The Consumer's Rent vs. Buy Decision: The Case of Home-Video

George Knox Jehoshua Eliashberg*

Marketing Department
The Wharton School,
University of Pennsylvania,
Philadelphia, PA 19104

{gknox, eliashberg}@wharton.upenn.edu

February 2005

*George Knox is a doctoral candidate and Jehoshua Eliashberg is the Sebastian S. Kresge Professor of Marketing, at the Wharton School, University of Pennsylvania. Address: 700 Huntsman Hall, 3730 Walnut Street, Philadelphia PA 19104. Corresponding author: George Knox, Email: gknox@wharton.upenn.edu, Tel: (215) 573-4831, Fax: (215) 898-2534. We thank TLA Entertainment for graciously providing us with the data used in this study and the help of Eliot Jarrett in organizing the data. Support from the Wharton e-Business Initiative (WeBI) is gratefully acknoledged. We also thank Eric Bradlow and Ye Hu for suggestions concerning the analysis, as well as to participants at the 2004 Marketing Science Conference in Rotterdam for helpful comments.

The Consumer's Rent vs. Buy Decision:

The Case of Home-Video

Abstract

The home video market represents more than $10 billion in sales, currently more than half the revenues of a major motion picture release. Increasing interest has turned to how consumers decide whether to rent or buy a particular movie title. In this paper, we develop and test two plausible models describing the consumer's decision to rent or buy a given title. We utilize a unique transaction data set obtained from a large video store chain to estimate relevant parameters and make inferences about the rent/buy behavior of individuals and the rentability/purchasability of movies. We use these estimates to demonstrate how the retailer could generate customized purchase prices for consumers based on their transaction histories.

We find that accounting for diminishing value of the video adds significant predictive power, both in and out of sample, to the model. To a large extent, observable characteristics of movies do not explain well differences in movie quality. We also find, in line with anecdotal evidence reported in the popular press, that movies which belong to the action genre and are rated R are more likely to be bought.

Key Words: Entertainment Marketing, Choice Models, Real Options, Customized Pricing

Word Count: 11,959

1 Introduction

There has been a growing interest in research and modeling issues related to the entertainment industry in general and the movie industry in particular. Most of the extant literature has focused on distribution (Litman and Ahn, 1998; Eliashberg, Jonker, Sawhney and Wierenga, 2000) and theatrical exhibition (Jones and Ritz, 1991; Eliashberg, Swami, Weinberg, and Wierenga, 2001) of feature films in the U.S market (De Vany and Walls, 1997; Radas and Shugan, 1998; Ravid, 1999; Orbach and Einav, 2001) or in foreign markets (Neelamegham and Chintagunta, 1999; Elberse and Eliashberg, 2003). Much less research effort has been made in studying the home-video market, and in particular, the consumer decision making in this space.

In 2003, the home-video market represented a $12 billion market, an average of 52% of the revenue for a major motion picture release, compared to approximately 1% only 8 years earlier (Economist, 2004). This amount exceeds the U.S. theatrical feature films market ($9.4 billion). The significance of the home video market has been recognized by the studios, as manifested, for example, at MGM, Universal, and New Line Cinema, where home-video sales executives now exhibit more power in the initial stage of green-lighting new projects (Kirkpatrick, 2003). Some studios see the home video market as the "corporate A.T.M. machine," particularly helpful in rescuing movies that are risky bets at the box office (Gertner, 2004).

A combination of factors, such as the rapid adoption of DVD players (46.7% penetration by the end of 2003) and the affordable pricing of videos for sale have made the home video market grow faster than expected in the last few years. Two of the largest retail chains, Blockbuster Video and Hollywood Video, saw sell-through revenues grow 3.7% while rental revenues shrank by 3.6% (Reinan, 2004). While studios appear to understand that aggregate drivers of success in the home video market are different from the box office (Gertner, 2004), how consumers decide whether to buy or rent a particular movie title is relatively unknown both to entertainment executives and academics.

Companies have addressed the consumer's decision of whether to rent or buy through a variety of business models. Some of them focus only on the rental opportunity, others only on the sell-through opportunity, and yet others are trying to capitalize on both opportunities. Netflix, for instance, represents a DVD online rental only service with 2 million subscribers who create their movie list (from a library of 25,000 titles), receive 3 DVDs from the list by mail, return the viewed movies by mail, and then receive more. This 3-at-a-time rental service costs the consumer $21.99 per month. Columbia House with their DVD Club represents the sell-through only business model. It offers 5 DVDs for $0.49 each for consumers who join the club and who then become obligated to buy 5 more DVDs within the next 2 years, at regular club prices. Blockbuster represents the 'rent and buy' business model with both 'brick and click' presence. On its web site, the subscriber can find a rental service very similar to that of Netflix, while at the same time, the consumer can purchase the same title. Wal-Mart also offers a similar 'brick and click/buy and rent' service. Recently (Reuters, October 15, 2004) Netflix shares tumbled 43% after Amazon.com, known for its only sell-through business model, announced its intention to enter the rental market as well. With advances in technologies such as Microsoft's Symphony, deigned to connect PCs, TVs, and other consumer electronic devices, consumers will be able to download and enjoy multimedia content such as movies in real time, without having to wait for the mail delivery of the content. Their decision to rent or buy will remain, however, an important one. It is the focus of this paper.

In this paper we adopt a micro-level modelling approach addressing this decision, and we calibrate it using actual transactions for a sample of customers subscribed to a particular 'buy and rent' type home-video store over a six-month period. Specifically, we develop and empirically test two models, including two different specifications of the movie effects for each model. We then use the parameters of the best-fitting model to develop relevant managerial implications such as customizing purchase prices to consumers. Unlike most previously published entertainment research studies that take the perspective of the feature film distributor/ studio (e.g., Krider and Weinberg 1998; Neelamegham and Chintagunta, 1999; Eliashberg et al. 2000), we are interested in the home-video retailer's perspective. Our concern is with modelling the consumer choice between renting and buying and estimating managerially relevant individual and movie level parameters from consumers' observed choices. For example, according to Ben Feingold, president Columbia TriStar Home Entertainment, some consumers are “collectors,” while others are merely “watchers” (Gertner, 2004). We can identify such customers. We are also interested in categorizing movies as rentable (e.g., About Schmidt) or buyable (e.g., Lord of the Rings: Two Towers). This work, therefore, fills the gap in the literature on entertainment research, and it is also relevant to practitioners seeking how to match the content with the consumer decision to buy or rent it.

At the heart of any decision to buy or rent is an estimate of how much value - either expected value or number of viewings - the consumer anticipates he or she will derive over subsequent movie viewings. While this estimate is crucial to the consumer’s decision, our data set only allows us to observe choices. Both of the models we develop easily lend themselves to deriving these latent constructs.

In the first model, the consumer decides between investing in an uncertain sequence of future values that diminish over time as the video is watched and a one-shot value that requires less investment. In this case, the consumer’s value of buying is treated as a real option (Dixit and Pindyck 1994; Brennan and Schwartz, 1985), specifically a “machine abandonment” problem. Part of the value the consumer attaches to buying is the value of flexibility from choosing when to stop watching, or “abandoning,” the movie. The consumer estimates total expected utility from buying by averaging over all sequences of values from watching under the optimal policy. She compares the expected utility of buying with a one-shot value from renting and chooses the option that maximizes expected utility.

We compare this model with another model from cognitive psychology that has been used in analytical models in the context of renting and buying. In this model, the consumer is concerned with the number of times a movie is watched, not value or utility. The consumer is modelled as calculating a threshold level of viewings, based on the costs of renting and buying the title (known to the consumer and the modeler). If the consumer’s anticipated number of viewings is greater than the so determined threshold level, the consumer decision is to buy the home-video. Otherwise, s/he will rent it. From the modeler’s standpoint, not being able to observe the anticipated number of watchings, the buy/rent choice is treated probabilistically.

The complexity of the models and the limited observations for some individuals in the data set present an empirical challenge which we overcome by using a Hierarchical Bayes estimation procedure. In a simulation data set which closely mimics the actual data set, we demonstrate the estimation procedure's ability to recover the model's parameters. We adopt (separate) priors over the population of individuals and movies. Using Bayesian estimation techniques allow us to borrow information across individuals and movies. The posterior individual and movie level parameters also provide us with measures of uncertainty, which are useful in assessing forecasting accuracy as well as in the customized pricing task. Predictive ability of both versions of both models is measured both within and out of sample.

In addition to testing the predictive capabilities of the models, another objective of the paper is to obtain diagnostic information explaining the differences in movie sales across individuals, using a number of readily available movie characteristics such as genre, box office grosses, and MPAA Ratings. We contrast the impact of these variables with current industry thinking. For example, studios executives believe that male-oriented action movies do the best in the home video sell-through market (Kirkpatrick, 2003). We provide empirical evidence related to such beliefs. We also study the relative effects, across both models, of the box office gross, genre, and MPAA ratings.

The paper is organized as follows. In Section 2 we review the relevant literature. Section 3 describes the two models. In Section 4 we described the data employed to test the models and their estimation. The empirical results are provided in Section 5. Section 6 demonstrates how a video retailer can customize purchase prices to the consumer using the model. Section 7 summarizes the work and provides suggestions for future research.

2 Relevant Literature

Our analysis draws on three areas of research. The literature on home video release strategy shares our context - the home video market - but posits reduced form aggregate demand models, usually in order to test implications of a theoretical model of home-video distributor (studio) conduct. Also related is the growing literature that models aggregate box office demand using discrete choice methods and conducts policy simulations given these estimates. Lastly, we draw on the literature that examines consumer choice between flat pricing-based unlimited consumption, and per-use pricing plans, when usage or value is uncertain from the consumer's perspective.

