Manufacturing & Service Operations Management

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

Priority Allocation in a Rental Model with Decreasing Demand

Apurva Jain, Kamran Moinzadeh, Aussadavut Dumrongsiri

MANUFACTURING & SERVICE OPERATIONS MANAGEMENT
Vol. 17, No. 2, Spring 2015, pp. 236–248
ISSN 1523-4614 (print) | ISSN 1526-5498 (online)
http://dx.doi.org/10.1287/msom.2014.0515

Priority Allocation in a Rental Model with Decreasing Demand

Apurva Jain, Kamran Moinzadeh
Foster School of Business, University of Washington, Seattle, Washington 98195
{apurva@u.washington.edu, kamran@u.washington.edu}

Aussadavut Dumrongsiri
Sirindhorn International Institute of Technology, Thammasat University, Bangkadi,
Muang Pathum Thani 12000, Thailand, aussadavut@siit.tu.ac.th

We analyze a model of rental and return process where limited inventory of a product is rented to two customer classes that differ in their return behavior and penalty costs. The rental demand is a decreasing function of time. We consider two cases: where a demand that is not met is lost and where an unmet demand returns. We show that to minimize penalty cost, the optimal allocation policy may give priority to different classes at different points in time and may decline lower-class demand for some time. Computational results show the benefit of the optimal allocation policy over a priority scheme reportedly used in practice.

Keywords: inventory theory and control
History: Received: July 21, 2011; accepted: September 27, 2014. Published online in *Articles in Advance* February 27, 2015.

1. Introduction

We model the rental process for a firm that acquires multiple units of a product and rents them to its customers. In case the firm is not able to satisfy a rental demand, it is charged a penalty cost. Customers belong to one of the two classes that differ in rental durations and penalty costs. A demand that is not met may be lost or may return; we consider both cases. To minimize the total penalty cost, an allocation policy must decide which class should be awarded limited on-hand inventory at any given time. In line with the fluid models of rental and return process in literature, we develop and analyze an optimal control formulation to determine the allocation policy for rental inventory with an objective to minimize the total penalty cost. Our focus is on understanding the theoretical and practical implications of our model's distinguishing feature—decreasing rental demand from customers of both classes.

Rental businesses are prevalent in practice, but the literature on rental inventory models is comparatively limited. Such models must capture detailed dynamics of the return process. For stationary stochastic demand, issuing and return of rental inventory can be thought of in terms of queueing theory. In that context, there is a well-developed stream of literature devoted to studying the optimal demand admission or inventory allocation policy. Even though our fluid model is different in some details and assumptions, the conceptual issues and the managerial question of interest are well rooted in this literature; that is, at different states of system evolution, what types of customers should be allocated limited rental inventory? The trade-off in making this allocation decision goes back to weighing the impact of allocating a unit, making it unavailable for the rental duration, against the impact of incurring a penalty for not satisfying the demand. We ask what would be the influence of a decreasing demand pattern on resolving this trade-off.

Business publications regularly report on the rise of rental economy, including the part that is our focus in this paper—lending of physical goods by a firm. For example, Miller (2014) reports on a variety of rental business models and refers to lending-by-firm models as Netflix economy. While the Netflix DVD-by-mail rental business remains an early example, similar business models are being launched in a variety of areas. Recently, Griffith (2014) reported on Rent the Runaway, which rents designer dresses. The company offers a subscription-based pricing option that allows a customer to rent and use a limited number of items at a time. It is reported to have a customer base of five million with a variety of preferences on how much they use the service from special-occasion wear to regular wear. Another popular item that appears to fit this business model is jewelry. McDowell (2014) reports on several firms that operate in this space. Borrowed Bling offers different subscription price levels and allows subscribers to rent a limited number of jewelry items at a time. Haute Vault, too, has a subscription model and adds watches to its list of items.

Olney (2013) adds Rocksbox and Le Tote for renting jewelry to this list. Other examples of recent rental business of similar type include handbags (bagborrowandsteal.com), and books at BookFree.com, which offers monthly subscription plans that allow the customer to borrow (with no shipping costs) a limited number of paperbacks and audio books and keep them as long as they want. Getting back to the early Netflix example, even as Netflix's online streaming business grows, the number of DVD-by-mail subscribers appears to have reached a plateau, and new blockbuster titles are almost exclusively available on DVD when they are first offered by Netflix (see Stenovec 2014).

Most of these businesses deal with goods that have a *fashion goods* component in the sense that new designs are released periodically and demand tends to be high at the time of release before it decreases over time. The primary research question we address is how the firm should allocate its limited inventory between different customer segments in presence of *decreasing* demand over time. This research question is rooted firmly in the existing literature and, at the same time, allows us to add new insights to both theory and practice. Before presenting that general discussion, we briefly note a specific example of this question. Rental demand for a new movie title displays a decreasing pattern over time (see Lehmann and Weinberg 2000, Pasternack and Drezner 1999). Netflix DVD rentals usually have limited inventory of newly released titles (see Stross 2007, Lilly 2008). Customers are segmented by their rental frequency (see *KOMO News* 2006); "heavy" customers rent and return more frequently than "light" customers and cost more to firm in delivery cost. In this setting, Rosales (2006) refers to the allocation decision of prioritizing heavy customers as "throttling." By addressing our research question, we generate insights into this specific example as well as into the broader research issue of allocating limited rental inventory when demand is decreasing.

Before discussing our theoretical findings and managerial insights, we start with reviewing the literature on rental models with stationary demand. Earliest rental models like Tainiter (1964) and Whisler (1967) drew parallels between the units of rental inventory and number of servers in a queuing system. Miller (1969) advanced to multiple arrival classes that differed in per-unit revenue and arrival rates but not in service times. In Albright (1977), the servers were heterogeneous, and the reward from each class became dependent on the server assigned to it. Ormeci et al. (2001) considered a two-class system where classes differed in required service times. In Altman et al. (2001), each class required different service times and different number of servers. In a two-class system, Savin et al. (2005) further differentiated the classes based on a lump-sum penalty cost. They derived the properties of the optimal policy and offered an approximation to compute policy parameters.

The main modeling tool in this stream of research is the Markov decision process (MDP) formulation of a system with Poisson demand arrivals and exponential service times for different classes. Given the difficulty of the model, a fluid approximation was employed, as in Altman et al. (2001) and Savin et al. (2005). To serve our purpose of exactly analyzing the case of decreasing demand, our model is somewhat simpler in that we directly start with a fluid model. We use optimal control methods to determine the optimal policy. Rather than capturing customer revenue, we model a penalty cost for not satisfying a customer's demand. Classes differ in rental durations and penalty costs. These modeling choices allow us to not only exactly analyze the decreasing demand case but to also capture the case of unsatisfied demand returns, a feature not discussed in the literature described above.

The video rental business context, which has motivated our model, has also attracted considerable research attention. Tang and Deo (2008) model a video rental and return process by assuming that a fixed fraction of a period's rental will be returned in each period before the due date. However, they assume that demand follows a stationary process, customers are similar in their return behavior, and they focus on the competition between rental firms. Pasternack and Drezner (1999) and Gerchak et al. (2006) capture the decreasing demand feature in video rental business. These models, however, do not explicitly model the return process and fail to differentiate between customers based on their characteristics. A recent work that is closest to our model is Bassamboo and Randhawa (2007). In their model, all the new demand arrives at time 0, and whereas our model considers different penalty costs for two classes, they treat both classes equally when formulating the objective function, which is to minimize the total waiting time.

