{it}}(k-A_{ijt})}{k!}\right]. \quad (6)$$

Using Equation (6), we can show that the expected stockout increases with the probability of having nonzero consumption and mean consumption needs, but decreases with the number of movies available for consumption.

Because consumers make plan choices at the monthly level, we need to compute the number of expected monthly stockouts by aggregating across all days in the month. More specifically, $E[SO_{ijm}] = \sum_{t=1}^{T_m} E[so_{ijt}]$, where T_m is the number of days in a month. Similarly, the monthly expected consumption realizations are given by

$$E[C_{ijm}] = \sum_{t=1}^{T_m} E[c_{ijt}]. \quad (7)$$

3.2.4. Purchase Decision and Utility. Let D_{ijm} represent consumer i's plan choice in month m

$$D_{ijm} = \begin{cases} 1 & \text{if consumer } i \text{ chooses plan } j \text{ in month } m, \\ 0 & \text{otherwise.} \end{cases} \quad (8)$$

We assume that the consumer makes plan choices based on the benefit from consumption, C_{ijm}, the monetary costs of the subscription price of the chosen plan, P_{ijm}, and the nonmonetary stockout cost. The consumer also incurs a switching cost if she switches to a different plan or defects. We assume that the utility function can be approximated by a multiattribute, additive compensatory utility model (Lancaster 1966)

$$U_{ijm} = \beta_{0i} + \beta_{1i}C_{ijm} + \beta_{2i}SO_{ijm} + \beta_{3i}P_{ijm} + \beta_{4i}SW_{ijm} + \varepsilon_{ijm}. \quad (9)$$

The first term in (9), β_{0i}, measures the baseline utility of the subscription service (e.g., convenience of getting movies in the mail). The next two terms are related to usage. The second term, C_{ijm}, is the realized consumption for consumer i during month m, conditional on plan choice j. The third term SO_{ijm} is the stockout amount, included to capture the possibility that the amount of unmet consumption needs affects consumer plan choice. Both are plan specific, as determined by the quota of each plan. We also note that unlike consumer-packaged goods (e.g., ketchup) the BBPD service quota usually cannot be carried over to the next period. Consequently, both C_{ijm} and SO_{ijm} are affected only by the quota (plan) chosen for the current month. The next variable, P_{ijm}, is the price of the chosen plan. Although the listed price of each plan, P_j, is identical for all consumers, the actual paid prices, P_{ijm}, vary across consumers because the company occasionally offers small price discounts. In addition, a 6.6% sales tax is charged to those consumers who reside in the same state as the company. Thus, P_{ijm}, the actual price paid by consumer i for plan j in month m is

$$P_{ijm} = P_j - DSCT_{ijm} + TAX_i \times 0.066 \times P_j, \quad (10)$$

where P_j is the listed monthly price for plan j, $DSCT_{ijm}$ is the price discount received by consumer i in month m for plan j, and TAX_i, assumed to be known with certainty, is a dummy variable that equals 1 if the consumer pays sales tax, and 0 otherwise. Finally, we allow for possible switching costs, despite the fact that the focal company does not impose any explicit monetary penalty on switching or early termination. The inclusion of switching costs is based on two considerations. First, the automatic continuation of payments and the separation of payment and consumption occasions may lead to the "status-quo" bias, which is a psychological switching cost that has been found in previous papers (e.g., Goettler and Clay 2011), even where there is no explicit monetary penalty. Second, Zauberman (2003) shows that even a small switching cost can lead to persistent lock-in for the current choice. Thus, we include SW_{ijm} to allow for potential costs required to switch to another plan: it is a dummy variable that is 0 if consumer i chooses to stay with her current plan in month m, and 1 if she chooses a different plan, or when she drops out. We leave it to the data to show whether psychological switching cost is significant.

Regarding the coefficients, β_{1i} measures the unit benefit of movie consumption; β_{2i} represents consumer sensitivity to stockouts; β_{3i} measures price sensitivity; β_{4i} represents switching cost, incurred if the consumer switches to a different plan or leaves the company. Finally, ε_{ijm} represents the random errors related to choosing plan j, observable by the consumer, but not by the researcher.

As discussed above, given the advance purchase, we need to account for the uncertainty in consumers' consumption needs. Such uncertainty is inherent because of many factors, such as the amount of time available for consumption, all of which cannot be perfectly anticipated by the consumer. Thus, following the literature (e.g., Miravete 2002a, Narayanan et al. 2007, Lambrecht et al. 2007) we let consumers make plan-choice decisions based on their expected utility for plan j, given by

$$E[U_{ijm}] = \beta_{0i} + \beta_{1i}E[C_{ijm}] + \beta_{2i}E[SO_{ijm}] + \beta_{3i}P_{ijm} + \beta_{4i}SW_{ijm} + \varepsilon_{ijm},$$

where the expectations are taken over uncertainty in usage.

The utility of the outside option ($j = 0$) is specified as

$$U_{i0m} = \kappa_{1i} \cdot NFLXP_m + \kappa_{2i} \cdot BBL_m + \varepsilon_{i0m}, \tag{11}$$

where $NFLXP_m$ is the price of the most popular Netflix plan at month m, and BBL_m is a dummy variable that is 1 if Blockbuster Online is operating at month m, and 0 otherwise.[4] Thus, κ_{1i} and κ_{2i} measure the extent to which the outside option becomes more or less attractive in the presence of the two major competitors. Consumers are assumed to have no uncertainty regarding the utility of the outside option at month m; thus, $e[U_{i0m}] = \kappa_{1i} \cdot NFLXP_m + \kappa_{2i} \cdot BBL_m + \varepsilon_{i0m}$.

Given a consumer's rational expectations on her future consumption needs, each plan implies a stream of positive utilities from consumption and a stream of negative utilities from stockout and payment. When evaluating her current plan choices, as well as deciding on whether to switch plans, the consumer trades off the benefit of higher consumption and lower stockout with the higher costs of the subscription payment and switching. The decision calculus of a forward-looking consumer can be conceptualized as the following dynamic programming problem. At the beginning of time t, consumer i observes the state variables ($S_{im} = [Ngenre_{im}, PastConsump_{im}, AccConsump_{im}, D_{i,m-1}]$), where the first three terms can be summarized by the probability of nonzero consumption $1 - \pi$ and the mean of consumption needs λ. The consumer then forms expectations on her future consumption needs. When evaluating each plan, she takes into account the instantaneous quota that limits her daily consumption and estimates the amount of consumption that can be realized, and hence, the associated stockout risks. Consumers then form expectations about the expected utility, $E[U_{ijm}]$, and choose a plan that maximizes the discounted value of a stream of utilities defined by consumption, stockout, payment, and switching cost.

The consumer intertemporally trades off among the different utility components so as to maximize long-term utility. Specifically, she recognizes that the high switching cost also implies that a purchase decision now may become a commitment to the same service plan in the future. For example, the consumer can enjoy a higher number of movies consumed and avoid stockout situations from upgrading her plan; however, the extra utility comes at a cost: she needs to incur an additional switching cost (mental costs from remembering to do so, time spent signing in and making the plan change) and a larger payment. She will not upgrade if the downside dominates the upside. For example, to avoid the future disutility associated with stockout and to avoid future switching costs, the consumer may be willing to stay with a high plan, even at the cost of a larger financial payment.

Formally, we model the consumer as a forward-looking decision maker who makes advance plan choices to maximize her total discounted future expected utility over an infinite horizon

$$\max_{D_{ijm}} \left\{ E\left[\sum_{\tau=m}^{\infty} \delta^{\tau-m} U_{ij\tau} \right] \right\}, \tag{12}$$

where $U_{ij\tau}$ is the single-period utility function, and δ is the discounting factor that measures the trade-off between current and future expected utilities. Because we do not have the exclusion restriction as in Chung et al. (2014), we follow the convention (e.g., Erdem and Keane 1996, Kopalle et al. 2012, Sun 2005) and fix the monthly discount factor at 0.98.

The Bellman equation of the consumer is given by

$$V_{im}(S_{im}) = \max_{D_{ijm}} \{V_{ijm}(S_{im})\} = \max_{D_{ijm}} \{E[U_{ijm} \mid S_{im}] + \delta E[V_{i,m+1}(S_{i,m+1})]\}, \tag{13}$$

[4] Arguably, Netflix and Blockbuster Online are the two major competitors of the focal company during the observation period. Ideally, we would use the sales figures for Blockbuster Online, which was launched in August 2004. However, for our observation period, only the total Blockbuster revenue from both online and offline channels was available at the quarterly level. Although Bloomberg now records Blockbuster Online revenue, they started doing so only after the second quarter of 2006.

where S_{im} denotes the set of state variables, including the probability of zero consumption needs (π), the mean consumption needs (λ), and the consumer's previous plan choice, $D_{ij,m-1}$ $(j = 1, \ldots, J)$. The state transition for π and λ are estimated from the data. The optimal plan choice is given by

$$D^*_{ijm} = \arg\max_{D_{ijm}} \left\{ \sum_{j=0}^{J} D_{ijm} V_{ijm}(S_{im}) \right\}. \tag{14}$$

3.3. Heterogeneity and Estimation

To control for unobserved heterogeneity, we employ the latent-class approach (Kamakura and Russell 1989). Suppose there are N distinct latent segments, and each consumer has a probability $q(n)$ of belonging to segment n. Then, the vector

$$\Theta = \big[\theta_0(n), \theta_1(n), \theta_2(n), \alpha_0(n), \ldots, \alpha_5(n), \beta_0(n), \ldots, \beta_4(n), \kappa_1(n), \kappa_2(n), q(n)\big]$$

represents all parameters to be estimated and the segment size $\pi(n)$ for all n.