When studios release movies sequentially - first on the big screen, next for home video, pay-per-view, etc. - sales in a later channel can partially cannibalize sales in the current channel. The premise of the home video release literature is that movies in the theater and on DVD are partial substitutes in the eyes of the consumer, because a consumer can always decide when to see the movie. For example, Lehmann and Weinberg (2000) use exponentially decreasing demand curves both for box office and the rental market in order to find the optimal time to release a movie in the home video rental market. Movies that are available simultaneously for buying and renting also serve as partial substitutes for each other. A strategic decision on the part of studios is whether to release a video simultaneously to both rental and sales markets, or to release it only to the rental market for a period before making it available for purchase. While approximately 90% of VHS videos are released sequentially - first to the rental market, then to the retail market - DVD's are released simultaneously to both rental and retail markets (Mortimer 2004a, 2004b; Hu, Eliashberg, and Raju 2004). As Mortimer (2004b) shows, studios are in effect choosing whether to use intertemporal price discrimination to segment high and low value customers. She estimates a model where purchasing and renting the same movie are vertically differentiated products, a la Shaked and Sutton (1982), which explains why studios choose to adopt different pricing policies for the same movie on two different formats. She then uses the parameter estimates to simulate a few policy experiments on DVD adoption patterns and copyright law changes. Hu et al. (2004) develop a model of the consumer that incorporates heterogeneity and forward looking behavior. They find that the optimal release strategy is influenced by both movie characteristics and heterogeneity in consumers' expected number of times to watch the video. The consumer level model draws on Varian (2000), who examines whether firm without the ability to intertemporally price discriminate should price to sell or rent an information good.

Our analysis differs substantially from these papers in that we develop and empirically test an individual level model from transaction level data. This affords a better look at how consumers actually choose between buying or renting a video. For example, both Varian (2000) and Hu et al. (2004) model the consumer as using a reduced form decision rule which is reasonable from the consumer's point of view if the value at each watching remains the same and is precisely known beforehand to the consumer. We compare the fit of such a model with a structural model in which the value stochastically diminishes as the consumer watches the video more. Another important distinction is that we take the view of the retailer, and not the studio; we forecast individual level choice, not aggregate demand, and demonstrate how a retailer could use these estimates to customize purchase prices based on consumer transaction history (in the spirit of Rossi, McCulloch and Allenby, 1996). To our knowledge, this is the first paper in the entertainment modeling literature that takes the perspective of the video retail store.

While there are many papers that estimate box office demand for motion pictures (for example, Sawhney and Eliashberg 1996, Krider and Weinberg 1998, Radas and Shugan 1998, Neelamegham and Chintagunta 1999, and Elberse and Eliashberg, 2003), more recently some researchers have turned to discrete choice estimation methods to model competition between movies (Ainslie, Drèze, and Zufryden, 2004), competition between multi-product studios (Moul, 2004), and underlying seasonal effects (Einav, 2004). One characteristic these papers share is that they model aggregate market shares using individual level choice models. Ainslie et. al. (2004) create a sliding window logit model to predict weekly movie market shares. They compare this to a demand model that treats each movie in isolation and find that accounting for competition via the logit model increases fit dramatically. Einav (2004) uncovers underlying seasonal demand from the observed seasonal pattern of sales by accounting for the number and quality of movies using a nested logit model of weekly demand. Moul (2004) uses a modified principles of differentiation generalized extreme values (PD GEV) to model weekly demand and market power between studios and theaters. The modeling approach taken in this paper is similar in spirit to Neelamegham and Jain (1999), who develop an individual level discrete choice model that accounts for various psychological variables and is estimated on individual level data. [1] While we share with these models the use of discrete choice models to estimate movie parameters and evaluate policy changes based on these estimates, our data allows us to examine actual individual level choices rather than aggregate market shares.

The decision of whether to buy or rent shares characteristics with a broader, more abstract set of decisions encountered in daily life: whether to pay a flat fee for unlimited consumption or to pay at each time of consumption. Some common examples of firms that offer both flat-rate and pay-per-use options found in the literature include health club membership, online grocery shopping, and telephone calling plans. The goal of this literature is to draw conclusions from initial contract choice and subsequent usage. Here, the literature has found a consumer bias for the flat fee in most cases and a bias for the pay-per-use in a few cases (for a good overview of the literature and explanations for these biases, see Lambrecht and Skiera, 2004). For example, Della Vigna and Malmendier (2003) show that gym members who choose the flat monthly fee end up paying more per visit than the offered per-visit price. They model the consumer as forward-looking, but with time-inconsistent preferences, and examine consumers' sequence of price and usage choices. Miravete (2002) finds a bias toward the pay-per-use option based on transactional data from a tariff experiment. Nunes (2000) presents a descriptive model of a consumer who uses simplified heuristics to estimate usage in the context of online grocery shopping, gym membership, and swimming pool fees. In his model, a consumer estimates a threshold and matches the probability that usage exceeds the threshold with the probability of choosing the flat-fee option. This helps explain why consumers habitually overestimate their usage and choose the flat fee option. Another explanation for the flat fee bias found in the literature is that unlimited consumption includes an option value if consumers are uncertain of their preferences (Kridel, Lehman and Weisman, 1993). This explanation is more in line with how we model the buyer in the first model.

[1] Their concern, however, is with modelling pre-choice and post-choice expectations of movies which have just been released in the theater, which they validate by conducting an experiment.

3 Context and Models

The scenario we model is a segment of consumers for whom renting and buying a title are mutually exclusive and exhaustive options. [2] We present two models of varying complexity meant to capture the consumer's choice between buying or renting a particular movie. Consumers assess the value of buying in one of two ways. In the first model, the consumer understands that the value of watching a movie diminishes over time and treats the problem of how many times he should watch the movie as an optimal stopping problem. S/He estimates expected utility of buying by averaging the values received from watching under the optimal policy and chooses the option (rent or buy) that maximizes total expected utility. In the second model, the consumer estimates a threshold level from the costs of renting and buying. The consumer then estimates whether or not the number of times the movie is watched will be greater than the threshold. We present both models in more detail below.

3.1 Model I

Every time a movie is watched, the consumer receives a value. Our model of values is directly motivated by two observations. First, from the perspective of the consumer, the value of watching a movie again is not certain. That the consumer chooses to watch a movie again at all implies that there is still enough value to the experience. We assume that this value is stochastically related to the previous value. We accomplish this by making the value of watching the movie now equal to the value of watching it last time plus a random term that may be either positive or negative. The second observation is that viewers on average derive less value from watching a movie each time they watch it. The value of the experience on average decays over time, and in practice this decay may vary by movie type. For certain "classics" like Lord of the Rings, the decay may be slow; for movies like Daredevil and 2

[2] Another segment that may exist in the market place is comprised of consumers for whom there are three options: (1) rent only, (2) buy only, and (3) rent first and buy next. This segment, if it exists, is not well represented in our data. Of over 17,000 transactions, only 713 involve consumers transacting with the same movie multiple times (renting, then buying; renting more than once; buying more than once).

Fast 2 Furious, the utility of future enjoyment may decay much more rapidly.

In line with our two observations, we model the value to a consumer of watching a particular movie (suppressing subscripts) for the t^{th} time as the value of the movie the previous time plus a random term.

$$X_t = X_{t-1} + \epsilon_t \tag{1}$$

For tractability we assume the random components at each viewing, ϵ_t, are independent and identically distributed according to a normal distribution with mean $-\mu$ and variance σ^2. [3,4]

From the consumer's perspective at the video store, a particular sequence of movie viewing valuations is a realization of the random walk outlined in equation 1. The consumer's problem is whether to pay an amount k (roughly \$3) for a one-shot value of X_0 or to invest a larger amount $K > k$ (roughly \$20) for a stochastically diminishing sequence of values $X_0, X_1, \ldots, X_T$. While the utility from renting is simple, the consumer needs to calculate the utility of buying in two steps. First, she chooses when to stop watching, T, by following an optimal policy, which is to calculate a threshold based on discounted future movie valuations. Once this threshold has been calculated the consumer estimates expected utility of buying by summing over all possible sequences of valuations above the threshold. [5]

In order to maintain tractability, we need only to modify the above process slightly

[3] This process has been suggested as a "general law" of clicks on the web, among other scenarios (Lukose and Huberman, 1998; Huberman, Pirolli, Pitkow, and Lukose, 1998)

[4] Our formulation nests the case where there is no uncertainty ($\sigma^2 \longrightarrow 0$), though the empirical application strongly rejects this.

[5] The proposed model assumes a sophisticated forward-looking consumer capable of calculating this threshold and estimating expected total utility of watching until this threshold is reached. This is in line with the consumer models of Erdem and Keane (1996) and Gonul and Srinivasan (1996), which assume that consumers solve dynamic problems as optimal decision makers. For example, Gonul and Srinivasan (1996) model the consumer as optimally planning diaper purchases over likely coupon drops in the future. Note that these models are for frequently purchased consumer goods with high repeat purchases, i.e., they have multiple household-product observations over time. In our context we do not have repeat individual-movie observations, as mentioned before (see footnote 2). In addition, since we do not observe the consumer's utilization of the video, we choose the optimal stopping rule as a rational first approximation to a consumer's actual behavior.

from a discrete-time, continuous-state random walk to a continuous-time, continuous-state process. In Appendix A, we show that the following continuous-time Brownian motion is a continuous interpolation (i.e., has the same distributional properties) of the discrete time process outlined above in equation 1:

$$\begin{aligned} dX(t) &= -\mu\, dt \,+\, \sigma\, dz \\ X(0) &= X_0 \end{aligned} \tag{2}$$

where dz is a Wiener process, i.e. $dz = \eta\sqrt{dt}$, where η is a unit normal. At each point in time, the distribution of video watching utility to the buyer is:

$$X(t) \sim N(X_0 - \mu\, t\,, \sigma^2\, t)$$

For the consumer, the decision of how long to watch a movie can be characterized as a machine abandonment problem (see Dixit and Pindyck 1994, pg. 110-2, or Brennan and Schwartz, 1985), which is a specific form of an optimal stopping problem (Bertsekas, 2000). The decision, when to scrap a machine whose quality of production stochastically diminishes over time, is binary. In the framework of the current problem, the consumer maximizes enjoyment over the infinite lifetime of the purchased movie by choosing when to stop watching.