Our contribution is primarily in the area of developing structural results about the optimal allocation policy in our optimal control model with decreasing demand. We also develop managerial insights for a video rental business based on computational experiments.

The structure of the optimal policy confirms some insights from earlier literature and adds new ones. Given that demand is decreasing over time, we find it natural to describe the optimal policy in terms of evolution of time rather than as a function of state space. At the beginning, it is optimal to satisfy demands from both classes; this is sometimes called the complete sharing policy in the literature on stochastic loss systems. We then show the existence of a time-point at which complete sharing is no more optimal in the sense that even though inventory is available, demand for one class is not satisfied. This is similar to rationing in anticipation of future random high-class demand in stochastic inventory models. In our model, this phenomenon occurs without randomness in demand. In effect, the optimal policy shuts out lower-class demands and saves rental inventory to delay the time when it is no longer able to satisfy all the higher-class demand. As we go forward, such a time eventually arrives; initial inventory runs out, and the system enters a no-inventory region. Now it is optimal to use the returning units to satisfy as much of the higher-class demand as possible. We call this a priority allocation policy. In the case where unmet demands are lost, we show that identity of the class that gets high priority can be determined by a simple "$c\mu$-type" rule; however, in the case where unmet demands return, a simple rule is not sufficient, and we outline a procedure.

As time progresses, since demand rate is decreasing, return rate is eventually more than the demand rate. The system enters a positive-inventory region, and we do not have to make allocation decisions anymore. However, before entering this positive-inventory region, we find, to our surprise, that the optimal allocation policy may switch the identity of class that is given high priority.

Since the rental literature has focused on stationary demand, the existing insights from that literature do not immediately explain why the optimal policy in our model switches priority between classes. At the beginning of no-inventory period, there is value in getting a rental unit back early so that it can be reissued. This favors the class that keeps the unit for shorter time. As time progresses and demand decreases, the value of getting a unit back faster also decreases because there may not be enough demand to reissue it. Thus, the importance of rental duration in determining the optimal priority class diminishes and the importance of penalty cost increases, leading to the switching of the priority class. The combination of return process and decreasing demand results in this new property of the optimal allocation policy. We believe that the paper's contribution lies in identifying these new features of optimal allocation policy in presence of decreasing demand. Moving beyond what we find theoretically interesting in the results, we use numerical experiments to highlight the advantage of our optimal policy over a simple allocation policy reportedly used in practice.

The rest of this paper is organized as follows. In §2, we describe the basic model. In §3, we develop the optimal priority allocation policy for the model with lost demand. In §4, we consider the model with returning demand. In §5, we present computational results. We close the paper in §6 by summarizing the results and suggesting avenues for future work.

2. The Basic Model

Consider a firm that rents a product to two classes of customers. The rental volumes are large, and rented units are received at and dispatched from the rental firm continuously; therefore, we capture the dynamics in a continuous-state, continuous-time model. As motivated by the discussion in §1, the customer classes are differentiated based on their rental behavior: heavy renters (denoted by subscript H) return the rented unit faster, and the light renters return the rental slower (denoted by subscript L). We will use i, as in "class i" or as in a variable with subscript i, to refer to any one of the two classes, $i = H, L$. We capture the return behavior of a class with return fraction p_i, the fraction of outstanding units that are returned at any time. Heavy renters return faster; thus, their return fraction is higher than that of light renters, $p_H > p_L$. The notion of return fraction is similar to exponentially distributed rental durations, which have a constant hazard rate property, a common feature in stochastic versions of this problem. Let state variables $k_i(t)$ denote the number of outstanding rental units with customers of class $i = H, L$ at time t. The return rate for class i at time t is then given by $p_i k_i(t)$.

If a customer's demand for renting a unit is not met, the demand may be lost or the customer may return in the future. Correspondingly, we consider two models: "lost demand" model in §3 and "returning demand" model in §4. In either model, a penalty cost π_i is charged to the firm for not meeting the demand. Furthermore, we consider cases where the penalty cost for heavy renters can be either more or less than the penalty cost for light renters. However, in this study, we do not explicitly model a reward for meeting a customer demand. There are two reasons for this. Frist, in the context that motivated us, video rentals, the customers either pay a fixed price per rental or only pay a fixed fee per period without paying anything per rental. Second, as we will observe in the lost demand model presented in §3, including a per-rental reward for satisfied demand will only require us to redefine the penalty cost. Thus, the structure and the analysis of the model will remain the same. We note that the penalty cost in our model is primarily capturing the loss in good will. Because of subscription-based payments, there is no actual money that the firm loses even though it would be indirectly penalized for dissatisfying the customers.

We model the demand rate from class i at time t, as deterministic continuous function $d_i(t)$ with

$dd(t)/dt = d_i'(t) < 0$. Note that there are no other restrictions on the shape of the demand curve. Keeping demand deterministic allows us to add other features to the model that are not considered in the literature; that is, the decreasing nature of demand and the possibility that the demand, if not met the first time, may return (returning demand model). As discussed in §1, our assumption on demand is motivated by observations in movie rental business. Lehmann and Weinberg (2000) plot data that show video rental demand to be declining exponentially over time. They show that the temporal demand pattern (both the initial intercept and the slope) is quite predictable based on demand data from theaters. Ainslie et al. (2005) fit various decreasing functions to box office sales. Pasternack and Drezner (1999) and Gerchak et al. (2006) point out the predictability of movie rental demand. This supports our modeling of demand as decreasing deterministic functions.

The firm acquires an initial inventory of M copies of the rental product. For a fixed value of M, the firm's objective is to minimize the total penalty cost over the life of the product. Since the two classes differ in their return fractions p_i, the problem is to determine the optimal allocation policy that will minimize the total penalty cost.

3. Analysis of the Lost Demand Model

We start with a general formulation of the problem. The formulation introduces control variables $u_i(t)$, as the instantaneous fraction of class i demand $d_i(t)$ that is satisfied. Then, $(1 - u_i(t))d_i(t)$ represents the rate of class i demand lost at time t. We assume that lost class i demand is charged a penalty π_i. The objective is to minimize total penalty cost.

$$\underset{u_H(t),\,u_L(t)}{\text{minimize}} \int_0^\infty [\pi_H(1-u_H(t))d_H(t)+\pi_L(1-u_L(t))d_L(t)]\,dt$$

$$= \underset{u_H(t),\,u_L(t)}{\text{maximize}} \int_0^\infty [\pi_H u_H(t)d_H(t)+\pi_L u_L(t)d_L(t)]\,dt$$

subject to

$$\text{state equations:}\quad k_H'(t)=u_H(t)d_H(t)-p_H k_H(t), \tag{1}$$

$$k_L'(t)=u_L(t)d_L(t)-p_L k_L(t); \tag{2}$$

$$\text{inventory constraint:}\quad h(t)=M-k_H(t)-k_L(t)\geq 0; \tag{3}$$

$$\text{starting state:}\quad k_H(0)=0, \quad k_L(0)=0; \tag{4}$$

$$\text{control variables:}\quad 0\leq u_H(t)\leq 1, \quad 0\leq u_L(t)\leq 1. \tag{5}$$

Let $k_i'(t) = dk(t)/dt$. The formulation captures the dynamics of state evolution by ensuring that the rate of change $k_i'(t)$ of outstanding rental units for class i equals new rental units allocated to this class $u_i(t)d_i(t)$ minus the number of units $p_i k_i(t)$ returned from this class. The rental process begins at time 0 with an inventory of M units and with no outstanding units. The inventory constraint $h(t) \geq 0$ ensures that the sum of outstanding rental units is never more than the available inventory M. We will refer to $u_H(t)$, $u_L(t)$ as allocation or control policy and to the plot of control variable and state variables over time as control trajectory and state trajectory, respectively.