Define $V^*_{ijm} = V_{ijm} - \varepsilon_{ijm}$ as the deterministic part of the value function. The error term ε_{ijm} captures the unobservables affecting plan utilities and is assumed to be independently and identically extreme-value distributed. We obtain the probability of consumer i choosing plan j in month m in the familiar multinomial logit formula

$$\text{Prob}(D_{ijm} = 1 \mid \Theta) = \sum_{n=1}^{N} \pi(n) \frac{e^{V^*_{ijm}}(n)}{\sum_{j=0}^{J} e^{V^*_{ijm}(n)}}. \tag{15}$$

Our calibration sample consists of 800 randomly selected consumers, and the holdout sample contains an additional 800 consumers. We use simulated maximum likelihood (Keane 1993, McFadden 1989) for the estimation. Because some of the state variables in Equation (13) are continuous, we encounter the problem of a large state space. We first discretize the continuous spaces and then adopt the Keane and Wolpin (1994) interpolation method to calculate the value functions for a few state-space points, which we then use to estimate the coefficients of an interpolation regression. The interpolation regression function provides values for the expected maxima at any other state points for which values are needed in the backward-recursion solution process.

3.4. Identification

Identification for the Zero-Inflated Poisson Model. The ZIP model has two parts, characterized by Equations (1), (2), and (4). We discuss intuitions on how the parameters in these three equations are identified separately. Equation (1) implies that on any given day, there is a positive probability that the consumer has no need for movies. Equation (2) implies that conditional on having a consumption need, the number of movies that the consumer would like to watch is characterized by a Poisson distribution. Equation (4) suggests that actual consumption needs may be capped by the number of movies available. The identification of parameters in Equation (2) uses observations with a positive number of available movies. This identification depends on the variation in the explanatory variables included (e.g., recent consumption rate and accumulated consumption). The parameters in Equation (1) (e.g., θ_1 and θ_2) are identified by occasions where movies were available, but the consumer chose not to watch any. For example, consider a consumer with the same number of DVDs available, but on different days of the week. The different frequencies of *excessive* zero consumption (i.e., relative to what can be explained by a standard Poisson model) on weekdays versus weekends help to identify θ_1.

Identification for the Plan Utility Model. Identification of β_1 and β_2 are achieved by variations in expected consumption and stockout, both of which arise from the interaction between the quota and time-varying consumption needs. Identification of the price coefficient, β_3 is based on the variation in effective prices, which varies both across consumers in the same segment (differences in the tax amount between in-state and out-of-state consumers) and within consumers (discounts). Identification of the switching cost, β_4, is based on the observed switching patterns in plan choices. In particular, the identification is facilitated by the consumer's likelihood of staying with the service in months when the realized consumption is low (i.e., when both consumption needs and stockout costs are close to zero).

4. Results

4.1. Model Comparison

To see whether incorporating consumption uncertainty, switching costs, stockout risk, and forward-looking better explains the observed dynamic consumer purchase decisions, we estimate five models. In the first benchmark model (Model 1), we assume that there is no uncertainty regarding the expected usage or stockout, and consumers' plan-specific utility includes the instantaneous quota and price. This is very similar to most existing models used to study consumer plan choices in the telecommunications industry (e.g., Danaher 2002). The second benchmark model (Model 2) introduces uncertainty by incorporating expected consumption without the expected stockout cost. By introducing uncertainty, this model recognizes that consumers make their advance-purchase decisions with imperfect information. This model is similar to Lambrecht et al. (2007); however, it does not account for the

Table 4 Model Comparisons

	Model 1	Model 2	Model 3	Model 4	Model 5
Calibration sample					
Log-likelihood	−48,241.5	−47,111.9	−46,516.0	−44,995.8	−44,541.8
AIC	96,541.0	94,281.8	93,094.1	90,057.7	89,149.6
BIC	96,826.0	94,566.9	93,398.8	90,382.1	89,474.0
Hit rates (%)	84.12	85.61	86.60	90.44	91.60
Holdout sample					
Log-likelihood	−49,254.1	−48,073.4	−47,597.2	−45,909.4	−45,069.2
AIC	98,566.1	96,204.9	95,256.4	91,884.8	90,204.4
BIC	98,850.9	96,489.7	95,560.9	92,208.9	90,528.5
Hit rates (%)	83.60	85.12	85.72	89.43	90.52

stockout risk, which is conceptually unique to BBPD. The third benchmark model (Model 3) recognizes both expected consumption and stockout, but does not include the switching cost. The fourth benchmark model (Model 4) extends Model 3 by including switching costs. In Models 1–4, consumers are assumed to be myopic rather than forward looking, such that they maximize their current purchase utilities without accounting for future switching costs. The fifth model (Model 5) is our proposed model with usage uncertainty, stockout, switching costs, and forward-looking consumers.

As we use the latent-class approach to account for consumer heterogeneity, we must determine how many segments best fit the data for each of the five models. We estimate each of the competing models with various segments ($N = 1$, 2, and 3). The results suggest that Models 1 and 2 with three segments and Models 3–5 with two segments are the best fits. For example, the log-likelihood, Akaike information criterion (AIC) and Bayesian information criterion (BIC) of our proposed model are $-44{,}754.5$, 89,543.1, and 89,710.2, respectively, for one segment, $-44{,}541.8$, 89,149.6, and 89,474.0 for two segments, and $-44{,}467.2$, 89,182.6, and 89,614.0 for three segments.

In Table 4, we report the log-likelihood, *AIC*, *BIC*, and hit rates of the five competing models with the model-specific optimal numbers of segments. Model 1 has the worst fit ($LL = -48{,}241.5$, $AIC = 96{,}541.0$, $BIC = 96{,}826.0$). This result is not surprising, given that the instantaneous quota (IQT_j) is the same across all consumers and across time. Using it as a static explanatory variable fails to capture the individual-level dynamics in consumption needs. Replacing the instantaneous quota with the individual-specific and time-varying $E[C_{ijm}]$ results in a significant improvement in the model fit of the second baseline model ($LL = -47{,}111.9$, $AIC = 94{,}281.8$, $BIC = 94{,}566.9$). Model 3 further improves the model fit by including the expected stockout $E[SO_{ijm}]$, which confirms the importance of capturing consumers' high disutility for stockout situations. Model 4 incorporates switching costs, which helps explain the quite persistent patterns of choices observed in the data. The model fit increases significantly, confirming the significance of switching costs in our empirical context. Model 4 assumes that consumers are myopic and oblivious to future lock-in by setting the discount factor to 0. Model 5 differs from Model 4 by allowing the consumer to be forward looking. The improvement of model fit from Models 4 to 5 suggests that a forward-looking model better explains consumer plan choices than its myopic counterpart. Finally, the proposed model significantly outperforms all four benchmark models, with the highest hit rates for both the calibration sample (91.6%) and the holdout sample (90.3%). In addition, we test alternative discount factors for the dynamic model.[5] The results confirm the robustness of the main model.

To sum, model comparisons show that it is important to allow for uncertainty in consumption needs, stockout, and switching costs, and forward-looking consumers in modeling consumer purchase decisions under BBPD. The comparisons also suggest that adding switching costs contributes the most to improving the data fit, followed by incorporating expectations of future consumption needs, expected stockout, and forward-looking consumers. Because Model 5 is the best-fitting model, our subsequent discussions focus on it.

4.2. Parameter Estimates

The latent-class estimates indicate that 86.4% of consumers belong to the first segment, and 13.6% belong to the second segment.[6] We start with examining the parameter estimates and t-statistics for the zero-inflation equation, shown in the upper part of Table 5. We find that both segments are much less likely to

[5] We tested two models with the discount factors 0.99 and 0.95. Both models fit less well, compared to the model with the assumed discount factor of 0.98: ($LL_{\delta=0.99} = -44{,}566.2$, $BIC_{\delta-0.99} = 89{,}522.8$; $LL_{\delta=0.95} = -44{,}654.3$, $BIC_{\delta=0.95} = 89{,}699.0$).

[6] We follow Kamakura and Russell (1989) and assign consumers into each of the two segments, using 0.5 as the cutoff probability (e.g., Bucklin and Gupta 1992).