Because the buyer's problem is over an infinite time horizon, the solution is a stationary threshold, $\bar{X}$, above which the consumer will continue to derive value from the movie and below which she will decide to stop watching (see Appendix B):

$$\begin{array}{ll} \text{Keep Watching if} & X(t) \,>\, \bar{X} \\ \text{Stop if} & X(t) \,<\, \bar{X} \end{array} \tag{3}$$

In our context, the consumer balances the current cost of watching the same movie again with the future possibility of positive value (seeing some favorite scenes again). When the current value is negative enough so that the costs of continuing to watch outweigh the possible benefits, the consumer stops watching. Due to the random term in the valuation process, an optimal consumer would continue to watch even as the valuations become negative. While in expectation the value of watching is decreasing, there is a chance that the random part of the valuation may drive the process into positive territory. In Appendix B, we show in detail how to derive $\bar{X}$ and we prove that it is non-positive. We can write the boundary as a function of the drift, noise and discount rate:

$$\bar{X} = \frac{\mu}{r} + \frac{\sigma^2}{\mu - \sqrt{\mu^2 + 2\sigma^2 r}} \tag{4}$$

We illustrate the boundary as a function of the decay rate and variance in Figure 1. The boundary $\bar{X}$ is always negative except when the rate of decreasing enjoyment is extremely fast ($\mu \to \infty$), at which point the optimal threshold is zero. Since μ is the rate at which the enjoyment is decreasing, large μ means the costs of watching another time before quitting are greater to the buyer. The boundary becomes shallower the more costly it becomes to watch the same movie again. The variability of the valuations, σ^2, makes the boundary deeper. In the limit, as we completely remove the uncertainty ($\sigma^2 \to 0$), the threshold becomes zero. [6] As σ^2 increases and the process becomes more variable, the viewer becomes more reluctant to stop watching, even if current watching is not giving much value. This is because the noisiness of the process may bump up the value of watching the next time.

Once the consumer determines $\bar{X}$ (in the video store), she finds the distribution of the number of times she will watch the movie in order to calculate the total expected utility of buying. The time it takes the process of equation 2 to reach the boundary of equation 4 is a random variable $\tilde{T}$. A general result to derive the distribution of $\tilde{T}$ has been developed (see Karlin and Taylor 1975, pg. 363, theorem 5.3). They show that for Brownian motion with drift $-\mu \leq 0$ that starts at X_0, the time that the process first reaches $\bar{X}$ is distributed inverse Gaussian with probability density function and expected value:

$$\begin{aligned} f(\tilde{T} = t) &= f(t \mid X_0, \mu, \sigma) = \frac{X_0 - \bar{X}}{\sigma\sqrt{2\,\pi\,t^3}} \exp\left[-\frac{(X_0 - \bar{X} - \mu t)^2}{2\sigma^2 t}\right] \\ \mathsf{E}[\tilde{T} \mid X_0, \mu, \sigma] &= \frac{X_0 - \bar{X}}{\mu} \end{aligned} \tag{5}$$

There is a natural interpretation to $X_0 - \bar{X}$ and μ as the "distance" and "speed" of the movie enjoyment process. A large value of $X_0 - \bar{X}$ means a large capability for enjoying movies. A movie that has a low value of μ travels "slowly," or decays slowly, through $X_0 - \bar{X}$

[6] To see this, an application of L'Hopital's Rule to the second term yields $\lim_{\sigma^2 \to 0} \bar{X} = \frac{\mu}{r} - \frac{r}{\sqrt{\mu^2}} = 0$.

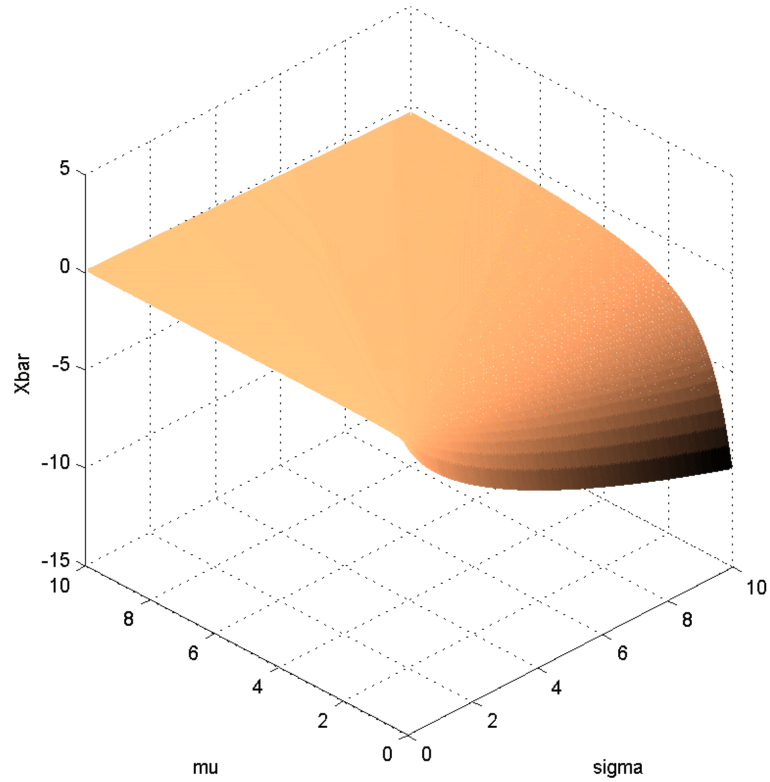


Figure 1: $\bar{X}$ as a function of μ and σ^2 holding r = 0.04 constant

gives more enjoyment, and hence is more likely to be bought. The amount of time one spends enjoying the movie is therefore the distance over the speed, which is precisely the mean in equation 5. In general, the inverse Gaussian has a wider tail than the normal distribution. This implies that there may be a significant number of people who may watch a movie many times, more than predicted by a normal distribution.

The consumer calculates the expected total utility of buying by integrating over all paths of valuations that start at X_0 and end at $\bar{X}$. This is accomplished in two steps. For each stopping time $\tilde{T} = T$, the consumer calculates the average utility by integrating the expected valuation over time. Because $\tilde{T}$ is a random variable, the consumer forms expectations of total expected utility by integrating this average valuation over all possible times. In order to maintain tractability, we keep the initial value X_0 as a discrete amount and leave subsequent valuations as continuous from equation 2:

$$\mathsf{E}[TU_B] = X_0 + \int_{0^+}^{\infty} \int_0^t e^{-r\,\tau}\, \mathsf{E}[X(\tau) \mid X(0)]\, d\tau\, f(t)\, dt \;-\; K \tag{6}$$

Some straightforward algebra presented in Appendix C shows the above integral can be written as:

$$\mathsf{E}[TU_B] = X_0 + \left(\frac{X_0}{r} - \frac{\mu}{r^2} \right) (1 - \phi) + \frac{\mu \, \phi \, (X_0 - \bar{X})}{r \, \sqrt{\mu^2 + 2\sigma^2 r}} - K \tag{7}$$

where $\phi(r \mid X_0, \mu, \sigma) = \int_0^\infty e^{-rt} f(t) dt = \mathsf{E}[e^{-r\tilde{T}}]$ is the Laplace transform of the inverse Gaussian random variable $\tilde{T}$.

As mentioned before, the renter's total utility is just the one-shot value X_0 minus a lower fee $k < K$. This one-shot value is the initial utility the consumer receives during the rental time window (typically 2-3 days). [7]

$$\mathsf{E}[TU_R] = X_0 - k \tag{8}$$

If we take the difference of equations 7 and 8, we see that the first term X_0 drops out from both sides.

$$\mathsf{E}[TU_B] - \mathsf{E}[TU_R] = \left(\frac{X_0}{r} - \frac{\mu}{r^2} \right) (1 - \phi) + \phi \, \frac{\mu \, (X_0 - \bar{X})}{r \, \sqrt{\mu^2 + 2\sigma^2 r}} - K + k \tag{9}$$

Since T is always positive, the Laplace transform is $0 < \phi < 1$ and acts as a weight between receiving a perpetually decreasing stream of values (never stopping) in the first term and the value of stopping when the process becomes too negative in the second term. [8] So the difference in expected utility of renting and buying is a weighted sum between a movie paying a decreasing amount into the future and the option value of stopping.

Denote subscripts i for individual and j for movie. There are four parameters in the model: X_0, μ, σ^2, r. We allow X_0 to vary over individuals i (henceforth X_{0i}), μ to vary over movies j, σ^2 to remain the same across people and movies, and finally r is held fixed at 0.04.[9] X_{0i} as the starting point of the movie valuation varies across individuals for all

[7]An extension to the model would be to propose a finite horizon machine abandonment problem as the renter's problem, though this would make the problem intractable and much harder to estimate.