The optimal control theory provides optimality conditions for this formulation. We first write the Hamiltonian function H and the Lagrangian function L using adjoint variables $\lambda_i(t)$ and Lagrange multipliers. We then specify complementary slackness, Hamiltonian maximization, and transversality conditions. Complete details of these are provided in the online appendix (available as supplemental material at http://dx.doi.org/10.1287/msom.2014.0515).

Although optimality conditions provide a beginning point for our analysis, it is usually difficult to use them to determine optimal control variables at all points in time. As Sethi and Thompson (2000) point out, constraints involving only state variables, such as our inventory constraint, are even more difficult to deal with. Computational solutions are possible, but they usually do not provide much insight into the structural properties of the optimal solution. We are interested in understanding the *form* of the optimal control policy and the reasons behind it. The form of the policy can then be used to simplify the computational effort. Toward that end, our first result is to show that optimal values for each control variable are limited to only two possibilities.

Proposition 1. *Optimal control variables $u_i^*(t)$ take either the maximum possible value under constraints* (1)–(5) *or take the minimum possible value* (*zero*).

All proofs are available in the online appendix.

The result is useful because it allows us to conveniently divide the continuous optimal control space in four discrete possibilities: both classes are allocated maximum possible inventory, either one is allocated maximum possible and the other nothing, or no class is given anything. The specific value of control $u_i(t)$ is dependent on time and state variables, but this result allows us to narrow down the *form* of the optimal control. Whereas this is helpful, it is not a complete specification of how much to allocate to each class, especially in the case when there is not enough inventory. To consider this question, the next section classifies the evolution of rental process into two types of intervals.

3.1. Classification of Time Intervals

In the terminology of optimal control, inventory constraint (3) is a pure constraint defined only on state variables $k_i(t)$ without including any control variables $u_i(t)$. A pure constraint can be used to divide the state trajectory into two types of intervals: boundary intervals (time intervals in which the pure constraint hits the boundary $M - k_H(t) - k_L(t) = 0$) and interior intervals (where $M - k_H(t) - k_L(t) > 0$). When the process enters a boundary interval, it is called an entry time. When the process exits a boundary interval, it is called an exit time. We focus only on cases where initial inventory is scarce enough for the process to enter at least one boundary interval. In general, the evolution of the process may alternate between these two types of intervals, creating multiple interior and boundary intervals.

We start by focusing on demand functions for which we can derive a useful property that considerably simplifies the complexity created by the possibility of multiple boundary intervals. We show that for demand functions listed in Proposition 2 below, the optimal state trajectory can have at most one boundary interval. When all units are rented out and the firm runs out of inventory, the rental process enters the boundary interval. Decreasing demand in our model ensures that the process must, at some point, exit the boundary interval. Proposition 2 shows that for listed demand functions, once the rental process exits the boundary interval and enters another interior interval, it never again enters another boundary interval. We define t_0 as the entry time and t_1 as the exit time; thus $[t_0, t_1]$ is the boundary interval, and $(0, t_0)$ is the interior interval before the process enters the boundary interval.

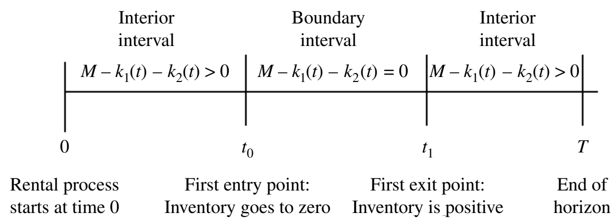
Proposition 2. *If* (a) *demand is linear or concave, $d_i'(t) < 0$, $d_i''(t) \le 0$ for both classes*, *or* (b) *demand is exponential with respect to time, $d_i(t) = a_i e^{-x_i t}$, $a_i > 0$, $x_i > 0$ for both classes, then for $t \ge t_1$ (the first exit point), the optimal control is such that the control variables are equal to one*; *that is, $u_H^*(t) = 1$, $u_L^*(t) = 1$.*

Single-boundary-interval condition: When demand functions satisfy (a) or (b), we refer to them as satisfying single-boundary-interval condition.

The single-boundary-interval property allows us to present our results about optimal policy without undue complexity. We can now limit our attention to cases with one boundary interval that is preceded and succeeded by different interior intervals, as marked in Figure 1. In cases where demand functions continuously decrease but never hit zero, the proof of the above result shows that an end-of-horizon T can be defined as the unique solution to $d_H(T) + d_L(T) = p_L M$. Starting at the exit point t_1, rental process is in an interior interval for $t \ge t_1$ in which, as the above result shows, a policy of satisfying all demand is feasible and, therefore, optimal.

Our approach in the rest of this section is as follows: starting at t_1 in Figure 1, we move to the left by first analyzing the boundary interval in §3.2 and then address the first interior interval in §3.3. Both of these sections address the case of single-boundary-interval. In §3.4, we discuss the case where multiple boundary intervals may exist.

Figure 1 **Illustration of Interior and Boundary Intervals, and Entry and Exit Points**



3.2. Optimal Control in a Boundary Interval

We are now ready to analyze the optimal control in the boundary interval $[t_0, t_1]$. Starting with $k_H(t_0) + k_L(t_0) = M$, a boundary interval, by definition, has $M - k_H(t) - k_L(t) = 0$. To satisfy this requirement, we must ensure that the slope of $M - k_H(t) - k_L(t)$ in the boundary period is never negative. The requirement $d(M - k_H(t) - k_L(t))/dt \ge 0$ translates into

boundary constraint:

$$p_H k_H(t) + p_L k_L(t) - u_H(t) d_H(t) - u_L(t) d_L(t) \ge 0. \quad (6)$$

It is useful to interpret (6) as a relationship between return rate and allocation rates. The term $p_H k_H(t) + p_L k_L(t)$ is the total rate of return from outstanding units, and the term $u_H(t) d_H(t) + u_L(t) d_L(t)$ is the total allocation of units to satisfy demand. Once the inventory is zero at the time of entry to the boundary period, the constraint (6) states that the allocation rate cannot be more than the return rate in the boundary interval. Whenever the process is in a boundary interval, we can replace inventory constraint (3) with boundary constraint (6).

Proposition 2 shows that the optimal control at the exit point is $u_i^*(t_1) = 1$. The form of optimal policy at the exit point is to allocate maximum possible amount to both classes. A corollary to the above result is that at any time in the boundary interval, a class can have zero allocation only if return rate is less than the other class's demand, because otherwise this will become an exit point and violate the above result. In effect, it is optimal to satisfy as much of both classes' demand as possible in a boundary interval. This shows that, out of four possible policies outlined below Proposition 1, only one is optimal in a boundary interval.