Table 5 Estimation Results of the Proposed Model

Parameters		Convenience segment Coefficient	Convenience segment t-value	Value-seeking segment Coefficient	Value-seeking segment t-value
		Consumption needs—zero-inflation model			
Constant	θ_{0i}	−0.496**	−9.88	−0.876**	−9.74
Weekend	θ_{1i}	−0.394**	17.07	−0.218**	−9.40
Number of genres	θ_{2i}	−0.017**	−4.00	−0.031**	−2.46
		Consumption needs—Poisson model			
Constant	α_{0i}	−1.096**	−50.78	−1.128**	32.45
Weekend	α_{1i}	0.453**	21.27	0.246**	15.44
Lag consumption week 1	α_{2i}	−0.335**	−29.89	0.095**	6.12
Lag consumption week 2	α_{3i}	0.152**	15.91	0.208**	15.29
Lag consumption week 3	α_{4i}	0.188**	19.27	0.179**	12.82
Accumulated consumption	α_{5i}	−0.0016**	−33.66	0.001**	5.50
		Plan-utility equation			
Constant	β_{0i}	2.225*	2.84	1.53**	3.25
Expected consumption	β_{1i}	0.461**	4.35	0.672**	5.14
Stockout amount	β_{2i}	−2.421**	−37.10	−1.813**	−15.93
Price	β_{3i}	−0.293**	−34.08	−0.354**	−51.16
Switching cost	β_{4i}	−7.360**	−9.80	−3.248**	−5.97
Netflix price	κ_{1i}	−0.02**	−3.28	0.03	1.06
Blockbuster dummy	κ_{2i}	0.14	0.77	0.10	0.62
Estimated segment size (%)		86.4		13.6	

$^{*}p < 0.05$; $^{**}p < 0.01$.

have zero consumption needs on weekends. Notably, the weekend effect is much higher for the first segment ($\theta_1 = -0.39$ versus -0.22). As expected, consumers who watch more movie genres are less likely to have zero consumption needs ($\theta_2 = -0.017$ and -0.031). For the Poisson model (middle part of Table 5), the weekend effects are positive for both segments ($\alpha_1 = 0.45$ and 0.25). Combined with previous estimates of the weekend variable in the zero-inflation equation, consumers in the first segment are more likely to cluster their consumption at convenient times—i.e., weekends—whereas consumers in the second segment are more flexible in planning their consumption over weekdays. Based on this finding and for the ease of exposition, we refer to the first segment as *convenience* and the second segment as *value seeking*. The coefficients of average realized consumption in the past weeks are positive and significant for the second segment, indicating strong inertia in consumption needs. Interestingly, the effect of accumulated consumption is different for the two segments. For the convenience segment, accumulated consumption has a small, yet significantly negative effect on consumption needs ($\alpha_5 = -0.002$). A possible explanation for this "fatigue" effect is that cumulative consumption may exhaust the consumer choice set, and hence reduces the need for future consumption. By contrast, there is a significant positive, "reinforcing" effect of accumulated consumption for the value-seeking segment ($\alpha_5 = 0.001$). A likely explanation is that value-seeking customers are more variety seeking for different movie genres. A further check confirmed that, consistent with this explanation, the value-seeking segment watched an average of 6.0 genres throughout their tenure, significantly more than the 4.1 viewed by the convenience segment.

We next turn to the parameter estimates in the expected plan-utility equation, shown in the lower part of Table 5. First, the Blockbuster dummy (BBL_m) does not have a significant effect on outside utility; and the price of Netflix has a small and marginally significant effect for the convenience segment ($\kappa_1 = -0.02$, $t = -3.28$). A possible explanation is that by operating in the niche market of content-edited movies, the focal company was able to avoid direct competition with Blockbuster, but less so with Netflix. As we expected, the consumption-benefit coefficients are positive and significant for both segments ($\beta_1 = 0.461$ and 0.672). Stockout coefficients are significant and negative for both segments, which means that consumers are sensitive to the negative utility caused by consumption needs capped by the instantaneous quota. We emphasize that caution should be exercised in interpreting the stockout coefficient, given that we assume rational expectations for consumption needs. Such a coefficient may be biased if consumers have dynamic inconsistency in their consumption needs (e.g., Laibson 1997, Milkman et al. 2009, Read et al. 1999) and cannot foresee such inconsistency. For example, at the time of purchase, if a myopic consumer overestimates (underestimates) the expected stockout (compared with the actual stockout), the stockout coefficient will be biased downward (upward).

Switching costs are negative and statistically significant for both segments ($\beta_4 = -7.36$, $t = -9.80$; $\beta_4 = -3.25$, $t = -5.97$). These switching costs are also economically significant: the monetized switching cost is about \$25 for consumers in the first segment and \$10 for those in the second segment. Such a magnitude of the switching cost is interesting, considering that the focal company does not charge any fee for plan changes or service termination. However, the auto-payment mechanism of the company may have induced substantial psychological and transactional costs favoring the company. The relatively higher switching cost for the convenience segment is likely due to their higher opportunity cost of time, which is consistent with the fact that their movie consumption is more likely to occur on weekends rather than weekdays.

Further comparisons of the two segments' coefficients reveal more interesting differences. Compared with the value-seeking segment, consumers in the convenience segment have a higher intrinsic preference for the online DVD rental service (2.23 versus 1.53), are more averse to stockout (−2.42 versus −1.81), and are less price sensitive (−0.29 versus −0.35).

4.3. Drivers and Consequences of Overpurchase

In this section, we focus on the drivers of overpurchase, which is a unique phenomenon under BBPD. We first show how overpurchase is related to stockout and switching costs (§§4.3.1 and 4.3.2). We then present evidence on an inverse relationship between overpurchase and customer retention (§4.3.3). This last finding suggests a trade-off between overpurchase and retention, such that a large overpurchase may not be beneficial for the company; it also provides a motivation for counterfactual exercises, where the company may use alternative pricing decisions. In a separate exercise not included in the manuscript, but available on request, we also find that the company can fine-tune its pricing strategy to achieve better profitability.

4.3.1. Drivers of Overpurchase. Stockout is a salient feature of BBPD that arises because of the constraint of the instantaneous quota. The magnitude of the stockout cost estimates suggests that consumers have a strong incentive to overpurchase as a way of avoiding the disutility of stockout. Unlike consumer-packaged goods (e.g., Ailawadi and Neslin 1998, Erdem et al. 2003, Sun et al. 2003), the consumer cannot stockpile. Consequently, the consumer cannot decrease stockout risk unless she purchases higher plans. Switching costs may also increase the overpurchase, because high switching costs may prevent the consumer's timely plan-choice adjustments in the presence of dynamic consumption needs. In this simulation, we compute the overpurchase rate as the percentage of the unused purchased quota, averaged across consumers and months, where the overpurchase for a representative consumer-month (i, m) is $\sum_{t=1}^{T_m}(A_{ijt} - C_{ijt}) / \sum_{t=1}^{T_m} A_{ijt}$.

Figure 3 (Color online) Overpurchase, Stockout, and Switching Costs

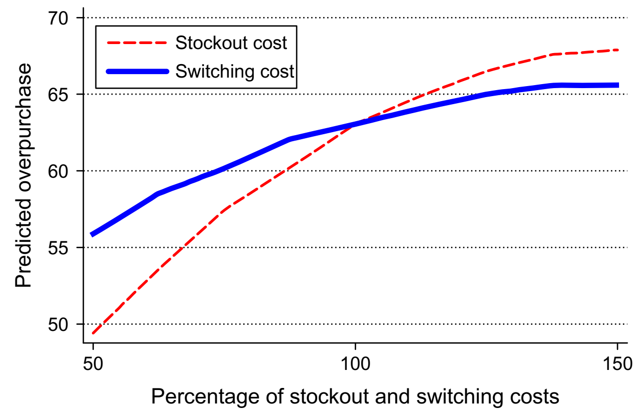


Figure 3 plots the average overpurchase rate against the different magnitudes of the stockout and switching costs. On the horizontal axis, each type of cost is varied from 50% to 150% of its respective parameter estimates, based on real data. We find that, consistent with the intuition discussed above, the overpurchase increases with both stockout and switching costs. More specifically, increasing the stockout cost by 25% leads to a 6% increase in the overpurchase and a 25% increase in the switching cost increases the overpurchase by 4%. At the current parameter estimates, the change in the overpurchase is more sensitive to the change in the stockout costs. This indicates that the main reason behind consumers' sacrifice in paying for a high-quota plan is to ensure an adequate stream of consumption utilities.