[8]The value of stopping, $\frac{\mu \, (X_0 - \bar{X})}{r \, \sqrt{\mu^2 + 2\sigma^2 r}}$ is always positive, because the stopping region occurs when the value of watching is non-positive, and so to reconcile the non-stopping process with the stopped process we have to add this positive factor.

[9]Since the discount rate is continuous, this corresponds to a mild discrete per-period discount rate of $\delta = e^{-0.04} = .96$. We tried other mild discount rates of 0.10, 0.05, and 0.01, all of which led to slightly worse off log-likelihood.

movies because certain customers are more likely to buy any videos overall, perhaps because they are collectors and are building a library. This parameter allows us to quantify how much a consumer is a "collector" or "watcher." A higher (lower) value of X_{0i} corresponds to an individual who is, ceteris paribus, more likely to buy (rent). The parameter μ_j is the speed at which value to the watcher diminishes. A higher (lower) value of μ_j corresponds to a movie which is more likely to be rented (bought).

We convert expected utilities into choice by adding a standard Gumbel (extreme value EV (0,1)) distributed error that is independent and identically distributed over individuals, movies, and choices. This error has the usual interpretation of being observed by the consumer but not by the analyst. Then we can write the probability of individual i buying movie j as:

$$P_{ij}(\text{Buy}) = \frac{\exp\{\mathsf{E}[TU_{ijB}] - \mathsf{E}[TU_{ijR}]\}}{1 + \exp\{\mathsf{E}[TU_{ijB}] - \mathsf{E}[TU_{ijR}]\}} \tag{10}$$

where from equation 9:

$$\mathsf{E}[TU_{ijB}] - \mathsf{E}[TU_{ijR}] = \left(\frac{X_{0i}}{r} - \frac{\mu_j}{r^2}\right)(1 - \phi_{ij}) + \phi_{ij}\,\frac{\mu_j\,(X_{0i} - \bar{X}_j)}{r\,\sqrt{\mu_j^2 + 2\sigma^2 r}} - K_{ij} + k_i \tag{11}$$

where $\phi_{ij} = \phi(r \mid \mu_j, \sigma, X_{0i})$.

In section 4, we report the estimation results from a simulated data set that closely mimics the actual transaction data set to show how this model can uncover parameter estimates. Since the coefficient of price is one, all units of analysis are in dollar terms.

3.2 Model II

The complexity of the last model calls into question whether consumers are capable of making such forward-looking calculations. As an alternative, we develop a simpler model based on a cognitive process model from psychology. This modeling framework, applicable to the general context of a choice between paying at each use and paying a flat fee, has been employed by Varian (2000), Nunes (2000), and in the specific context of the rent vs. buy (home video) choice by Hu et al. (2004). Accordingly, the consumer is concerned with the number of times a movie is watched, without regard to the decaying utility of multiple watchings. As in Nunes (2000), the consumer calculates a threshold level based on the cost of buying, K_{ij}, and the cost of renting, k_i. Under Model II, the consumer is assumed capable of estimating whether the number of times the movie will be watched is greater or less than the threshold:

$$\begin{array}{lll} \text{Buy} & \text{if} & N \geq \frac{K}{k} \\ \text{Rent} & \text{if} & N < \frac{K}{k} \end{array} \tag{12}$$

Comparing this threshold with that of Model I (equation 4), the threshold of Model II is exogenous and based on a simple ratio of store prices. While the consumer in Model II thinks in terms of raw number of times a movie is watched, the consumer of Model I thinks about the sequence of values from watching a movie.

In Model II, while the consumer is capable of estimating whether N is greater or less than a threshold, from the perspective of the modeler, N, the number of times up until the final watching of the the movie, is a random variable. We therefore estimate the probability of that this random variable is greater than the threshold by assigning N a distribution. We choose the geometric distribution as a reasonable modeling approximation for parsimony and interpretability. It has only one parameter, p, interpreted as the probability that the next watching is the final one. The break even number is the ratio of buy to rent price, which we denote $\tau_{ij} = \text{floor}[K_{ij}/k_i]$. In the empirical application we make p_{ij} a function of individual i and movie j covariates. Using some simple tools from a geometric series we can write (see Appendix D):

$$P_{ij}(\text{Buy}) = \sum_{x \geq \tau} p_{ij} \left(1 - p_{ij}\right)^x = \left(1 - p_{ij}\right)^{\tau_{ij}} \tag{13}$$

A higher (lower) value of p_{ij} means that the consumer will stop watching the video sooner (later). This naturally pushes the probability of buying lower (higher).

4 Data and Estimation

In this section, we briefly discuss the nature of the data we use to calibrate both models, how we estimate both models, and a simulation study for model I.

4.1 Data

The data we use to calibrate both models come from TLA Entertainment Group, a home-video retail chain which operates 6 stores in Philadelphia and New York City, rents and sells a theatrical and direct to consumer film labels, and produces two film festivals each year. In 2002, TLA was ranked 15^{th} in revenue across the U.S. for specialty stores (exclusively DVD/VHS) according to Video Store Magazine. Each store carries at least 20,000 movies for rent and 2,000 movies for sale.

Titles not available in the store for immediate sales can be ordered and shipped to subscribers within one week. The data in our sample come from the largest store, and cover the last two quarters of 2003. Within this store TLA maintains two separate areas, one for renting videos, the other for buying videos.

Of the 75,000 transaction records, containing over 9,000 movies and 5,000 members, we chose to focus on 76 most transacted movies over the 6 month time horizon of our data. A list of the movies used in the analysis is provided in Appendix E. These movies were available at the same time for renting and buying within the store (simultaneous release).

We only focus on DVD transactions for two reasons. First, DVD's are priced to sell in the \$10 - \$25 region compared to VHS tapes which can be priced as high as \$100. Because studios are more inclined to adopt sell-through pricing for DVD's than VHS (cf, Mortimer 2004, Hu et al. 2004), the DVD price makes the option of purchase more reasonable to the consumer. Secondly, given the scenario we model, both options - renting and buying - have to be available at the same time for the choice to be legitimate. DVD's are most commonly released for rental and purchase simultaneously, whereas VHS tapes are more often released sequentially.

A note on the prices is in order. K_{ij} is captured in the data as the variation in purchasing prices for movies over time. That is, time is nested within the individual-movie combination, depending on when the individual transacted with the movie. The price of renting, k_i, is different for consumers depending on whether they advance purchase rentals or not. The price of renting a movie was \$3.50 in our time horizon, but the range of prices we observe is \$2.50 - \$3.50. This is because TLA gave consumers mild quantity discounts if they pre-purchased rentals in bulks of 10, 20, or 50. If a customer pre-purchased 50 rentals, her cost of renting would be \$2.50. In our data, even though consumers do not pay to rent if they have already pre-purchased, TLA reports the price at the transaction.

Restricting ourselves to 76 movies provides us with over 24,000 observations from 4,279 customers, with only 343 recorded buys from 193 customers. Of these 4,279 customers, we estimate the model on 500 of them. Since the number of customers who have bought at least once is small in the stores population, we include all 193 customers who buy at least once, and randomly sample 307 customers from the non-buying population. The final data set we use includes 2,913 transactions from 500 customers. Of these 500 customers, 160 bought a movie at least once in the 6 month time period, and 112 both bought and rented at least once. [10] These 112 customers are the most promising in terms of "converting" more of their rentals into purchases. We focus on this subsegment of customers in the customized pricing section. The remaining 340 in our sample only rented. We present histograms of the number of purchases and rentals across consumers in Figures 2. We display the analogous histograms

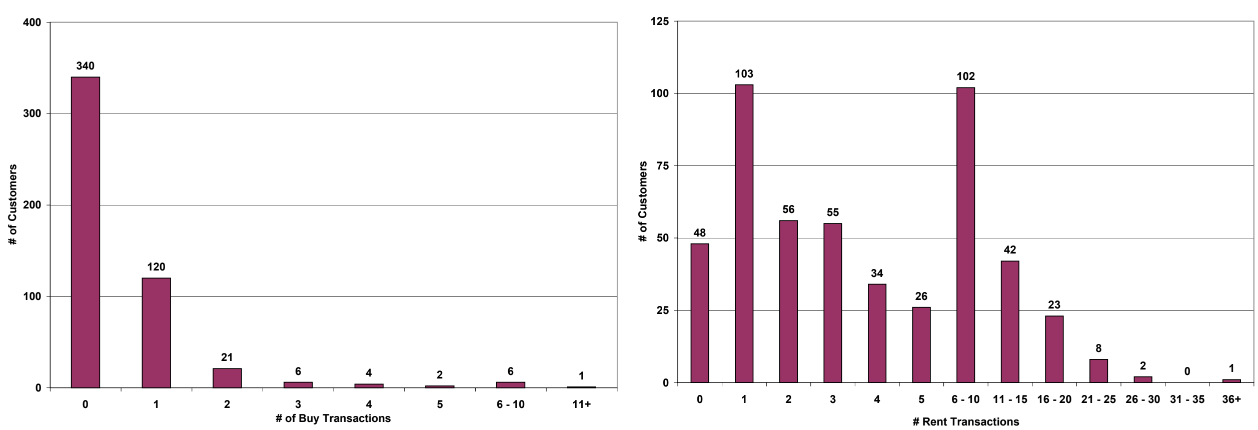
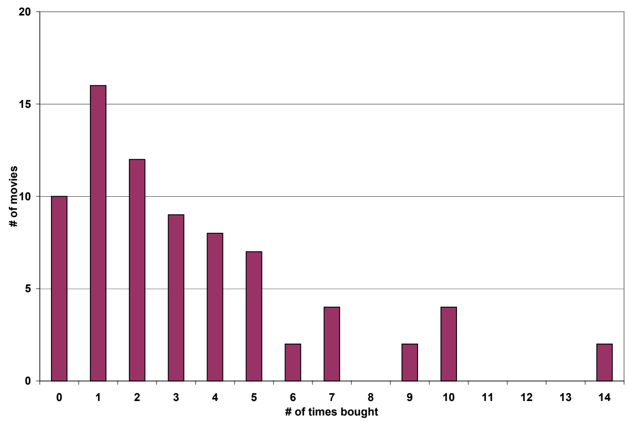


Figure 2: Histograms of Buying and Renting in the Population (Out of 500 Consumers)

across movies in Figure 3 ($n = 76$). We also present, illustratively, a snapshot of data for a particular customer (from the 112 above) who rents 10 and buys 1 video in Figure 4.