The above result is not yet a complete description of the form of control policy in the boundary interval. By the definition of boundary interval, there is not enough return to completely satisfy both classes' demand (if there were, the process would exit from the boundary interval). Therefore, the form of the optimal control must also answer the question of how to allocate the limited returns. We show in the next result that a "priority" control policy determines which class's demand is satisfied first. Recall that $p_H > p_L$. The statement of the next result is organized according to the value of the other parameter that differentiates between classes; that is, the penalty cost π_i. Before stating the result, we formally define the priority policy. A control policy that gives priority to class L is defined as

$$u_L^*(t) = \frac{\min\{(p_H k_H(t) + p_L k_L(t)), d_L(t)\}}{d_L(t)},$$
$$u_H^*(t) = \frac{p_H k_H(t) + p_L k_L(t) - u_L^*(t) d_L(t)}{d_H(t)}.$$

A control policy that gives priority to class H is defined as

$$u_H^*(t) = \frac{\min\{(p_H k_H(t) + p_L k_L(t)), d_H(t)\}}{d_H(t)},$$
$$u_L^*(t) = \frac{p_H k_H(t) + p_L k_L(t) - u_H^*(t) d_H(t)}{d_L(t)}.$$

Proposition 3. *If demand functions satisfy single-boundary-interval condition, then the optimal control in the boundary interval $[t_0, t_1]$ is as follows:*

Case 1: $\pi_H < \pi_L$, $p_H\pi_H < p_L\pi_L$. *The optimal control policy is to give priority to class L.*

Case 2: $\pi_H < \pi_L$, $p_H\pi_H > p_L\pi_L$. *There exists a unique $t_0 \le t_{sw} < t_1$, such that the optimal control policy is to give priority to class L in the interval $[t_{sw}, t_1]$ and priority to class H in the interval $[t_0, t_{sw})$.*

Case 3: $\pi_H > \pi_L$. *The optimal control policy is to give priority to class H.*

The "priority" policy can simply be understood as giving as much as possible to the priority class and then giving leftovers to the other class. In Case 2, we will refer to t_{sw} as the switching time because it requires us to switch the priority from one class to the other.

The results we have here fit the existing literature but also add new insights. The analysis highlights the role that the index $p_i\pi_i$ plays in determining the optimal policy. When class L carries a higher penalty cost, prioritizing class L demand over class H demand will save the penalty cost. On the other hand, class H has a higher return fraction; therefore, a rental to class H will be back into circulation quickly. This trade-off is captured by the index $p_i\pi_i$. As Case 1 shows, if this index is higher for class L, the firm should give priority to class L. The result confirms intuition because it is similar to the well-known $c\mu$ rule, even though the setting here is quite different. An intuition based on that rule will, however, suggest that in Case 2, where the index $p_i\pi_i$ is larger for class H, then the firm should always give priority to class H. As we show in Case 2, surprisingly, that is not the case. This observation brings into focus the role decreasing demand plays in this situation. As demand decreases, fewer inventories are required to satisfy it. At some point, the benefit of having a unit returned quickly is not significant enough to give priority to class H and incur additional penalty cost. As we reach switching time t_{sw}, the benefit of quick return decreases because very soon there will be enough returns to satisfy both classes' demand. Therefore, just before the exit time, there is a period of nonzero length in which class L always gets priority.

Finally, in Case 3, $\pi_H > \pi_L$, $p_H\pi_H > p_L\pi_L$, we have no reason to give priority to class L even when the system is very close to exiting the boundary period. There will be no penalty cost saving in allocating limited inventory to class L over class H.

We also find it interesting to look at this result from the perspective of related dynamic systems in the literature. In an inventory rationing model where all deterministic demand arrives at time 0, the optimal policy would simply prioritize the class with the highest π_i value. If the deterministic demand arrives at a constant rate, the system would behave like a queueing model, and the optimal policy would maximize the rate at which cost is incurred by prioritizing on index $p_i\pi_i$. In this paper, we have a rental system that exhibits properties of a queueing system (when returns of items are important) and an inventory rationing system (when returns are not important). Because the demand is decreasing, there is a point beyond which the firm will always have surplus inventory, and so it stops caring about returns. In this period of time, the firm exhibits the inventory rationing properties and prioritizes on π_i. However, before this time, it exhibits the queueing properties and prioritizes on $p_i\pi_i$.

3.3. Form of the Optimal Policy and Its Parameters

A complete statement about the form of optimal policy requires one additional step, which is to analyze the interior interval $(0, t_0)$ that precedes the boundary interval. The analysis of this period crucially depends on the behavior of adjoint variables $\lambda_i(t)$ at the entry time t_0. Hartl et al. (1995) show that these variables exhibit a discontinuity (jump) at the entry time and provide jump conditions that the values of adjoint variables just before the jump $\lambda_i(t_0^-)$ to their values just after the jump $\lambda_i(t_0^+)$. Details of these conditions are available in the online appendix. Our analysis of this interior interval leads to two subresults that are proved as part of Theorem 1, which follows.

These subresults relate the identity of priority class at time $t = t_0^+$, the beginning of boundary interval, to the optimal policy in the interior interval $(0, t_0)$. The first result shows that if a class gets priority at t_0^+, then all demand of that class is satisfied when the firm still has inventory. This makes intuitive sense because it is the same class that gets the best service from the firm both before and after entering the boundary interval. The second result shows that the class that does not get priority at at t_0^+, may experience a unique time t_{so}, $0 \le t_{so} \le t_0^-$ in $(0, t_0)$ such that nothing is allocated to it in period $t_{so} \le t \le t_0^-$. Combining the preceding two results, the class that gets priority at $t = t_0^+$ will have its demand fully satisfied in the interior interval, but the other class may be allocated nothing after t_{so}. We refer to t_{so} as shut-off time, which is when the firm stops giving inventory to the class with lower priority at $t = t_0^+$. Interestingly, even though the firm has inventory, it may choose not to meet the demand from the lower priority class. By doing so, while it incurs the penalty cost for the lower priority class, but by delaying the entry into the boundary interval, it also manages to meet more of high priority class demand. The trade-off between these two effects determines the shut-off time.

In Theorem 1 below, we formalize these results and put them together with earlier boundary interval results to fully characterize the form of the optimal control policy into two discrete policy types. This is possible because optimal policy can take only two possible values at t_0^+ identified by the index i of the class that gets priority at t_0^+. Before stating Theorem 1, we define the following two policy types:

(i) Heavy Policy: Class H gets priority at t_0^+. The form of policy is $u_H^*(t) = 1$, $u_L^*(t) = 1$ for $0 < t < t_{so}$; $u_H^*(t) = 1$, $u_L^*(t) = 0$ for $t_{so} \le t < t_0$; priority to class H for $t_0 \le t < t_{sw}$; and priority to Class L for $t_{sw} \le t \le t_1$. Policy parameters are t_{so}, t_{sw} where $0 \le t_{so} \le t_0$ and $t_0 \le t_{sw} < t_1$.

(ii) Light Policy: Class L gets priority at t_0^+. The form of policy is $u_H^*(t) = 1$, $u_L^*(t) = 1$ for $0 < t < t_{so}$; $u_H^*(t) = 0$, $u_L^*(t) = 1$ for $t_{so} \le t < t_0$; priority to class L for $t_0 \le t < t_1$. Policy parameter is t_{so}, $0 \le t_{so} \le t_0$.

Finally, $u_H^*(t) = 1$, $u_L^*(t) = 1$ for $t \ge t_1$.