4.3.2. Dynamics in Overpurchase. We next examine how the dynamics in overpurchase is related to switching costs. Intuitively, consumers with higher switching costs are more likely to experience persistent overpurchases, especially after the consumers' consumption needs have declined. This intuition is confirmed in Figures 4(a) and 4(b), which show the distributions of overpurchase for the convenience segment (high switching costs) and value-seeking segments (low switching costs) at two different points in time: the initial month and after the 20th month. Whereas the amount of overpurchase of the two segments does not differ much in the initial month, the overpurchase is significantly higher after the 20th month for the convenience segment, indicating that consumers with high switching costs are more susceptible to lock-in situations. Overall, the convenience

Figure 4(a) (Color online) Overpurchase (by Segment) in the First Month

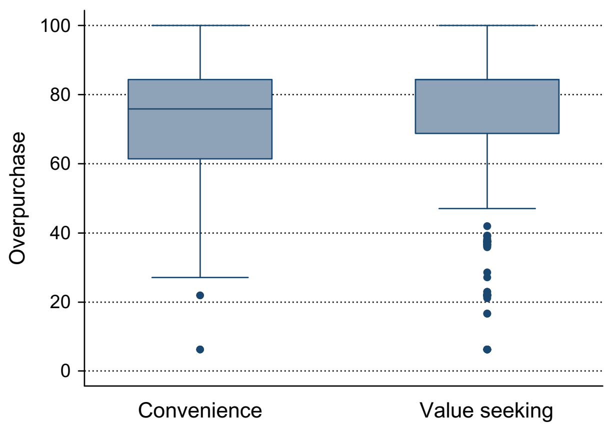
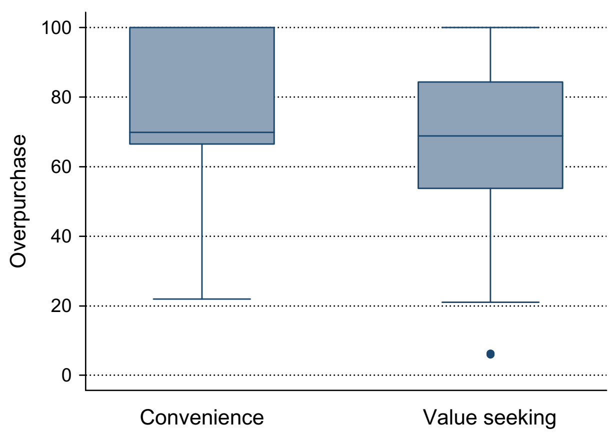


Figure 4(b) (Color online) Overpurchase (by Segment) After the 20th Month



segment incurred a larger overpurchase during its entire lifetime than the value-seeking segment (74.1% versus 65.9%).

Figure 5(a) presents a more complete picture of the overpurchase dynamics by segment. The significant downward trend of overpurchase for the value-seeking segment can be attributed to the interplay between the "reinforcing" effect of accumulated consumption on consumption needs ($\alpha_5 = 0.001$) and the relatively lower switching cost ($\beta_4 = -3.25$), which allows consumers in the value-seeking segment to adjust their plan choices and align their purchased consumption capacity with their evolving consumption needs. By contrast, the overpurchase of the convenience segment exhibits a significant and increasing trend over time. This is due to both the "fatigue" effect of accumulated consumption on consumption needs ($\alpha_5 = -0.002$) and the fact that consumers in the convenience segment do not adjust plan purchases frequently enough because of their higher switching costs ($\beta_4 = -7.36$).

Interestingly, during the early stage of their tenure, the convenience segment overpurchases *less* than the value-seeking segment (67.1% versus 71.2%). Such patterns seem puzzling at first, but can be rationalized by forward-looking consumers. On one hand, lock-in arises because of the consumers' significant switching costs and their tendency to delay switching (switching costs incurred in the distant future are discounted more than switching costs in the immediate future). On the other hand, by anticipating a higher probability of lock-in for the future, it is optimal for consumers with higher switching costs to choose a conservative (or lower) plan at the beginning of their subscription to avoid too much overpurchase from future lock-in. By contrast, consumers with lower switching costs are less susceptible to future lock-in; as a result, they are likely to start with a higher plan and are able to avoid excess overpurchase by adjusting their plan choices later.

4.3.3. Overpurchase and Attrition. Although a larger overpurchase implies a higher per-period profit for the firm, the overall profitability also depends on consumer retention. Focusing now on consumers'

Figure 5(a) (Color online) Dynamics in Overpurchase (by Segment)

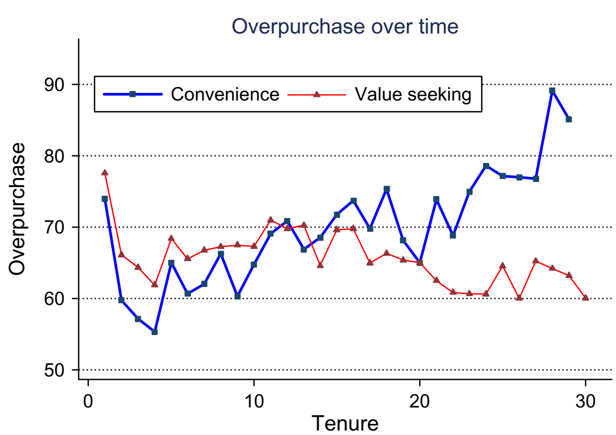
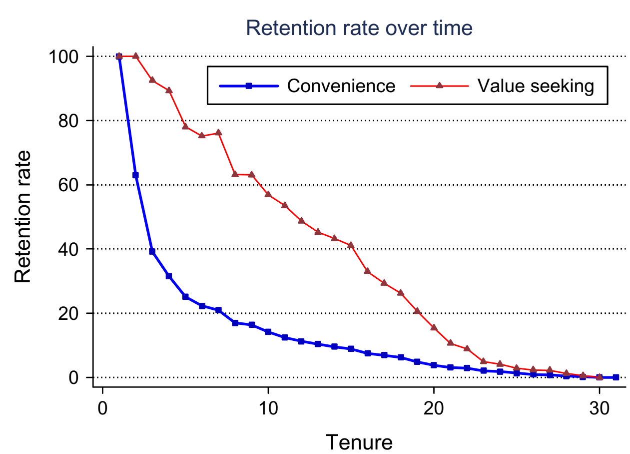


Figure 5(b) (Color online) Dynamics in Retention Rate (by Segment)



dropout decisions, we found that convenience segment switchers are more likely to leave the company early. As Figure 5(b) shows, the retention probability of the convenience segment is significantly lower than that of the value-seeking segment over time. An intuitive explanation is that the convenience segment values consumption less, and their consumption is more easily satiated. Given the increasing mismatch of the value and cost of payment, and the unwillingness to incur additional switching costs to change to lower-level plans, consumers in the convenience segment are more likely to end the lock-in by directly dropping out. By contrast, the higher retention rate of the value-seeking segment can be attributed to both higher consumption inertia and their ability to switch down.

In summary, the above analyses show that consumers with high switching costs are much less likely to adjust plan decisions over time to reduce overpayment. Furthermore, consumers with higher switching costs are also less likely to start with high-quota plans and are more likely to drop out sooner.

4.4. Understanding the Roles of the Instantaneous Quota

The two unique features of BBPD are the instantaneous quota and the stockout risk it induces. Two interesting questions to consider are the following: First, should the company stay with an instantaneous quota or use a monthly quota? Second, if the company allows the consumer to pay a marginal fee to cover the stockout risk, would the extra gain in marginal fees be enough to compensate for the (expected) loss in fixed fees? To explore these two questions, we use Monte Carlo simulations to compare the expected profits in three cases. Case 1 is the "standard" BBPD with the instantaneous quota. Case 2 is BBPD with the monthly quota, the standard practice of other nonlinear price formats, such as 3PT. Case 3 approximates 3PT in two ways: a monthly quota and the option for the consumer to pay a marginal fee for consumption in excess of the monthly quota.[7] Case 3 is different from the first two cases, in that the stockout risk of the consumer is eliminated. The simulations are based on the 800 consumers in the estimation sample. Next, we discuss the operationalization of the simulation exercises and the results.

Case 2. *Monthly quota with stockout risk.* We now explore whether the company would benefit from replacing the instantaneous quota with an equivalent monthly quota. Intuitively, a monthly quota would increase the consumer's flexibility in fulfilling her consumption needs because she can pick and choose when to consume to maximize her consumption needs and to reduce the stockout risk. As the simple examples in Figures A.1 and A.2 of the appendix illustrate the instantaneous quota "penalizes" peak consumption needs, so that for the same monthly consumption needs, consumers with high peak consumption needs are more likely to incur stockout. Thus, based on the estimated consumption needs, we let the consumer solve a new constrained optimization problem, with the equivalent monthly quota as the new binding constraint. Assuming that consumers take one day to watch a movie, the equivalent monthly quota of plan j, MQ_j, is approximated as follows: $MQ_j = IQT_j \times (\text{Number of business days in a month}/(1+T))$, where T is the number of days required for two-way shipping.

With the monthly quota, the utility function is specified as

$$E[U_{ijm}] = \beta_{0i} + \beta_{1i} E[C_{ijm} \mid MQ_j] + \beta_{2i} \cdot E[SO_{ijm} \mid MQ_j] + \beta_{3i} P_{ijm} + \beta_{4i} SW_{ijm} + \varepsilon_{ijm}. \tag{16}$$

Equation (16) differs from Equation (9) in that a daily level quota is replaced with an equivalent monthly quota (MQ_j). This change gives the consumer higher flexibility in consumption, manifested in plan-specific realized consumption $E[C_{ijm} \mid MQ_j]$ and consumption over the quota $E[SO_{ijm} \mid MQ_j]$, accounting for the change to the monthly quota. Figure A.3 of the appendix provides a simple illustration. Note that (16) is still BBPD, because it does not allow the consumer to cover the stockout with a fee.