We supplement our data with a set of movie covariates collected from IMDB.com and

[10] $160 - 112 = 48$ only bought in the time horizon.



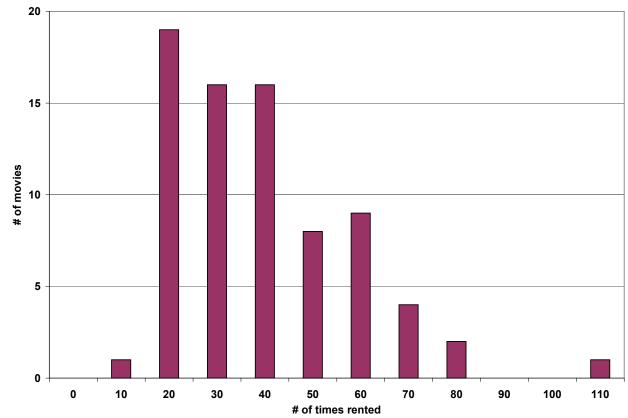


Figure 3: Histograms of Buying and Renting in the Movie Population (Out of 76 Movies)

CUST_ID	INV_DATE	TRANS_TYPE	DESC	Price
303	7/14/2003	R	PHONEBOOTH	3.5
303	7/18/2003	R	SHANGHAI KNIGHTS	3.5
303	7/24/2003	R	LIFE OF DAVID GALE	3.5
303	8/2/2003	R	GANGS OF NEW YORK	3.5
303	8/3/2003	R	DAREDEVIL	3.5
303	8/10/2003	R	OLD SCHOOL	3.5
303	8/13/2003	R	HEAD OF STATE	3.5
303	8/16/2003	R	BRINGING DOWN THE HOUSE	3.5
303	8/26/2003	B	LORD OF THE RINGS: TT	20.99
303	9/1/2003	R	HUNTED, THE (2003)	3.5
303	9/5/2003	R	CHICAGO	3.5

Figure 4: A Snapshot of Data for Customer 303

metacritic.com. IMDB.com provides raw data such as box office gross, MPAA rating, genre and distributor information. Metacritic.com aggregates reviews to come up with a 0-100 point scale, the Metascore, which is a weighted average of individual critic scores.

4.2 Model Estimation

We allow for heterogeneity across individuals and movies by adopting priors over the parameters for both models. Two features about our data set make it amenable to using a hierarchical Bayes estimation procedure. First, our data set is somewhat sparse. We observe many individuals who never buy and some movies that are never bought. Maximum likelihood estimates of X_{0i} and μ_j would be driven to the boundary of acceptable parameter values at 0 and infinity, respectively. Second, the number of observations we have over people and movies varies greatly. For individuals, the range is 1 - 39 observations, with an average of 6. For movies, the range is 8 - 112, with an average of 38. Sharing information across movies and people helps to alleviate this problem.

We estimate two versions of each model: one where we have a set of intercepts for each movie, and another where we describe differences in movies using M covariates Z_j. For model I, we have:

$$\begin{aligned}
\log \mu_j &= \alpha_j \\
\alpha_j &\sim \mathrm{N}(\mu_\alpha, \sigma^2_\alpha) \\
\log \mu_j &= Z_j \, \beta \\
\beta_k &\sim \mathrm{N}(\mu_\beta, \sigma^2_\beta) \\
\sigma^2 &\sim \mathrm{Unif}(0, 5000) \\
X_{0i} &\sim \mathrm{LogN}(\mu_X, \sigma^2_X)
\end{aligned} \tag{14}$$

For identification purposes we fix one of the parameters of the distribution across individuals, μ_X to 2.5. Note that this does not fix the mean of the log-normal, since σ^2_X also enters the expected value of a log-normally distributed random variable. We adopt a uniform prior over a finite range for σ^2 after experimenting with other distributions in our simulations and following some of the applied recommendations of Gelman (2004).

For model II, we have:

$$\begin{aligned}
\log \tfrac{p_{ij}}{1-p_{ij}} &= \gamma + v_i + \alpha_j \\
a_j &\sim \mathrm{N}(0, \sigma_a^2) \\
\log \tfrac{p_{ij}}{1-p_{ij}} &= \gamma + v_i + Z_j\, b \\
b_k &\sim \mathrm{N}(\mu_b, \sigma_b^2) \\
\gamma &\sim \mathrm{N}(\mu_\gamma, \sigma_\gamma^2) \\
v_i &\sim \mathrm{N}(0, \sigma_v^2)
\end{aligned} \tag{15}$$

For identification, we set γ as the grand mean and impose a zero mean prior on u_i and a_j.

We observe individuals $i = 1, \ldots, I$ who choose to either rent or buy $j = 1, \ldots, J_i$ movies. Define c_{ij} as the binary indicator that equals 1 if i buys j otherwise zero. Define the observation matrix $\mathbf{C} = (c_{ij})$, and the probability matrix (from either model I or II) $\mathbf{P} = P(\mathrm{Buy}_{ij})$. The likelihood for the model is a product of Bernoulli probabilities.

$$L(\mathbf{C} \mid \mathbf{P}) = \prod_i^I \prod_j^{J_i} P(\mathrm{Buy}_{ij})^{c_{ij}} (1 - P(\mathrm{Buy}_{ij}))^{1-c_{ij}} \tag{16}$$

Denote θ as the generic set of parameters with priors over them. For model I, $\theta = \{\alpha_j, \beta_k, X_{0i}, \sigma^2\}$ and for model II, $\theta = \{a_j, b_k, v_i, \gamma\}$. The marginal posterior for a particular parameter θ_q amounts to integrating the product of the likelihood and prior over all other parameters, θ_{-q}.

$$f(\theta_q \mid \mathbf{C}) \propto \prod_i^I \prod_j^{J_i} \int P(\mathrm{Buy}_{ij} \mid \theta)^{c_{ij}} \, (1 - P(\mathrm{Buy}_{ij} \mid \theta))^{1-c_{ij}} \, \pi(\theta) \, d\theta_{-q} \tag{17}$$

Because we cannot compute these integrals in closed-form, we use Monte Carlo sampling from a Markov Chain (Gelfand and Smith, 1990). For both models, the complete conditional densities are non-standard, and so Gibbs sampling can be ruled out. In practice, we allow the expert system of WinBUGS to choose the best sampling method. We fit both versions of model I outlined in equation 10 with priors from equation 14 as well as both versions of model II outlined in equations 13 and 15. We obtain the marginal posteriors of interest by using the freely available Bayesian inference software WinBUGS. [11] Inferences reported for all parameters are based on the combined draws of three independent chains run for 5000

[11] Available at http://www.mrc-bsu.cam.ac.uk/bugs.

iterations each, discarding the first 20,000 iterations as burn-in, from over-dispersed starting values. Convergence of the Markov chains (burn-in) was assessed by making use of the F-test statistic of Gelman and Rubin (1992a, 1992b). In practice, we saw rapid convergence within the first 1000 draws for both versions of both models. The code we use to estimate both models in any form is available from the authors upon request.

4.3 Simulating Model I's Estimation

Given the complexity of Model I we perform a simulation study to see how well the model does recovering parameter values. We report the procedure and results of the study in this section.

In keeping close to the size and range of observations in our data, we simulated 500 consumers making buy/rent decisions over 5 different movies. The details of the procedure are as follows:

1. We simulated 500 values of $X_i \sim \text{LogN}(2.5, .028)$
2. We set values of $\mu_j = \{4.5, 5.0, 5.5, 6.0, 6.5\}$
3. We set $\sigma^2 = 2.0$
4. Calculate $P(\text{Buy}_{ij})$ for each of the 2500 observations.
5. Draw a uniform random variable $Z \sim \text{Unif}(0, 1)$. If $Z \leq P(\text{Buy}_{ij})$, $c_{ij} = 1$, otherwise $c_{ij} = 0$.

The results of the simulation are shown in Table 1. All the true values shown below lie easily within 95 % posterior density interval of the estimated distributions. Of all 500 X_{0i} simulated values, 498 of those were within the 95 % posterior density interval of the estimates. We also looked at the population average $\bar{X_{0i}} = \frac{1}{500} \sum_i X_{0i}$.

Parameter	True Value	95% Interval	Post Mean	Post Std Dev	# Buys (out of 500)
μ_1	4.50	[4.17, 4.85]	4.59	0.17	163
μ_2	5.00	[4.88, 5.62]	5.30	0.19	64
μ_3	5.50	[5.33, 6.20]	5.79	0.22	61
μ_4	6.00	[5.71, 6.79]	6.25	0.27	7
μ_5	6.50	[6.27, 8.20]	7.08	0.50	3
σ^2	2.00	[0.16, 6.25]	2.10	1.64	–
$\bar{X_{0i}}$	12.21	[12.14, 12.29]	12.21	0.038	–

Table 1: Simulation results of Model I

5 Results

We now turn to discuss the empirical findings from the models. We first compare models in and out of sample, proceed to talk about explainable differences across movies using the covariates mentioned earlier, and end discussing Model I in greater detail.