Figure 2 provides a visual representation of the two types of optimal policies. For each type of policy, it shows the values of control variables at different points in time. We now prove the optimality of each type of policy for different parameter combinations.

Theorem 1. *For demand functions that satisfy the single-boundary-interval condition,*

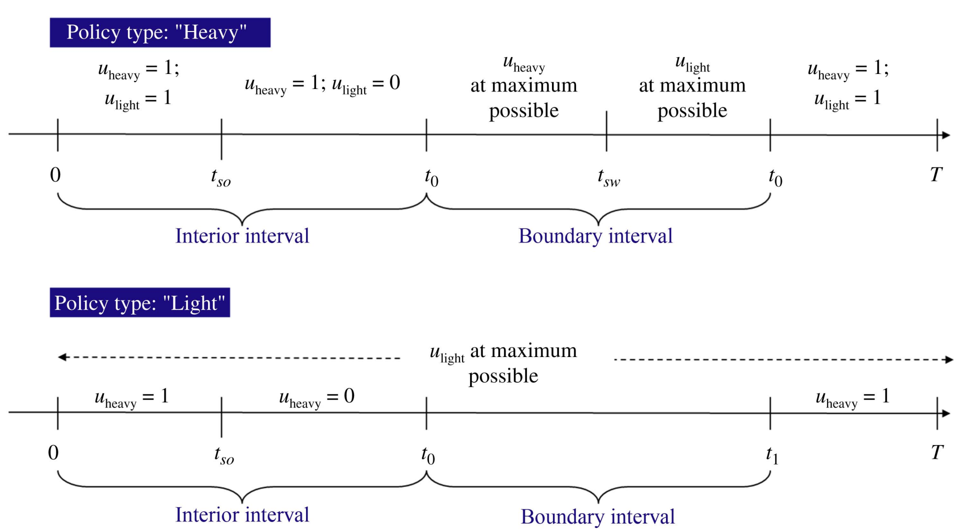
(i) if $\pi_H > \pi_L$ or $\pi_H < \pi_L$, $p_H \pi_H > p_L \pi_L$, then the optimal control policy type is Heavy;

(ii) if $\pi_H < \pi_L$, $p_H \pi_H < p_L \pi_L$, then the optimal control policy type is Light.

The above result allows us to focus on three parameters that together completely characterize the optimal policy. These parameters are *policy type*, t_{so}, t_{sw}. *Policy type* can take two possible values—*Heavy* or *Light*. Theorem 1 allows us to identify the optimal policy type based on the values of p_i, π_i. Given a specific value of *policy type*, and for any value of shut-off time t_{so}, state equations determine an entry time t_0. Given a *policy type* and entry time t_0, the proof of Theorem 1 (see the online appendix) shows how the combination of switching time t_{sw} and exit time t_1 can be computed. Finally, a search over one variable, that is, the shut-off time t_{so} (≥ 0), determines the optimal policy with the minimum cost.

Note that the process has four important time epochs: shut-off time t_{so}, entry time t_0, switching time t_{sw}, and exit time t_1. Location of t_{sw} and the gap $t_1 - t_{sw}$ represent the length of the period in which optimal

Figure 2 **(Color online) Illustration of Optimal Policy**



policy is less influenced by the value of fast returns; therefore, priority is switched to the low-return-rate class. Intuitively speaking, there are two major factors that determine this gap. First, relative values of parameters p_i, π_i for the two classes must influence the point in time when optimal policy starts putting less value on fast returns of rented units. The second factor that influences the determination of $t_1 - t_{sw}$, is the shape of the demand curve. If the demand curve decreases rapidly just before t_1, optimal policy may stop valuing fast returns sooner than it would if the demand were flatter. The gap $t_1 - t_{sw}$ in the optimal policy can be small in some cases but, for any given instance, it depends on relative values of p_i, π_i as well as the shape of demand curves.

3.4. Multiple-Boundary-Intervals Case

In our analysis, we have assumed that demand functions satisfy the single-boundary-interval condition. In such cases, the optimal state trajectory, after exiting from the boundary interval, is always able to meet all demand. This has allowed us to focus on developing insights into the structure of the optimal policy in that boundary interval. For demand functions that do not satisfy the single-boundary-interval condition, it is possible that the optimal state trajectory has N (≥ 1) boundary intervals. We can extend our existing terminology to this more general case by defining t_0^n as the nth entry time and t_1^n as the nth exit time such that $[t_0^n, t_1^n]$ is the nth boundary interval (where $n = 1, \ldots, N$). We now briefly discuss how to interpret our results in the case of multiple boundary intervals. Note that Proposition 1 does not depend on boundary intervals and, therefore, continues to hold. Clearly, conditions in Proposition 2 are not satisfied, but we can still show that beyond the last exit point $(t \geq t_1^N)$, the optimal control is to fully meet both demands. In the proof of Proposition 2, we rigorously show why this is true. Now, in the proof of Proposition 3, we can replace t_0 with t_0^N and t_1 with t_1^N and show that it continues to hold in the last boundary interval. Moreover, the statement of Theorem 1 also holds for the last boundary interval. Combining these results, we can computationally determine the optimal policy in cases where demand functions do not satisfy single-boundary-interval condition. We have carried out extensive numerical experimentations on a variety of demand function, and we will summarize our computational observations on this issue in §5.

4. Analysis of the Returning Demand Model

In this section, we analyze the case where unmet customer demand is not lost but returns after some time. We allow the possibility that the demand may not be met again and may have to return multiple times before it is satisfied. Every time a class i demand—either new or returning—is not met, a penalty cost π_i is charged to the firm. At any time, there are now two sources of demand: new demand and returning demand. New demand still follows $d_i(t)$. To keep track of returning demand, we introduce new state variables $B_i(t)$ that represent the class i demand that arrived in the past but was not met. We assume that a fraction p_i of $B_i(t)$ will return at time t as returning demand. This assumption means that we model the return behavior for both sources of demand the same way. For both, the same fraction p_i returns.

Our extension to returning demand model is motivated by the existence of customer preference lists in online DVD rental businesses like Netflix. The practice of issuing second-choice DVD if customer's first choice is not available means that the demand for first-choice title will come back when the customer returns the second choice; this is the feature we are capturing by modeling returning demand in our model. The rate at which a customer returns a rented video does not depend on whether it is her top-choice title or not. This supports our use of the same return rate for both outstanding units and accumulated unmet demand.