Revenue, Cost, and Profit for Cases 1 *and* 2. These marketing outcomes for Cases 1 and 2 are computed in the same way. For each of these consumers, we iteratively simulate purchase decisions made at the beginning of every month. More specifically, at the beginning of each month, we compute the daily level expected consumption and stockout for the consumer and then aggregate to the monthly level. Using these quantities, as well as the parameter estimates, we compute the expected utilities for each of the available plans. We then apply the multilogit formula to compute the probability of choosing each of the available plans,

[7] We emphasize that Case 3 is an approximation of 3PT, with the assumption that the consumer is only concerned about maximizing consumption utility, subject to the budget constraint (Lambrecht et al. 2007). Ascarza et al. (2012) found evidence that consumers also value the "free" units of 3PT, above and beyond its effect on the budget constraint. To ensure comparability of the three scenarios, we make the following assumptions: First, the consumer's latent consumption needs are the same in all scenarios. Second, we assume that the structural parameters in the utility function, such as the utility from consumption (β_{1i}), the price sensitivity (β_{2i}), and the switching cost (β_{3i}) are identical for both price formats. Third, all three scenarios have the same operational characteristics. Specifically, the consumer must wait for the same mailing time to receive the new rental products. Consequently, the marginal price for the approximate 3PT is charged for movies that arrive in the mail, not for immediate viewing.

as well as the outside option. These choice probabilities allow us to simulate the plan choice (or leave the service). After a plan choice is drawn, we simulate the daily level consumption, subject to the quota constraints of the simulated plan choice, until we come to the next payment period. We continue this procedure until the outside option (i.e., dropping out) is drawn, at which point the consumer's tenure, total revenue, cost, and profit are recorded; then, we move on to the next consumer. For each consumer, tenure is counted as the total number of months she chooses to stay with the company. Total revenue is determined by the consumer's simulated sequence of plans and payments: it is completely driven by the demand for the quota, and is not related to the actual consumption. The total cost is computed based on the realized consumption, conditional on realized consumers' consumption needs and the chosen plan's quota. Total profit is simply the difference between the total revenue and total cost. These quantities are summarized below

$$Rev = \sum_{i=1}^{I}\sum_{m=1}^{M_i}\delta^m\sum_{j=1}^{J}D_{ijm}P_j,$$

$$Cost = \sum_{i=1}^{I}\sum_{m=1}^{M_i}\delta^m\sum_{j=1}^{J}D_{ijm}E[C_{ijm}]\cdot mc,$$

$$Profit = Rev - Cost,$$

where mc is the marginal cost for each DVD consumed, approximated by the two-way postage fee of \$0.90 and an estimated handling fee of \$1.10 per DVD. In these equations, δ is the discount factor that the company uses to discount profits accruing from future sales and is set to be the same as in the consumer model.

Case 3. *Monthly quota with marginal fee.* In this case, the company uses both the monthly quota, as well as a marginal fee, for consumption over the quota. With both changes, the price format closely resembles a three-part tariff. A representative service plan j consists of the same fixed fee (P_j), and monthly quota (MQ_j) as in Case 2, and a marginal fee (r). The marginal fee r is set at \$2.99, the same as the rental price Blockbuster charged for newly released movies.

Consumer i's expected monthly utility for plan j is

$$E[U_{ijm}] = \beta_{0i} + \beta_{1i}E[C_{ijm}\mid MQ_j] + \beta_{3i}P_{ijm} + \beta_{3i}\cdot rE[SO_{ijm}\mid MQ_j] + \beta_{4i}SW_{ijm} + \varepsilon_{ijm}. \quad (17)$$

Apparently, the key difference between (16) and (17) is that the disutility of stockout ($\beta_{2i}\cdot E[SO_{ijm}\mid MQ_j]$) is replaced with the marginal fee ($\beta_{3i}\cdot r\cdot E[SO_{ijm}\mid MQ_j]$). For Case 3, the computations of the company's total revenue and cost are also different. In particular, the total revenue for 3PT consists of both the fixed fee and the marginal revenue collected from the avoided stockout (the difference between the realized consumption and the quota of the chosen plan). The additional consumption is also added to the company's operating cost. The revenue, costs, and profits are

$$Rev = \sum_{i=1}^{I}\sum_{m=1}^{M_i}\delta^m\sum_{j=1}^{J}D_{ijm}[P_j + E[SO_{ijm}]\cdot p_r],$$

$$Cost = \sum_{i=1}^{I}\sum_{m=1}^{M_i}\delta^m\sum_{j=1}^{J}D_{ijm}[E[C_{ijm}] + E[SO_{ijm}]]\cdot mc,$$

$$Profit = Rev - Cost.$$

We compare the marketing outcomes of Cases 1–3 in three ways: tenure (the total number of months consumers stay with the company), total costs, and total revenue. Table 6(a) shows the comparison between Cases 1 and 2. First, if the company uses a monthly quota, the simulated tenure increases by 9.0%: consumers are also likely to stay with the company longer because of increased flexibility. Second, a monthly quota allows consumers to significantly increase their consumption rates, which translates to an 11.2% increase in the company's operating costs. This is expected because consumers' consumption is less constrained by the monthly quota. Third, the total revenue also decreases by 1.3%. This is because with the monthly quota, consumers with large peak consumption are less constrained by the instantaneous quota, and thus have less incentive to purchase higher plans. Overall, a switch to the monthly quota would reduce the total profit by 3.2%.

Table 6(b) compares Cases 1 and 3. With Case 3, the simulated average tenure increases modestly by 5.0%, and the average costs increase by 9.0%. The additional revenue collected from the marginal fee in Case 3 (\$25.4 per consumer) is equivalent to 5.3%

Table 6(a) Comparison Between Case 1 and Case 2

	Case 1	Case 2	Percentage difference
Tenure (months)	7.9	8.6	9.0
Cost (\$)	65.7	73.1	11.2
Revenue (\$)	475.6	469.8	−1.3
Total profit (\$)	409.8	396.7	−3.2

Table 6(b) Comparison Between Case 1 and Case 3

	Case 1	Case 3	Percentage difference
Tenure (months)	7.9	8.3	5.0
Cost (\$)	65.7	71.6	9.0
Revenue from fixed part (\$)	475.6	432.4	−9.1
Revenue from marginal part (\$)	NA	25.4	NA
Total profit (\$)	409.8	386.3	−5.7

of the total revenue (which consists purely of subscription fees) in Case 1. Third, the fixed fee that the company expects to collect from the fixed subscription fee is 9.1% less compared with BBPD. Overall, the net profit decreases by 5.7% if the company uses a monthly quota and gives the consumer the option of paying the marginal fee. The intuition is that by allowing consumers to go over the quota at the cost of a marginal fee, Case 3 also effectively eliminates consumers' stockout risks, further reducing their incentive to purchase high-level plans, and the company's ability to extract revenue from the consumer diminishes significantly. Indeed, given the high disutility from stockout (~\$8), BBPD induces customers to pay higher fixed subscription fees and incur a high overpurchase.[8] Finally, we should note that although Case 3 is conceptually similar to a three-part tariff, we should not interpret the results as conclusive evidence on the aggregate optimality of BBPD over 3PT for several reasons. First, the subscription and marginal prices we assumed for Case 3 are not necessarily optimal for 3PT. Second, since it is impractical to estimate the marginal price coefficient, we made the strong assumption that it is the same as that for the fixed part. In practice, these are not necessarily the same (e.g., Ascarza et al. 2012). Third, we assumed exogenous consumption, characterized by a Poisson model. This modeling choice, however, does not fully capture the bunching of consumption around the quota, an interesting phenomenon that is expected to arise for 3PT (e.g., Lambrecht et al. 2007). Fourth, this counterfactual is based on the focal company's idiosyncratic plan-switching policy.

4.5. Alternative Ways to Improve the Design of BBPD

Having identified how the instantaneous quota of BBPD can help the company, we now investigate ways by which the company can fine-tune other aspects of its current BBPD design. The three counterfactuals we consider are informed by observed industry practices. First, the focal company may allow the consumer to switch more often than its current policy (once per month). The second counterfactual is motivated by Netflix's now well-known "throttling" practice of intentionally delaying the turnaround time for heavy (and less profitable) users. Because this tactic specifically targets heavy users, we assume that it is applied to the company's 10% heaviest users (based on the average monthly consumption rates). The third counterfactual is motivated by the fact that all BBPD services we know of are continuous subscription service: the customer lifetime value depends not only on per-period profit but also on retention. As demonstrated previously, overpurchase can be a double-edged sword. A high overpurchase increases the company's short-term profit (consumers overpay for the service and the cost of serving them is low); however, persistent overpurchases can also reduce the consumer retention rate. Thus, it is important for the focal company to optimally balance overpurchase, costs, and retention to optimize long-term profitability. Because overpurchase is readily observable, the company could actively monitor and manage it. A direct way to mitigate the downside of overpurchase is to give *targeted* price discounts to consumers who are observed as having excessive levels of overpurchase. The idea is based on the rationale that consumers who overpurchase more (because of the increasing gap between their reduced consumption needs and the quota) have higher risks of defection. Note that all of these counterfactuals can be readily implemented by the company.