5.1 Comparing Models

We distinguish among the models' ability to describe actual buying-renting behavior in the video store, using a sample of TLA's transaction files. To that end, we compare both models in and out of sample. Of the 2,913 transactions made available, we fit the models to a randomly chosen two-thirds, or 1,983 transactions. We leave the remaining 930 transactions for model validation. The percentage of purchases in the calibration sample is 10% and 7% in the validation.

We list a number of fit measures to consider in comparing the performance of the models. First, in order to compare the models' log marginal likelihoods, we use the harmonic log likelihood (Newton and Raftery, 1994) as an estimator. Second, we compare their mean absolute error (MAE), defined as the mean absolute difference between the probability and the actual choice. [12] Lastly, we compare hit rates, defined as the fraction of times when there was a purchase where the model predicted $P_{ij}(\text{Buy}) > 0.5$. We compare both versions

[12] $MAE = \frac{1}{N} \sum \mid c_{ij} - P_{ij}(\text{Buy}) \mid$

of both models on all of these measures. All of these are reported in Table 2. Models which perform better have greater log marginal likelihoods, smaller mean absolute errors, and greater hit rates.

		Within Sample			Holdout Sample	
Model	Version	Log-Marginal Likelihood	MAE	Hit Rate	MAE	Hit Rate
1	Intercepts	-190.99	0.035	0.811	0.056	0.687
2	Intercepts	-403.65	0.112	0.341	0.102	0.328
1	Covariates	-243.65	0.040	0.788	0.058	0.681
2	Covariates	-399.96	0.112	0.349	0.102	0.327

Table 2: Model Comparison

Model I significantly outperforms Model II both in-sample and out of sample. Since we are comparing within versions, across models the number of parameters is being held constant. The marginal log likelihood of Model I is markedly higher than Model II. The hit rate of Model I is more than double that of Model II and the mean absolute error of Model I is about half that of Model II. We conclude that Model I does a better job of describing our data than Model II.

At this point, it might be fruitful to recall the structural differences between the models. Model I assumes a stochastically diminishing valuation at each watching. Model II can be interpreted as a "value" model only if we assume that the consumer receives a constant value at each watching. In Model I, the threshold of when to stop watching is endogenously defined by an interaction of the starting point of the process (specific to individual i), and the coefficients of the process (specific to movie j). In Model II, this threshold is exogenous, and depends on the current renting and buying costs.

We now turn to comparing the two specifications of Model I. Specifically, we want to know whether we can reduce a set of 76 dummy variables, one for each movie, to an intercept and 8 explanatory movie covariates Zj, as described in section 4.1. We choose not to turn to the traditional penalized measures of fit such as the AIC or BIC, because the hierarchical structure makes counting the number of parameters in the model a difficult task. Hence,

Version	Raw parameters	p_e	DIC	BIC
Intercepts	577	142.1	447.7	774.1
Covariates	510	133.0	550.3	855.8

Table 3: Version Comparison for Model 1

we follow Spiegelhalter et al. (2002) and Congdon (2003) and estimate the effective number of parameters, p_e. The basic idea is to penalize parameters differentially, with respect to how much variance they contribute to the sample deviance mean. p_e is estimated as the difference between the sample average deviance and the deviance measured at the posterior mean of the parameters.[13] We then input p_e into the standard AIC and BIC calculations. The modified AIC is called the Deviance Information Criterion (DIC).

$$\begin{aligned} DIC &= -2\,\mathrm{LogL} + 2p_e \\ BIC &= -2\,\mathrm{LogL} + p_e \ln(n) \end{aligned} \tag{18}$$

In Table 3, we list both measures, raw number of parameters, and p_e. Although the covariate version has 67 fewer raw parameters, we see that the effective number parameters after controlling for parameter deviance is much closer. Both DIC and modified BIC calculations suggest that we are better off with intercepts version of Model I. The finding that observable movie characteristics do little to explain movie effects has been also shown by Einav (2004). We next illustrate the usefulness of this particular model by demonstrating how a retailer could customize prices to consumers in section 6.

5.2 Diagnostic Results

Before we discuss the coefficient-based diagnostic results, we briefly review the effect of the signs on both models. Note that a negative coefficient in either model increases the probability of the movie being bought. In Model I, a negative coefficient implies a lower decay rate over viewing; hence, the value gained from buying and viewing more often is

[13]Deviance is defined as as $-2\,\mathrm{LogL}$. For more details on estimating the effective number of parameters, see cites above.

greater. In Model II, a negative coefficient implies that the probability one stops watching at each trial is lower; again, this increases the probability the movie is bought.

In addition, we caution that these results represent the clientele of TLA and may not be representative of the home video market as a whole. Our objective, however, is not to determine which, if any, variables can strongly predict rental and purchase sales across the home video market at large (for an example of this, see Prosser 2002). We are interested in a particular retailer's (i.e., TLA's) prediction problem, given its own customer base, in order to customize its prices or manage its inventory.

In Table 4, we show the coefficient results for Models 1 and 2. Both models reveal consistently that box office gross (in $ millions) increases the likelihood that a particular movie is bought. In our sample of 76 movies, we observed some high grossing movies such as Lord of the Rings: Two Towers ($ 341.7 M), Catch Me If You Can ($ 164.6 M), and Die Another Day ($ 160.9 M). In the data, we observed 14 out of 93, 9 out of 45, and 5 out of 37 recorded buys for these three movies. On the other hand, box office gross could not explain why a movie like How To Lose A Guy in 10 Days ($ 106.1 M) garnered one buy out of 41 transactions, whereas a cult favorite, like About Schmidt ($ 65.0 M) saw 7 buys out of 54 transactions. Conventional wisdom holds that a movie can correct for a bad showing at the box office by subsequent DVD sales, because DVD releases provide more of a "level" playing field to all movies being released than at the theater, where only a handful of movies at any given time are seen as being successful (Fabrikant, 2001).

The potentially intriguing set of results are that Rated R and action movies, long considered to be some of the more unprofitable movies in the business, appear to significantly increase a movie's chances of being bought, at least according to Model I. For example, Medved (1992) shows that R-rated films were less than half as likely as PG-rated releases to reach $25 million in domestic box office revenues. De Vany and Walls (2002) examine more formally whether Hollywood produces too many R-rated movies by modeling the distribution of profits and accounting for risk across the ratings categories. They show that R-rated movies are stochastically dominated by G, PG, and PG-13 rated movies according to profits, revenues, costs, and returns on production costs. Notably, these studies have ignored home

	Model 1		Model 2	
Covariate	$\mathsf{E}[\beta]$	Std Error	$\mathsf{E}[b]$	Std Error
Box Office Gross ($ mil)	-0.0015**	0.000512	-0.0025**	0.000821
MetaCritic Rating (0-100)	-0.0012	0.00156	-0.0058*	0.00345
Rated R	-0.1636**	0.0657	-0.0237	0.103
Drama	-0.0103	0.0591	-0.0210	0.116
Action	-0.1774**	0.0743	-0.0523	0.118
Thriller	0.0819	0.0649	0.0464	0.108
Romance	-0.0166	0.0533	-0.0168	0.117
Major Distributor	-0.0309	0.04937	-0.0370	0.0885

Table 4: Coefficient Results for Models 1 and 2

** $p < 0.01$, * $p < 0.10$.

video profits, which have since grown to account for more than half of a movie's profit than the studio. Despite this, a few articles in the popular press have recognized the growing trend, that action movies (in particular R-rated action movies) are more "durable" in the sense that they can be enjoyed multiple times. For example, The New York Times recently reported that "young men spend the most on DVDs, so male-driven films are what home-theater bosses like best, as well as family themes - not adult drama (Kirkpatrick, 2003)."

5.3 Additional Model I Results

We now turn to examine additional results of Model I, given that appears to be the best fitting model. For the retailer, we are interested in ranking movies according to their "purchasability." It is important to note that our model is conditional on movie choice. In order to make stocking decisions, a retailer would want, in addition to the rent/buy analysis we provide, some estimates of how often the movies are transacted in the larger population.

At each draw from the posterior of a movie's decay parameter μ_j, we can calculate its rank relative to the 75 other movies (recall that a lower value of μ_j means that a movie is more likely to be bought). We can repeat this procedure for each draw, and collect the ranks into a distribution. This procedure allows movie and customer uncertainty to be reflected in the rankings. The distribution of ranks for the movie Chicago is shown below in Figure 5. The median of this distribution is 13. Since we compare movie parameters based on the median of these distributions, there may be ties for each rank.