In the optimal control formulation for this model, $B_i(t)$, which represents the class i rental demand that occurred in the past but was not met, evolves over time. At time t, it decreases by $p_i B_i(t)$ (representing returning demand) and increases by $(1 - u_i(t)) \cdot (d_i(t) + p_i B_i(t))$ (representing the unsatisfied new and returning demand). There is no limit on how many times a customer can return if her demand is not satisfied. The optimal control formulation is as follows:

objective function:

$$\underset{u_H(t),\, u_L(t)}{\text{minimize}} \int_0^\infty [\pi_H(1 - u_H(t))\{d_H(t) + p_H B_H(t)\} + \pi_L(1 - u_L(t))\{d_L(t) + p_L B_L(t)\}]\, dt;$$

state equations:

$$k'_H(t) = u_H(t)(d_H(t) + p_H B_H(t)) - p_H k_H(t),$$
$$k'_H(t) = u_L(t)(d_L(t) + p_L B_L(t)) - p_L k_L(t),$$
$$B'_H(t) = (1 - u_H(t))(d_H(t) + p_H B_H(t)) - p_H B_H(t),$$
$$B'_L(t) = (1 - u_L(t))(d_L(t) + p_L B_L(t)) - p_L B_L(t);$$

inventory constraint: $h(t) = M - k_H(t) - k_L(t) \geq 0;$

starting state: $k_H(0) = 0, \quad k_L(0) = 0,$
$B_H(0) = 0, \quad B_L(0) = 0;$

control variables: $0 \leq u_H(t) \leq 1, \quad 0 \leq u_L(t) \leq 1.$

The analysis follows the same steps as earlier and relies on complementary slackness, Hamiltonian maximization, and transversality conditions. We first identify a point in time T after which it is feasible to satisfy all demand. We separate the state trajectory into a sequence of interior and boundary intervals. We show that it is optimal to satisfy as much demand as possible when in the boundary interval. This allows us to rewrite the formulation for the boundary interval and show that the optimal policy is a priority policy. We now have four state variables with four state equations, which makes the analysis more detailed and tedious compared to the lost demand model. Recall that t_0^n as the nth entry time and t_1^n as the nth exit time such that $[t_0^n, t_1^n]$ is the nth boundary interval (where $n = 1, \ldots, N$). We can show the following result for the returning demand model.

Theorem 2. *The optimal control policy in the last boundary interval* $[t_0^N, t_1^N]$ *preceded by the interior interval* (t_1^{N-1}, t_0^N) *has either of the two following types*:

(i) Heavy Policy: Class H gets priority at t_0^{N+}. The form of policy is $u_H^*(t) = 1$, $u_L^*(t) = 1$ for $t_1^{N-1} < t < t_{so}^N$; $u_H^*(t) = 1$, $u_L^*(t) = 0$ for $t_{so}^N \le t < t_0^N$; priority to class H for $t_0^N \le t < t_{sw}^N$; and priority to Class L for $t_{sw}^N \le t \le t_1^N$.

(ii) Light Policy: Class L gets priority at t_0^{N+}. The form of policy is $u_H^*(t) = 1$, $u_L^*(t) = 1$ for $t_1^{N-1} < t < t_{so}^N$; $u_H^*(t) = 0$, $u_L^*(t) = 1$ for $t_{so}^N \le t < t_0^N$; priority to class L for $t_0^N \le t < t_1^N$.

Finally, $u_H^*(t) = 1$, $u_L^*(t) = 1$ for $t \ge t_1^N$.

Unlike Proposition 3 and Theorem 1 for the lost demand model, the identity of priority class in the returning demand model does not follow directly from index $p_i\pi_i$. For the lost demand model, $p_L\pi_L > p_H\pi_H$ meant that class L will get priority. This is not necessarily the case in the returning demand model. We will use our computational experiment to develop further insight into this issue and report them in §5.

An interesting extension of the returning demand model is to consider the case where the return rate (say q_i) of the unmet demand B_i is different from the regular return rate p_i. The formulation needs only marginal modification. We can show that some of the basic results, as in Propositions 1 and 2, will still hold. The analysis in the boundary interval is not equally straightforward because it requires us to consider the integrated impacts of four different return rates.

5. Computational Results

This section uses our model to develop insights into several issues about understanding the policy behavior, the effect of parameters on the policy, and other managerial implications. These insights are based on an experiment carried over a range of parameters and are presented here by using a sample example or a selected graph from the complete experiment.

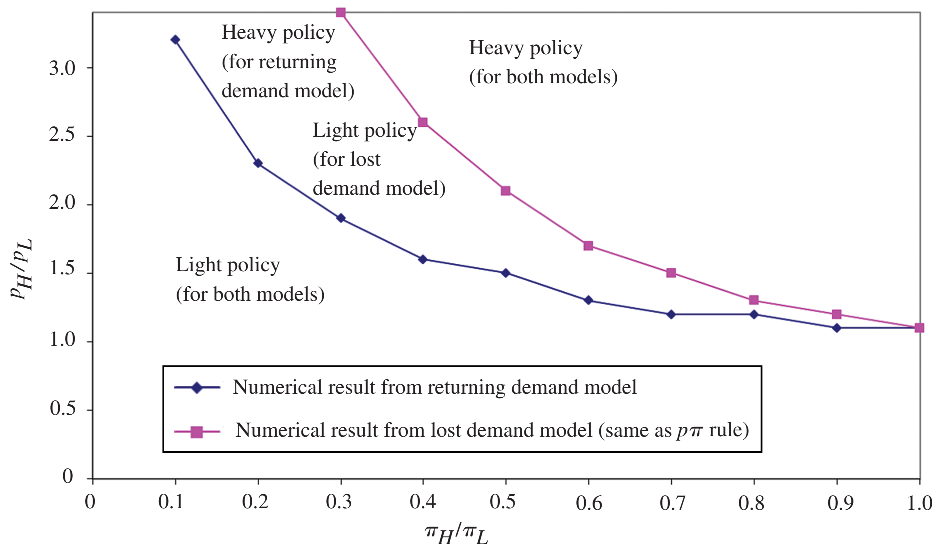
In the experiment, we use the normalizing values of $p_L = 0.25$, $\pi_L = 0.75$. For class H, we use $p_H/p_L = \{1.1, 1.2, 1.3, \ldots, 3.4\}$, $\pi_H/\pi_L = \{0.1, 0.2, 0.3, \ldots, 1, 1.1, 1.2, \ldots, 2\}$. To select demand functions for our experiment, we tried a variety of options. For functions that do not satisfy the single-boundary-interval condition, we used our analytical results for the last boundary interval in combination with exhaustive search (by discretizing time) to determine the optimal control. Based on our numerical tests, we can report that the only demand functions that led to multiple boundary intervals were nondifferentiable or piecewise differentiable. Continuously differentiable demand functions, such as combinations of negative exponential curves, always led to single boundary intervals. We would like to use this computational exercise to demonstrate the applicability of our results beyond the demand functions for which they have already been analytically proven. Therefore, we chose demand functions that are not linear, concave, or exponential but are combinations of exponential functions. This also allows us to choose demand functions that bring us closer to the spirit of our motivating context, DVD rental. That context is characterized by the existence of a preference "list" of movies maintained by the customers. To consider the demand for a new title, we model the number of customers with that title at the top of their lists as $N_i(t)$. At time t, $p_iN_i(t)$ of these customers will return the title they have and demand the title at the top of their lists; that is, $d_i(t) = p_iN_i(t)$. We model the number of customers adding the title to the top of their lists as an exponentially decreasing function of time, $A_i(t) = a_iA_{0i}e^{-a_it}$, where A_{0i} represents the total additions to the top of the lists after $t = 0$. This gives $dN_i(t)/dt = A_i(t) - p_iN_i(t)$. The solution to this equation completes the specification of $d_i(t)$. This boils down to solving $N_i(t) = ((a_ip_iA_{0i}e^{-a_it}/(p_i - a_i)) + (p_id_i(0)e^{-p_it}(1 - (a_i/p_i) - (a_iA_{0i}/d_i(0)))/(p_i - a_i)))/p_i$ for $N_i(t)$ and setting $d_i(t) = p_iN_i(t)$. In our experiment, we set $A_{0i} = 1{,}000$, $a_i = 0.01$, and $d_i(0) = 1{,}000$.