For each counterfactual, we compare the simulated overpurchase, average tenure, and profit, summarizing the results in Table 7. We refrain from conducting segment-specific counterfactuals because it is less practical for the company to implement segment-specific marketing actions. By combining the simulated purchase and consumption and summing them over all of the months, the simulated customer i's lifetime overpurchases can be written as

$$\sum_{m=1}^{M_i} D_{ijm} \sum_{t=1}^{T_m} (A_{ijt} - E[c_{ijt}]) \Big/ \sum_{m=1}^{M_i} D_{ijm} \sum_{t=1}^{T_m} A_{ijt}. \tag{18}$$

Alternative Plan-Switching Policy. One of the insights from our model is that persistent overpurchase can be explained by the substantial switching costs identified from the data. Although such switching costs are attributed to consumers' psychological inertia, and not an explicit penalty imposed by the company, it is interesting to investigate a hypothetical situation where the company allows its subscribers to switch more frequently. We conducted a counterfactual exercise where the company adopts the switching policy of Netflix, the largest online movie rental company. As of February 2015, Netflix allows consumers to upgrade their plan choice on any day, but only allows them to downgrade at the beginning of the next billing cycle. Row (A) of Table 7 shows the results of this simulation. We find that a hypothetical change to Netflix's plan-switching policy decreases the average overpurchase by 7.4%. Such a reduction is expected, as consumers would have greater flexibility in adjusting their plan

[8] Note that the preceding counterfactual analyses are based on the focal company's (once-every-month) switching policy, and the results should be interpreted with this assumption in mind. To extend our modeling framework to another firm, the consumer's purchase decisions must be modeled at a different time frame consistent with the company's switching policy.

Table 7 **Overpurchase Tenure and Profit Under Alternative Marketing Strategies**

	All consumers (%)		
	Percent change in overpurchase	Percent change in tenure	Percent change in profit per customer
A: Uses Netflix's most recent switching policy	−7.4	4.1	−2.9
B1: Throttling (reducing available DVDs by 10%) for the heaviest 10% of users	6.7	−5.0	8.9
B2: Throttling (reducing available DVDs by 20%) for the heaviest 10% of users	14.3	−12.2	12.4
C: Giving price discounts to consumers who exhibit excessive overpurchase	5.3	7.9	3.3

choices, and would consequently reduce their overpurchase amount. The overall profit is 2.9% less than the current switching policy.

Figures 6(a) and 6(b) further compare the simulated dynamics of overpurchase with the current and Netflix switching policies. The dynamic paths of overpurchase for both consumer segments remain qualitatively the same. The hypothetical switching policy results in a substantial decrease in overpurchase, especially for the value-seeking segment. Nevertheless, Netflix's switching policy does not completely eliminate overpurchase, which remains positive and significant. Intuitively, the restriction on the downgrade and the high switching costs continue to prevent consumers from adjusting their purchases quickly enough so that the need for overpurchase remains for consumers. Together, these results suggest that Netflix's somewhat restrictive plan-switching policy and high switching costs contribute to the observed overpurchase.

Figure 6(a) **(Color online) Overpurchase Dynamics, Convenience Segment**

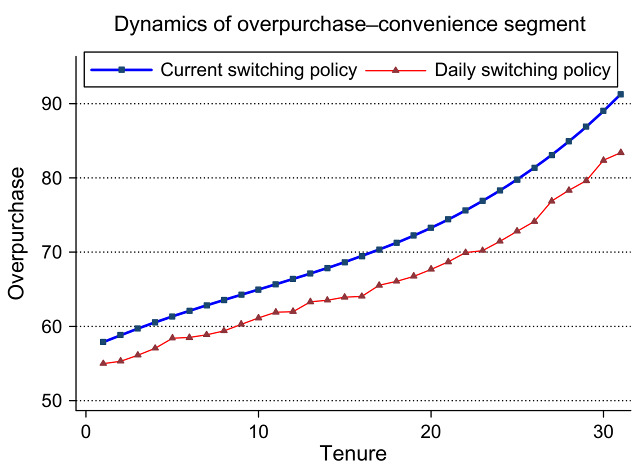
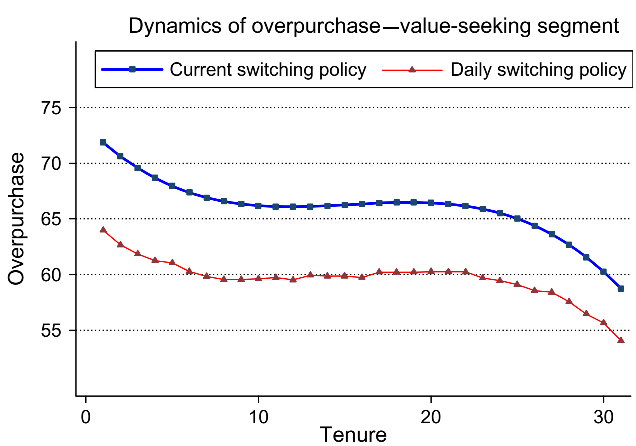


Figure 6(b) **(Color online) Overpurchase Dynamics, Value-Seeking Segment**



Throttling. Netflix's decision to deliberately increase the turnaround time for the heaviest users to reduce the operational costs of serving them has become a well-known anecdote. It is interesting to understand to what extent such a practice will affect the company's profits. In this counterfactual, we first identify the heaviest 10% of users, based on the average monthly consumption rates. We then approximate the practice of throttling by reducing the available movies, by either 10% or 20%. The other 90% of consumers are not affected. Rows B1–B2 of Table 7 show the results. The per-customer total costs for the company dropped 5.0% and 12.2%, respectively, following a reduction of available movies by 10% and 20% for the heaviest 10% of users. The cost savings turn out to be greater than the reduced revenues and eventually result in a per (affected) customer profit increase of 8.9% and 12.4% for the two levels of reduction. These results suggest that the practice of throttling may increase the net profit for the heaviest users. However, an important caveat is that such a practice has likely led to unintended sequences (e.g., negative publicity and a class-action lawsuit in the case of Netflix).

Targeted Price Discount. Row (C) of Table 7 shows the simulation of giving price discounts to consumers with excessive and persistent overpurchases. More specifically, a 50% price discount is given to consumers whose overpurchase has exceeded 80% for at least two months. The rationale is that although these consumers generate high profits in the short run, these customers are at a high risk of leaving the company permanently (recall that in the data, no consumer reinitiated the service). The company would rather keep these customers at a lower profit margin than lose them forever. We assume that the price discount will not affect their consumption decisions. Results show that the price discount is very effective, increasing tenure by 7.9%. As a result, the overpurchase is increased by 5.3%, and the total profit increases 3.3% for an average targeted consumer.

To summarize, the additional counterfactual exercises demonstrate various ways of improving the design of BBPD. The consistent theme across these exercises is that the company benefits by achieving a better trade-off between overpurchase and the quota.

5. Conclusions, Managerial Implications, and Future Research

Despite the economic significance of BBPD and its unique design, no previous research has systematically studied consumers' decision calculus in this context. We contribute to the literature by building an empirical model of dynamic purchase decisions under BBPD. Our model encapsulates several joint drivers of consumers' dynamic purchase decisions for continuous subscription services: uncertainty in consumption needs, switching costs, and consumer forward looking. Importantly, our model incorporates the instantaneous quota, an essential and unique feature of BBPD. We show how stockout risk induced by the instantaneous quota rationalizes persistent overpurchase, a seemingly irrational behavior observed under BBPD.

Broadly speaking, our research contributes to the nonlinear pricing literature, especially research in 3PT (e.g., Lambrecht et al. 2007). Whereas 3PT allows consumers to balance between upfront fixed subscription fees and marginal fees, BBPD introduces a new trade-off for consumers, i.e., between stockout costs and fixed fee payments by eliminating marginal fees. We show that the company also faces a new trade-off: between overpurchase (which drives short-term profitability) and customer retention (which drives long-term profitability). We conduct counterfactual analyses to generate some initial insights into how the company may benefit from the instantaneous quota and the ensuing stockout risk.

Our research generates several important managerial implications. First, we help managers gain deep insights into consumers' dynamic choices among competing BBPD plans. We propose two key measures for services that use BBPD—overpurchase and retention—both of which can be readily observed by the company. The company should be aware that both are driven by consumption uncertainty, stockout costs, and switching costs, and it should recognize overpurchase as a double-edged sword. Second, we show various ways in which the company can leverage insights from consumers' decision making to improve BBPD design. In addition to the counterfactual analyses in the paper, there are more fundamental ways by which the company can influence consumers' purchase decisions. For example, the company can either strategically increase consumers' switching costs (e.g., by charging explicit fees for plan switching), or decrease such costs (e.g., by sending reminder emails to consumers, or by making the switching process more friendly) when it is more profitable to do so. Apparently, the substantial heterogeneity among consumers forms a meaningful basis for targeted marketing actions. The common rationale behind the many possibilities available to the company (e.g., instantaneous versus monthly quotas, alternative prices) is that the company should balance overpurchase and retention (and manage the less-obvious downside of overpurchase) so as to increase overall profitability. Third, our research also sheds new insights into the differences between BBPD and 3PT, which is useful for companies that have already adopted either one of these two popular price formats, but may consider the other alternative.