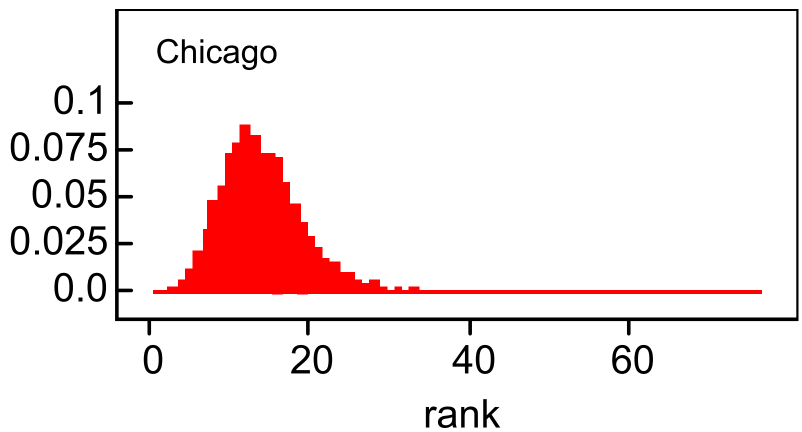


Figure 5: Histogram of ranks for the parameter μ_j for the movie Chicago

We first present the top 10 movies based on median rank of the posterior distribution of μ_j, posterior means and standard errors of μ_j, average purchase price paid by those who purchase the movie, and the number of times movie j is bought and rented in Table 5. Surprisingly, not all the movies in this list are sure bets. For example, Spirited Away, is transacted 23 times in our data set. It is bought 6 times, twice at full price ($26.99), twice at a mild discount ($17.99), and twice at a major discount ($12.99). Clearly, price can only explain some of this movie's top ranking, since there are four other movies on this list with higher average purchase prices. If we look at the 6 customers who bought the movie, 5 of them only bought one movie: Spirited Away. Lord of the Rings, Twin Towers, on the other hand, is bought more 8 times and at a higher average price, but by buyers who bought other movies. The niche appeal of Spirited Away is such that, conditional on a customer choosing that movie to watch, the likelihood that the customer is going to buy rather than rent is large relative to the other movies.

Movie	Median Rank	Posterior $\mathsf{E}[\mu_j]$	Posterior $SE[\mu_j]$	Average Purchase Price Paid ($)	Bought / Rented
SPIRITED AWAY	1	3.535	0.3732	18.16	6/17
JERRY SEINFELD COMEDIAN	2	4.106	0.6025	26.99	1/11
DIE ANOTHER DAY	4	4.418	0.5899	13.39	5/32
SHANGHAI KNIGHTS	4	4.426	0.5968	14.83	3/34
LOTR: TWO TOWERS	7	4.615	0.3158	21.11	14/79
OLD SCHOOL	7	4.641	0.4186	16.14	10/34
FRIDA	11	4.918	0.633	13.42	7/65
CATCH ME IF YOU CAN	12	4.936	0.4363	18.07	9/36
KID STAYS IN THE PICTURE	12	4.993	0.6176	21.88	2/20
VIEW FROM THE TOP	12	5.051	0.7983	29.99	1/31

Table 5: Top 10 Most Purchasable Movies, given transaction, based on μ

A similar story goes for Jerry Seinfeld Comedian, which is transacted only 12 times, but bought once for full price. More obvious favorites are Die Another Day and Shanghai Nights: these were transacted over 35 times each, and bought 3 or more times. Not all heavily bought movies have low values of μ_j: some will be bought opportunistically, only when the price is low, or when the consumer buying it is a collector who always buys movies. The model weighs these factors - from the individual consumer, the price, and the movie - in order to estimate μ_j.

We can compare with Table 6, which lists the bottom-ranked μ_j's, or the most rentable movies given selection. Even though some of the movies listed here are purchased more than 5 times, they are purchased on average at a deep discount. In addition, these movies are bought by consumers who are mostly collectors: their willingness to buy is due more to consumer rather than movie level factors.

Further, we can classify movies into 3 categories: “buyable”, “average”, and “rentable”. Recall lower μ_j increase a movie's buyability. We can show the classifications for each movie relative to the mean buying level. We estimated the probability that μ_j was in the upper or lower quartile. The results are presented in Figure 6.

Movie	Median Rank	Posterior $\mathsf{E}[\mu_j]$	Posterior $SE[\mu_j]$	Average Purchase Price Paid ($)	Bought / Rented
TWO WEEKS NOTICE	60	7.242	0.9269	7.99	1/13
CONFESSIONS	62	7.34	0.6458	11.39	7/46
RUSSIAN ARK	62	7.397	0.9718	–	0/24
RABBIT-PROOF FENCE	63	7.445	0.9179	–	0/22
TALK TO HER	64	7.453	0.7027	15.69	3/25
ANGER MANAGEMENT	65	7.534	0.893	7.50	2/42
PIANIST	67	7.65	0.8543	9.68	7/46
IDENTITY	67	7.679	0.698	11.62	4/60
SOLARIS	68	7.724	0.5979	15.09	5/52
HOW TO LOSE A GUY	70	7.99	0.9023	29.99	1/40

Table 6: Top 10 Most Rentable Movies, given transaction, based on μ

Buyable	*Average*	*Rentable*
ADAPTATION	ABOUT SCHMIDT	2 FAST 2 FURIOUS
ANALYZE THAT	ANTWONE FISHER	25TH HOUR
ANIMATRIX	BASIC	ANGER MANAGEMENT
BEND IT LIKE BECKHAM	BOAT TRIP	CONFESSIONS OF A DANGEROUS MIND
BULLETPROOF MONK	BOWLING FOR COLUMBINE	CONFIDENCE
CATCH ME IF YOU CAN	BRINGING DOWN THE HOUSE	HOLES
CHICAGO	CORE	HOURS
DAREDEVIL	DADDY DAY CARE	HOUSE OF 1,000 CORPSES
DELIVER US FROM EVA	DANCER UPSTAIRS	HOW TO LOSE A GUY IN 10 DAYS
DIE ANOTHER DAY	DARK BLUE	IDENTITY
FRIDA	DREAMCATCHER	JUST MARRIED
GANGS OF NEW YORK	DYSFUNKTIONAL FAMILY	LOVE LIZA
HEAVEN	FINAL DESTINATION 2	PIANIST
IRREVERSIBLE	GODS AND GENERALS	QUIET AMERICAN
JERRY SEINFELD COMEDIAN	GOOD THIEF	RABBIT-PROOF FENCE
KID STAYS IN THE PICTURE	GURU	RUSSIAN ARK
LORD OF THE RINGS: TWO TOWERS	HE LOVES ME/HE LOVES ME NOT	SOLARIS
OLD SCHOOL	HEAD OF STATE	TALK TO HER
PHONEBOOTH	HUNTED	TWO WEEKS NOTICE
PUNCH-DRUNK LOVE	LAUREL CANYON	
SHANGHAI KNIGHTS	LIFE OF DAVID GALE	
SPIRITED AWAY	LOST IN LA MANCHA	
SPUN	MALIBU'S MOST WANTED	
VIEW FROM THE TOP	MAN APART, A	
	MIGHTY WIND	
	NARC	
	RAISING VICTOR VARGAS	
	REAL WOMEN HAVE CURVES	
	RECRUIT	
	SECRETARY	
	SHAPE OF THINGS	
	SPIDER	
	TEARS OF THE SUN	

Figure 6: Movie classification based on quartile partitioning of posterior $\mathsf{E}[\mu_j]$

Instead of ranking customers, as we do with movies, we highlight the model's ability to discriminate consumers who bought exactly twice in the 6 month time period in Table 7.

Customer	Posterior $\mathsf{E}[X_{0i}]$	Posterior $SE[X_{0i}]$	Bought / Rented	Average Price Paid ($)	Average Posterior $\mathsf{E}[\mu_j]$ bought
36	17.47	1.046	2/1	28.99	5.92
26	14.3	0.563	2/6	19.99	5.64
29	14.15	0.742	2/4	17.25	4.27
21	13.18	0.664	2/10	13.99	6.49
20	12.47	1.174	2/6	11.49	5.49
34	12.15	1.141	2/5	9.99	6.36
23	12.1	1.016	2/8	11.49	5.74

Table 7: Comparing customers' by X_{0i} with 2 purchases

We present posterior mean and standard error of X_{0i} along with the number of rent and buy transactions, average purchase price paid, and average posterior mean μ_j of titles purchased. Customer 36 has the highest posterior mean X_{0i} since he or she purchases a movie at a relatively high price, rents only once and buys movies which are not across the board hits (average posterior mean μ_j is relatively high). Customer 29, though he or she rents less than customer 26, pays on average less money to purchase movies, and so is less of a valuable buyer to the store, assuming costs are equal, than customer 26.

6 Customized Purchase Prices

The results discussed earlier (e.g., purchasable vs. rentable movies) have implications concerting the type of products to be offered by the store to the customers. Additionally, we demonstrate the usefulness of the model by showing how the retailer can customize retail prices to a sample of customers with different histories of renting and buying titles. Such customized pricing applications are becoming available due to cheaper and more comprehensive technology. For example, IBM's Shopping Buddy can store grocery lists, loyalty program benefits, in addition to giving personalized discounts on preferred brands (Shermach, 2004). In the home video market, Blockbuster has also begun to use customized e-coupons to induce customers that rent online to purchase previously viewed movies from local offline stores. In the general context of customized pricing, Rossi, McCulloch and Allenby (1996) show how a retailer can use posterior distributions of a choice model to customize coupon values based on consumer transaction history. We follow such a strategy in customizing purchase prices for a sample of transactions from our data.

In customizing prices, we choose to focus on the profitability of customizing prices (without regard to menu costs, for example). We also ignore the price effects of switching movies and stores. The first effect would require a more extensive first stage choice model over 76 movies and 500 individuals, which would be extremely difficult to estimate. The second would require an augmented data set with store choice. We can partially address the absence of a movie choice stage by focusing on consumer transactions where a movie was already chosen to be rented. That is, we focus on converting rent transactions from customers who have bought in the past into purchases. In our example, all customized prices are for movies that were rented by the consumer in the sample.