5.1. Impact of Modeling Returning Demands

In Figure 3, we identify the optimal policy ("Heavy" or "Light") as the ratio of penalty costs, and the ratio of return fractions are varied. The demand parameters for two classes are the same. For a given π_H/π_L, there is always a critical p_H/p_L that divides the policy space. Clearly, as we increase the return fraction for class H, keeping everything else the same, the benefit of a quick return increases while the costs of penalty remain the same. If, for some p_H, we prefer the Heavy policy to Light policy, a higher p_H will only increase the attractiveness of the Heavy policy.

The top curve represents the optimal policy in the basic model with lost demand. Each point at the top curve satisfies $p_H\pi_H = p_L\pi_L$. Below it, $p_H\pi_H < p_L\pi_L$ and as predicted by our results, Light policy is optimal. Above it, $p_H\pi_H > p_L\pi_L$, and Heavy policy

Figure 3 Preferred Policy Type for a Range of Parameter Values



Note. $p_L = 0.25$, $\pi_L = 0.75$, $M = 300$.

is optimal. The bottom curve represents the separation between policies in the returning demand model. When both ratios are 1, the two classes are equal and, in either model, giving priority to either class will have the same cost.

As we discussed in §4, returning demand model does not allow us to predict the identity of the priority class based on a simple function of p_i, π_i. Computationally, we observe that the range over which the Heavy policy is preferred with returning demand is larger than the preference range of Heavy policy without returning demand. Intuitively, this is because returning demand model introduces an additional reason to prefer the high-return-fraction class with p_i playing a bigger role here than it did in the lost demand model. In returning demand model, each time a demand (either new or returning) is not met it will return offering another occasion to incur a penalty cost. As $p_H > p_L$, an unmet class H demand will return faster than an unmet class L demand. Since demand is decreasing over time, a demand farther in future is more likely to arrive after boundary interval is over and, therefore, is less likely to incur a penalty cost. This argues for further favoring class H in returning demand model.

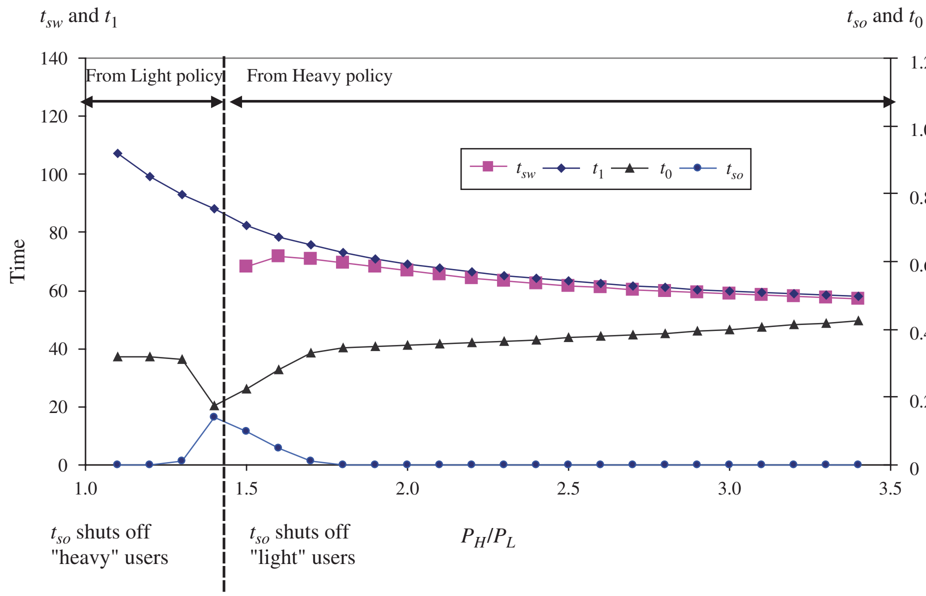
5.2. Time Epochs in Optimal Policy

Beyond the identification of Light or Heavy policy, the four time epochs, t_{so}, t_0, t_{sw}, t_1 play an important role in the optimal policy. Figure 4 depicts the behavior of these four time epochs as the ratio of return fraction is changed. Again, the demand parameters for two classes are the same. As the return fraction increases, t_1 occurs earlier and t_0 (with one exception) occurs later. The unmet demands are accumulated between these two epochs, and an increase in the return fraction should indeed suggest a smaller accumulation period. There is a threshold value of class H return fraction, which separates the Light policy optimal area to its left from the Heavy policy optimal area to its right. When the Light policy is optimal, there is no t_{sw}.

In the Light policy optimal range, when p_H is small and almost equal to p_L, there is no benefit in giving the unit to class H and therefore, the optimal policy shuts out class H by fixing t_{so} at zero. As p_H incrementally increases, t_{so} increases as the period $[t_{so}, t_0]$ over which class H is shut off decreases to almost zero at the boundary between Light and Heavy policies. In the region where the Heavy policy is optimal, the same behavior is observed in reverse. As p_H increases, the relative value of giving priority to class H increases, and the period $[t_{so}, t_0]$ over which class L is shut off increases. The switching time t_{sw} tracks t_1 very closely, indicating that there is a finite but very small period before t_1, where class L gets priority. With increase in p_H, there is comparatively less benefit in giving priority to class L, and therefore period $[t_{sw}, t_1]$ is further squeezed. As we discussed in §3.4, the gap between t_{sw} and t_1 depends on relative values of p_i, π_i parameters for two classes and the shape of demand curves.

Figure 4 shows an intuitive continuity in the policy, especially through t_{so}. In particular, starting from the left as p_H increases, it makes sense to give better service to heavy users so the length of time class H is shut off (difference between t_{so} and t_0) decreases, and eventually, as we observe in our experiments, t_{so}

Figure 4 **Time Epochs in Optimal Policy**



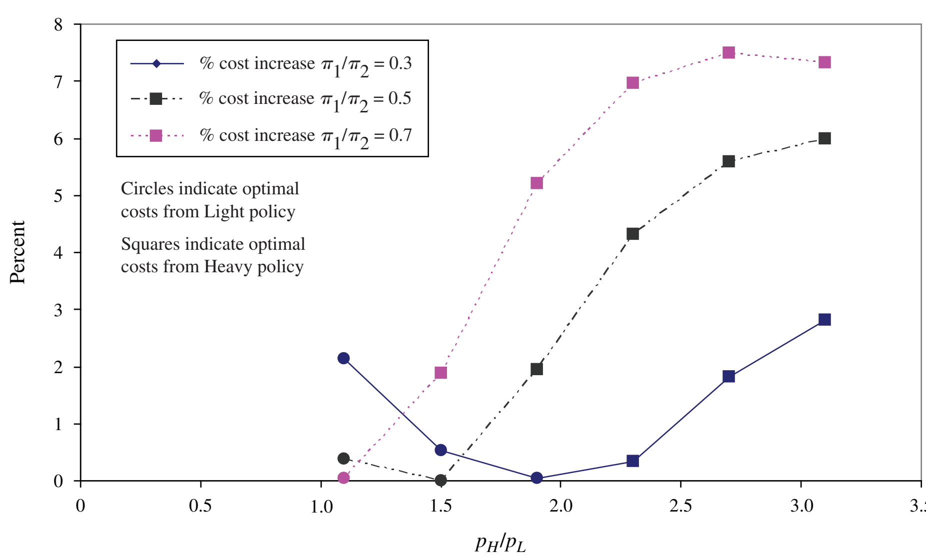
Note. $p_L = 0.25$, $\pi_L = 0.75$, $\pi_H = 0.375$, $M = 300$.

can exactly equal t_0, at which point there is essentially no rationing until the firm runs out. Then, as p_H increases further, the policy switches to prioritizing heavy users, and the shut-off point for light users decreases toward zero.