Our research is subject to several limitations, which provide promising avenues for further research. First, we treated the consumers' consumption needs as exogenous. Future research can consider consumers' endogenous consumption and purchase decisions in an integrated framework, to better understand consumption decisions in BBPD, and offer insights into possible bunching around quota as in a three-part tariff. Second, our findings are based on a specific company. It is worthwhile for future research to extend our framework to other service categories with different magnitudes of stockout and switching costs to examine the generalizability of these results. We also did not fully capture competitive effects, and our measures of competitive effects (Netflix price and Blockbuster presence) are arguably imperfect. Although we believe these are unlikely to be of serious concern for our focal company (which focuses on the niche market of family oriented movies and is not in direct competition with Netflix or Blockbuster), future research can more fully model competition facing the company. Third, we focus on modeling consumers' dynamic purchase decisions in a time-consistent framework. However, a limitation of our model is that we treat consumption needs as exogenous. For example, an alternative behavioral explanation of overpurchase is that at the time of purchase, the consumer simply does not account for the future fall in consumption needs.[9]

[9] The findings from the behavioral literature on movie consumption (e.g., Milkman et al. 2009, Read et al. 1999) suggest the possibility of dynamic-inconsistency in consumption preferences. If consumers are forward looking in the sense of correctly forming

Thus, caution is advised in interpreting the stockout coefficient.

We provide three suggestions for future attempts to apply our modeling framework to other BBPD services: (1) be sensitive to the possible presence of dynamic inconsistency and be aware that time inconsistency may be stronger for certain "vice" products (e.g., video games) than for other products (e.g., books); (2) try to collect additional data (e.g., survey data used in DellaVigna and Malmendier 2006) to measure the extent to which the consumer's expectation of usage is consistent with her actual behavior, or try to collect measures that meaningfully correlate with the degree of dynamic inconsistency (e.g., preference for "high-brow" versus "low-brow" movies, examined in Milkman et al. (2009) and Read et al. 1999); and (3) check other possible behavioral explanations of overpurchase, such as social comparison and bunching of consumption below the quota. Fourth, the rental service can be viewed conceptually as a durable good with a long selling horizon. A monopolistic firm that cannot commit to a specific price faces the dynamic inconsistency problem, which in general reduces firm's profitability (e.g., Coase 1972, DeGraba 1994). We do not explicitly consider a firm's dynamic inconsistency problem, based on the consideration that our focal firm did not implement any price changes during the observation period. Future research should investigate consumers' responses to the possible changes in the firm's decisions over its marketing mix (e.g., prices) and formulate the firm's decision as a solution to a dynamic (as opposed to a static, "one-shot") optimization problem, and explore conditions where the firm can ameliorate the dynamic inconsistency problem. Fifth, to facilitate model identification, we follow the previous literature and fix the discount factor, which is likely to be a strong assumption (e.g., Frederick et al. 2002). Given the recent advances in estimating the discount factors (e.g., Dubé et al. 2010), future research can use additional information (e.g., consumer surveys) for more accurate estimates of the discount factor. Sixth, because of the lack of demographic information, we used a latent-class approach to capture consumer heterogeneity. Future research with access to more information can apply the Bayesian estimation methods proposed by Imai et al. (2009). Finally, because of data restrictions, our counterfactuals focus on existing customers and do not consider the potential effects of various marketing strategies on customer acquisition. Future research can examine both retention and acquisition for BBPD services.

Acknowledgments

The authors thank Ajay Kalra, Vineet Kumar, Yanwen Wang, the editor-in-chief, associate editor, and two anonymous referees for comments on earlier drafts of this paper.

Appendix

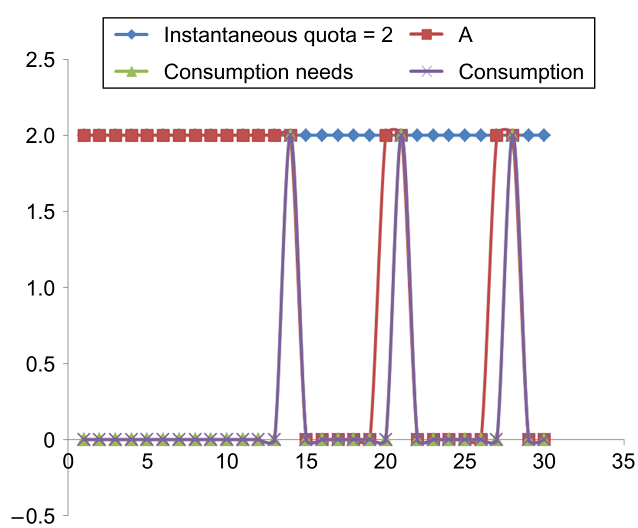
Instantaneous Quota (*IQT*) **and Available Inventory** (*A*)

Recall that the instantaneous quota (*IQT*) is the sum of A_{ijt}—the number of DVDs in the hands of the consumer immediately available for consumption—and T_{ijt} is the number of DVDs in the mailing process not immediately available for consumption

$$\begin{aligned} IQT_j &= A_{ijt} + T_{ijt}, \\ A_{ijt}, T_{ijt} &\geq 0. \end{aligned} \tag{19}$$

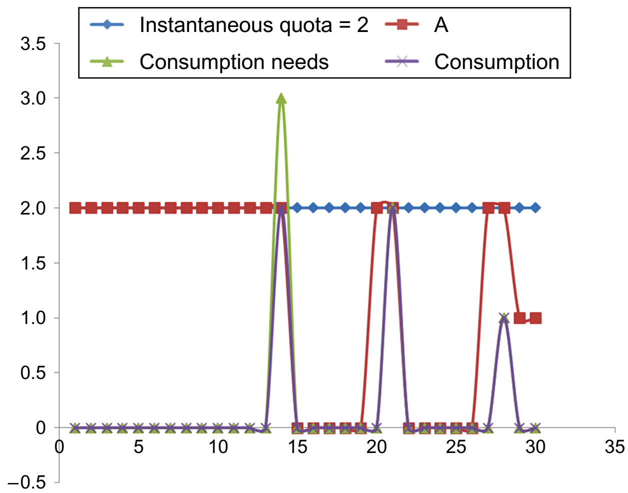
Figure A.1 uses a simple example to illustrate the relationship between the *IQT* and the available quota (*A*) for a hypothetical consumer "Alice" over a month for a BBPD plan with $IQT = 2$. The turnaround time T for Alice is six days. In this example, Alice has consumption needs of two DVDs on day 14, two DVDs on day 21, and two DVDs on day 28, and her consumption needs for all other days are zero. First, the instantaneous quota, represented by the blue series, stays constant throughout the month, whereas A, represented by the red series, varies across the month. Observe that on any given day, A is either equal to or smaller than *IQT*. Second, after Alice's consumption needs are realized, the available inventory for her is zero for the next six days (the time required for shipping the old movies, and for shipping the new movies), and then is fully restored to two on day 7 (i.e., the new movies come

Figure A.1 (Color online) Dynamics of Consumption, Available Quota for a Hypothetical Consumer, "Alice"



their expectations for future consumption needs, and they make purchase decisions accordingly, the parameter estimates in the purchase utility model will not be biased. However, if consumers are myopic in the sense of failing to anticipate the change in consumption preference between the purchase and consumption occasions, then dynamic inconsistency in conjunction with myopic consumers can potentially bias the magnitude of overpurchase and the estimated stockout coefficient along the same direction. Based on the findings of Milkman et al. (2009), we conduct a simple test (details available on request) for dynamic inconsistency, and we do not find strong evidence that consumers' overpurchase is systematically related to their movie preference.

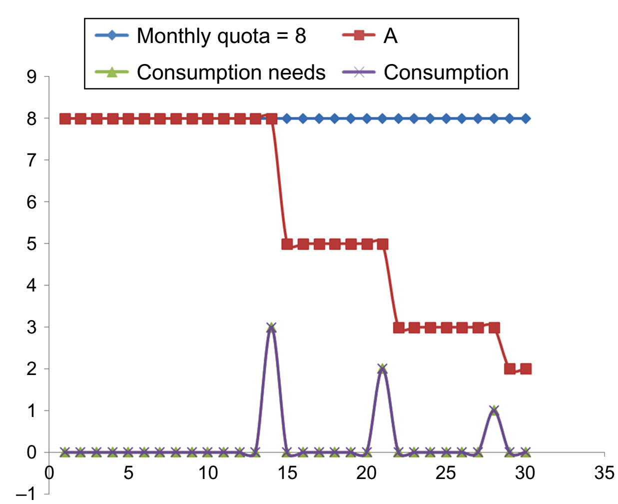
Figure A.2 (Color online) Dynamics of Consumption, Available Quota for a Hypothetical Consumer, "Samuel"



back in the mail). Given her consumption needs, Alice is able to avoid stockout completely by choosing the plan with $IQT = 2$.

To illustrate how BBPD penalizes consumers based on peak consumption, consider a second hypothetical consumer, "Samuel," who has consumption needs of three DVDs on day 14, two DVDs on day 21, and one DVD on day 28, and zero consumption needs for all other days. Thus, the total monthly consumption need for Samuel is six movies, or the same as Alice's. However, Samuel's peak consumption needs (three movies) are higher than those of Alice's. Suppose Samuel selects the plan with $IQT = 2$: the total realized consumption is five movies, with a stockout of one movie, which occurred at Samuel's peak consumption need (three movies). Notice that with BBPD, Samuel will need to increase his purchase capacity by 50% (i.e., switch from $IQT = 2$ to $IQT = 3$) to fully eliminate stockout. In this case, his realized consumption only increases by 16% (from five to six movies).

Figure A.3 (Color online) A Simple Illustration of the Monthly Quota and the Available Quota for "Samuel"



Instantaneous and Monthly Quota

Following the example discussed in Figure A.2, consider an alternative service plan with a monthly quota of eight movies. The "eight movie" quota is chosen to be the maximum consumption capacity that corresponds to a BBPD plan with a monthly quota of $IQT = 2$. In this case, Samuel incurs zero stockout, which is less than the stockout for the BBPD counterpart (one movie).

Figure A.3 illustrates why a monthly quota is less likely to induce stockout, compared with an instantaneous quota.

References

Ailawadi K, Neslin S (1998) The effect of promotion on consumption: Buying more and consuming it faster. *J. Marketing Res.* 35(3):390–398.

Ascarza E, Lambrecht A, Vilcassim N (2012) When talk is "free": The effect of tariff structure on usage under two- and three-part tariffs. *J. Marketing Res.* 49(6):882–899.

Bucklin R, Gupta S (1992) Brand choice, purchase incidence, and segmentation: An integrated modeling approach. *J. Marketing Res.* 29(2):201–215.

Chaloupka FJ, Grossman M, Saffer H (2002) The effects of price on alcohol consumption and alcohol-related problems. *Alcohol Res. Health* 26(1):22–35.

Chung D, Steenburgh T, Sudhir K (2014) Do bonuses enhance sales productivity? A dynamic structural analysis of bonus based compensation plans. *Marketing Sci.* 33(2):165–187.

Coase RH (1972) Durability and monopoly. *J. Law Econom.* 15(1): 143–149.

Danaher PJ (2002) Optimal pricing of new subscription services: Analysis of a market experiment. *Marketing Sci.* 21(2):119–138.

DeGraba P (1994) No lease is short enough to solve the time inconsistency problem. *J. Indust. Econom.* 42(4):361–374.

DellaVigna S, Malmendier U (2006) Paying not to go to the gym. *Amer. Econom. Rev.* 93(3):694–719.

Dubé JP, Hitsch GJ, Jindal P (2010) Estimating durable goods adoption decisions from stated preference data. Working paper, University of Chicago, Chicago.

Erdem T, Keane MP (1996) Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets. *Marketing Sci.* 15(1):1–20.

Erdem T, Imai S, Keane MP (2003) Brand and quantity dynamics under price uncertainty. *Quant. Marketing Econom.* 1(1):5–64.

Essegaier S, Gupta S, Zhang ZJ (2002) Pricing access services. *Marketing Sci.* 21(2):139–160.

Frederick S, Loewenstein G, O'Donoghue T (2002) Time discounting and time preference: A critical review. *J. Econom. Literature* 40(2):351–401.

Goettler R, Clay K (2011) Tariff choice with consumer learning and switching costs. *J. Marketing Res.* 48(4):633–652.

Hitt LM, Chen P (2005) Bundling with customer self-selection: A simple approach to bundling low-marginal-cost goods. *Management Sci.* 51(10):1481–1493.

Hobson M, Spady RH (1988) The demand for local telephone service under optional local measured service. *Bellcore Econom. Disc. Paper* No. 50.

Imai S, Jain N, Ching A (2009) Bayesian estimation of dynamic discrete choice models. *Econometrica* 77(6):1865–1899.

Iyengar R (2010) Pricing for a menu of service plans—An application to the digital music industry. Working paper, University of Pennsylvania, Philadelphia.

Iyengar R, Ansari A, Gupta S (2007) A model of consumer learning for service quality and usage. *J. Marketing Res.* 44(4): 529–544.

Jain S (2012) Marketing of vice goods: A strategic analysis of the package size decision. *Marketing Sci.* 31(1):36–51.

Kamakura WA, Russell GJ (1989) A probabilistic choice model for market segmentation and elasticity structure. *J. Marketing Res.* 26(4):379–390.

Keane MP (1993) Simulation estimation for panel data models with limited dependent variables. Maddala GS, Rao CR, Vinod HD, eds. *Handbook of Statistics*, Vol. 11 (North-Holland, Amsterdam), 545–571.

Keane MP, Wolpin KI (1994) Solution and estimation of dynamic programming models by simulation. *Rev. Econom. Stats.* 76(4): 648–672.

Kling JP, van der Ploeg SS (1990) Estimating local call elasticities with a model of stochastic class of service and usage choice. de Fontenay A, Shugard MH, Sibley DS, eds. *Telecommunications Demand Modeling* (North Holland, Amsterdam), 119–136.

Kopalle PK, Sun Y, Neslin SA, Sun B, Swaminathan V (2012) The joint sales impact of frequency reward and customer tier components of loyalty programs. *Marketing Sci.* 31(2):216–235.

Kridel D, Lehmann DE, Weisman DL (1993) Option value, telecommunications demand and policy. *Info. Econom. Policy* 5(2): 125–144.

Laibson D (1997) Golden eggs and hyperbolic discounting. *Quart. J. Econom.* 112(2):443–477.

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

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

Lancaster KJ (1966) A new approach to consumer theory. *J. Political Econom.* 74(2):132–157.

Lovelock CH, Wirtz J (2007) *Services Marketing: People, Technology and Strategy*, 6th ed. (Prentice Hall, Upper Saddle River, NJ).

McFadden D (1989) A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica* 57(5):995–1026.

Menard S (2002) *Applied Logistic Regression Analysis* (Sage, Thousand Oaks, CA).

Milkman KL, Rogers T, Bazerman MH (2009) Highbrow films gather dust: Time-inconsistent preferences and online movie rentals. *Management Sci.* 55(6):1047–1059.

Miravete E (2002a) Estimating demand for local telephone service with asymmetric information and optional calling plans. *Rev. Econom. Stud.* 69(4):943–971.

Miravete E (2002b) Choosing the wrong calling plan? Ignorance and learning. *Amer. Econom. Rev.* 93(1):297–310.

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

Mullahy J (1986) Cigarette smoking: Habits, health concerns, and heterogeneous unobservables in a microeconomic analysis of consumer demand. Dissertation, University of Virginia, Charlottesville.

Mussa M, Rosen S (1978) Monopoly and product quality. *J. Econom. Theory* 18(2):301–317.

Narayanan S, Chintagunta P, Miravete EJ (2007) The role of self-selection, usage uncertainty and learning in the demand for local telephone service. *Quant. Marketing Econom.* 5(4):1–34.

Oi WY (1971) A Disneyland dilemma: Two-part tariffs for a Mickey Mouse monopoly. *Quart. J. Econom.* 85(1):77–96.

Read D, Loewenstein G, Kalyanaraman S (1999) Mixing virtue and vice: Combining the immediacy effect and the diversification heuristic. *J. Behavioral Decision Making* 12(4):257–273.

Rochet J-C, Stole LA (2002) Nonlinear pricing with random participation. *Rev. Econ. Stud.* 69:277–311.

Schlereth C, Skiera B (2012) Measurement of consumer preference for bucket pricing plans with different service attributes. *Intern. J. Res. Marketing* 29(2):167–180.

Sun B (2005) The promotion effect on endogenous consumption. *Marketing Sci.* 24(3):430–443.

Sun B, Neslin SA, Srinivasan K (2003) Measuring the impact of promotions on brand switching when consumers are forward looking. *J. Marketing Res.* 40(4):389–405.

Train KE, Ben-Akiva M, Atherton T (1989) Consumption patterns and self-selecting tariffs. *Rev. Econom. Statist.* 71(1):62–73.

Train KE, McFadden DL, Ben-Akiva M (1987) The demand for local telephone service: A fully discrete model of residential calling patterns and service choices. *RAND J. Econom.* 18(1): 109–123.

Wertenbroch K (1998) Consumption self-control by rationing purchase quantities of virtue and vice. *Marketing Sci.* 17(4): 317–337.

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

Xie J, Shugan SM (2001) Electronic tickets, smart cards, and online prepayments: When and how to advance sell. *Marketing Sci.* 20(3):219–243.

Zauberman G (2003) The intertemporal dynamics of consumer lock-in. *J. Consumer Res.* 30(3):405–419.