5.3. Optimal Policy vs. Throttling

We discussed Netflix's practice of throttling in the introduction. An interpretation of that policy in this context is to always give priority to class L. In our model, this is equal to never shutting off any class in the interior interval, and giving priority to class L in the boundary interval, like a Light policy without the shut-off time. It is not difficult to use our computational infrastructure to determine the cost of throttling policy. Clearly, optimal policy will perform better, but we are interested in understanding the margin by which optimal policy outperforms a practically reasonable policy like throttling. Note that in the comparisons shown in Figure 5, the policies are evaluated at their corresponding values of M that minimize $cM + \mathrm{C}(M)$ where $\mathrm{C}(M)$ is the objective function in §4, and c is per-unit purchase cost.

Figure 5 **Cost Increase Due to Throttling Policy**



Note. $p_L = 0.25$; $\pi_L = 0.75$; $d_i(0) = 1{,}000$; $a_i = 0.01$; $A_{0i} = 1{,}000$; $c = 3$.

Figure 5 plots the percentage by which the throttling cost is higher than the optimal cost and shows how this difference changes with p_H/p_L and π_H/π_L. On each graph, circular dots are the cases where optimal policy is the Light policy, and rectangles show the cases where Heavy policy is optimal. Even when the Light policy is optimal, the throttling cost is still higher because the optimal policy may shut off class H for some time. When Heavy policy is optimal, throttling clearly performs much worse. As p_H/p_L increases, Heavy policy gets better, and the performance of throttling policy drops. Interestingly, the performance of the throttling policy also drops as p_H/p_L decreases because at lower p_H/p_L, shutting off class H plays a more important role because of the lower quick return benefits. The performance of throttling policy also drops on average with an increase in π_H/π_L. To summarize the differences, throttling as described in reports focuses only on the higher penalty cost of class L but ignores several other important influences that are captured in this paper. Those are (i) the role of quicker return characteristic of class H, (ii) the decreasing nature of demand over time that determines the shut-off and switching times in the optimal policy, and (iii) the role the class characteristic plays in determining the shape of the demand curve.

6. Conclusions and Future Work

We consider the impact of decreasing demand on the optimal allocation policy in a two-class rental model. Our analytical results show that the optimal policy is a priority policy, but the identity of the priority class may change over time. We believe that our results add to the theory by introducing the notions of switching time and shut-off time in the context of a decreasing demand model.

Based on the reporting of a real allocation decision at a video rental firm, we interpret our model to capture the main complexities involved in that situation: the rental process, the decreasing nature of demand over time, different customer classes, returning customers, and preference list-based demand. This interpretation allows us to draw several managerial implications of our model.

Several issues remain to be explored. On the technical side, it will be useful to analyze the stochastic version of the model. On the managerial side, our operations model can be embedded in a larger business model to address other contemporary issues of interest: satisfaction of customers and corresponding churn rates, and the alternatives to the current subscription-based pricing.

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/msom.2014.0515.

Acknowledgments

The authors are very grateful to the reviewers and the associate editor for not only making many clever suggestions but also for being very constructive in the review process.

References

Ainslie A, Drèze X, Zufryden F (2005) Modeling movie life cycles and market share. *Marketing Sci.* 24(3):508–517.

Albright SC (1977) Structure of optimal policies in complex queuing systems. *Oper. Res.* 25(6):1020–1027.

Altman E, Jimenez T, Koole G (2001) On optimal call admission control in a resource-sharing system. *IEE Trans. Comm.* 49(9): 1659–1668.

Bassamboo A, Randhawa RS (2007) Optimal control in a Netflix-like closed rental system. Working paper, Northwestern University, Evanston, IL.

Gerchak Y, Cho RK, Ray S (2006) Coordination of quantity and shelf-retention timing in the video movie rental industry. *IIE Trans.* 38(7):525–536.

Griffith E (2014) Rent the Runway unveils a Netflix subscription for your closet. *Fortune* (July 16), http://fortune.com/2014/07/16/rent-the-runway-unlimited-netflix-subscription-closet/.

Hartl RF, Sethi SP, Vickson RG (1995) A survey of the maximum principles for optimal control problems with state constraints. *SIAM Rev.* 37(2):181–218.

KOMO News (2006) "Throttling" angers Netflix heavy renters. (February 10), http://www.komotv.com/news/printstory.asp?id=41807.

Lehmann DR, Weinberg CB (2000) Sales through sequential distribution channels: An application to movies and videos. *J. Marketing* 64(3):18–33.

Lilly P (2008) Netflix subscribers frustrated by Blu-Ray shortage, light users move to front of the line. *Maximum PC* (December 16), http://www.maximumpc.com/print/4579.

McDowell M (2014) Jewelry rental websites offer more bling for fewer bucks. *San Francisco Chronicle* (February 1). http://www.sfgate.com/style/article/Jewelry-rental-websites-offer-more-bling-for-5193483.php.

Miller BL (1969) A queueing reward system with several customer classes. *Management Sci.* 16(3):234–245.

Miller CC (2014) Is owning overrated? The rental economy rises. *New York Times* (August 29), http://www.nytimes.com/2014/08/30/upshot/is-owning-overrated-the-rental-economy-rises.html?_r=0&abt=0002&abg=1.

Olney J (2013) Startups use Netflix rental model for clothing, jewelry. *ABC News* (May 20), http://abclocal.go.com/story?section=news/assignment_7&id=9107413.

Ormeci EL, Burnetas A, van der Wal J (2001) Admission policies for a two class loss system. *Stochastic Models* 17(4):513–539.

Pasternack B, Drezner Z (1999) The videotape rental model. *J. Appl. Math. Decision Sci.* 3(2):163–170.

Rosales E (2006) Heavy Netflix users pushed back in line: Policy in fine print. *ABC News* (February 10), http://abclocal.go.com/kgo/story?section=local&id=3898425.

Savin S, Cohen M, Gans N, Katalan Z (2005) Capacity management in rental business with two customer bases. *Oper. Res.* 53(4):617–631.

Sethi SP, Thompson GL (2000) *Optimal Control Theory: Applications to Management Science and Economics*, 2nd ed. (Kluwer Academic Publishers, Boston).

Stenovec T (2014) This is why you won't see Oscar blockbusters streaming on Netflix. *Huffington Post* (blog) (March 3), http://www.huffingtonpost.com/2014/03/03/netflix-oscar-movies_n_4892961.html.

Stross R (2007) Pass the popcorn. But where's the movie? *New York Times* (August 19), http://www.nytimes.com/2007/08/19/business/yourmoney/19digi.html?_r=0.

Tainiter M (1964) Some stochastic inventory models for rental situations. *Management Sci.* 11(2):316–326.

Tang CS, Deo S (2008) Rental price and rental duration under retail competition. *Eur. J. Oper. Res.* 187(3):806–828.

Whisler WD (1967) A stochastic inventory model for rented equipment. *Management Sci. Ser. A* 13(9):640–647.