The expected total profit to the video retailer from consumer i and movie j is:

$$\mathsf{E}[\tilde{\pi}] = P_{ij}(\text{Buy} \mid \tilde{\theta}, K, k)\ (K - c_B) + (1 - P_{ij}(\text{Buy} \mid \tilde{\theta}, K, k))\ (k - c_R) \tag{19}$$

where $\tilde{\theta}$ is a vector that denotes the model parameters $(X_{0i}, \mu_j, \sigma, r)$, c_B is the wholesale cost of a tape to the retailer and c_R is the marginal cost of renting a copy. At each value of K, we simulate the posterior distribution of expected total profit by drawing from the posterior distribution of $\tilde{\theta}$. The posterior distribution of $\tilde{\theta}$ allows us to incorporate uncertainty about the individual's willingness to purchase and the movie's purchasability and measure their effects on the posterior distribution of expected transaction profits to the retailer. For example, an individual who is not observed often in the data may have a wide posterior X_{0i} density, which in turn would generate a wider range of posterior expected profits to the retailer. We vary K in increments of \$0.25 from \$10.00 to \$25.00, and calculate averaged posterior expected profits from a sample of 2000 draws in order to find the profit-maximizing K for equation 19. We set the wholesale cost $c_B =$ \$14.00, in line with most sell-through title pricing (Gertner, 2004), and set the marginal cost of renting to zero. All the results were obtained using WinBugs.

We illustrate our methodology for finding optimal K by plotting the average and 95% confidence intervals of the posterior distribution of expected profit for a range of purchase values for a specific transaction in figure 7. We choose for this particular example customer 4, who buys 7 and rents 21 titles in our sample over the 6 month period. The movie in question is Final Destination 2, which customer 4 initially chose to rent. TLA's price for this movie was \$25.19. The optimal purchasing price according to our model for this customer is $K^* = \$19.50$, though in this region the expected transaction profit appears to be relatively flat in the \$18.50 - \$20.50 region.

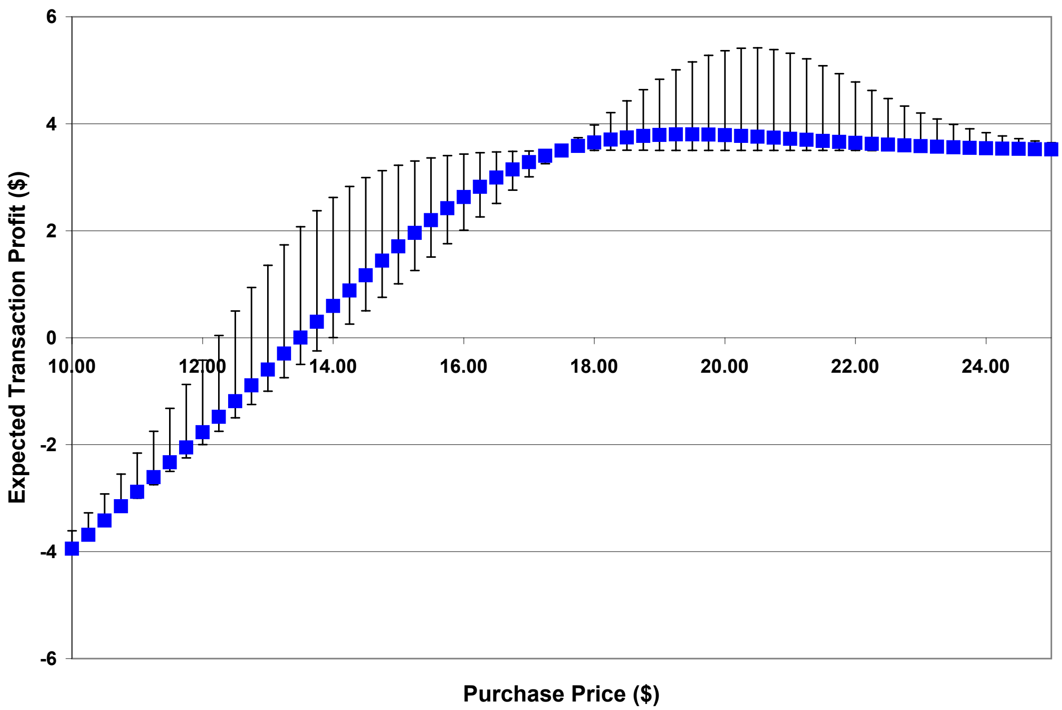


Figure 7: Posterior distribution of expected transaction profit ($\mathsf{E}[\pi]$) for various purchase prices (K)

It is worth pointing out from figure 7 that the response of expected profitability to the purchase price K is asymmetric. The lower 95% bound of the expected profit posterior distribution asymptotes at k_i (here \$3.50) as K gets large. This is due to the fact that the retailer is "guaranteed" renting the movie for k_i, because if the buy price K is extremely high the consumer will always rent with probability 1. The retailer can price a movie to

Cust	Movie	Rent Price (k)	Buy Price (K)	Customized $P_{ij}(Buy\|K^*, \theta)$	Customized Buy Price (K^*)	$\mathsf{E}[\pi]$
4	GOOD THIEF	3.50	25.19	0.175	19.50	3.85
4	VIEW FROM THE TOP	3.50	26.99	0.402	20.50	4.61
4	2 FAST 2 FURIOUS	3.50	19.99	0.094	19.00	3.64
4	FINAL DESTINATION 2	3.50	25.19	0.151	19.50	3.80
7	FINAL DESTINATION 2	3.50	25.19	0.270	20.50	4.31

Table 8: Customized Prices for Customers 4 & 7

rent only by setting K arbitrarily large. As K gets small, however, the expected profit can be less than the margin the retailer makes on renting the movie out, or negative. To see whether it is optimal for the retailer to price the movie for the consumer to rent or buy, we can look at the $P_{ij}(Buy)$ at the optimal K and compare across movies.

We turn to some other transactions of customer 4 in table 8. Customer 4 chose to rent The Good Thief at \$ 3.50 instead of buying it for \$ 25.19. According to the model, the optimal K^* TLA should charge is \$ 19.50. Notice that for another movie customer 4 decided to rent, View From the Top, the optimal price is only \$ 1 higher. That is because this movie is considered more "buy-able" in general, and so the retailer does not have to knock the price down to induce the consumer to buy. As another illustration we compare targeted prices of the same movie to different consumers. Consider also the history of customer 7, at the bottom of table 8, comparing the custom prices of Final Destination 2.

Customer 4 gets commands a lower price than customer 7, because customer 4 appears to be more willing to rent. The store would have to lower the purchase price for him more, then, in order to induce him to buy. At the same time, however, customer 7 has bought his movies at relative discount, whereas customer 4 has bought one of her movies at full price. This effect would make customer 4 appear more of a buyer than a renter, though the size of the effect may not be as large as the one just mentioned. The model weighs these two relative effects in deriving a custom price for each customer based on transaction history.

A simple heuristic that the store manager may employ is setting the price K such that the consumer is indifferent between renting and buying the movie, or $P_{ij}(\text{Buy}) = 0.5$. In practice, we use a sample of 2000 draws to find the highest price $\bar{K}$ such that $P_{ij}(\text{Buy}) \geq 0.5$. The heuristic price $\bar{K}$ for customer 4, was \$16.50 for Good Thief, \$19.50 for View from the Top, \$15.75 for 2 Fast 2 Furious, and \$ 16.50 for Final Destination 2. For customer 7, the price was \$ 18.25 for Final Destination 2. Though the heuristic prices are generally within \$3 of the profit maximizing prices in table 8, this heuristic generally under-prices the movies, because it misses the asymmetric response of the purchase price on profits as price increases.

7 Discussion

In this paper, we attempt to explain and predict the consumer's decision to rent or buy a particular movie at a video store. To our knowledge, this is the first academic paper that has attempted to model this particular decision. One approach we take from the real options literature in economics, the other from cognitive psychology. We find that that the real options model which assumes diminishing value over watching and forward looking consumers that know when to stop watching, fits our data substantially better. We also find that box office gross, an R rating, and action genre significantly increase the chance that a movie is bought. These findings confirm some of the industry intuition we see in the popular press.

There are a number of limitations to our analysis that represent future research directions. Firstly, we choose to model only the rent/buy decision conditional on movie choice. Future work could integrate such a model with a first stage movie choice model. Secondly, our data do not allow us to observe day-to-day availability of a title. To the extent that consumers substitute across titles given unavailability, we would not expect this to bias our results. However, if consumers substitute renting for buying the same movie or vice versa because the movie is unavailable in one format, we would expect our results to be biased. TLA provides more copies (both for buying and renting) of popular titles, however, because these movies are popular, they are more likely to be out of stock. Some retailers such as Blockbuster guarantee rental availability for popular titles. We can only partially control for this by choosing movies that TLA stocks heavily. Lastly, we do not attempt a global search of plausible models. Our two models are meant to capture extreme points on the continuum of theories of consumer choice. Future work could further develop and test more models.

References

Ainlie, Andrew, Xavier Drèze, and Fred Zufryden. Modeling Movie Lifecycles and Market Share. Working Paper, Department of Marketing, UCLA, 2004.

Bertsekas, Dimitri P. Dynamic Programming and Optimal Control. Athena Scientific, Belmont, MA: 2000.

Brennan, M. and E. Schwartz. Evaluating Natural Resource Investments. Journal of Business 58:2, 135-57, 1985.

Congdon, Peter. Applied Bayesian Analysis. Wiley Series in Probability and Statistics, Sussex: 2003.

Della Vigna, Stefano and Ulrike Malmendier. Overestimating Self-Control: Evidence From the Health Club Industry. Working Paper, University of California, Berkeley, 2003.

De Vany, Arthur and W. David Walls Does Hollywood Make Too Many R-Rated Movies? Risk, Stochastic Dominance, and the Illusion of Expectation. *Journal of Business*, 75:3, 425-51, 2002.

Dixit, Avinash K. and Robert S. Pindyck. Investment Under Uncertainty. Princeton University Press, Princeton, NJ: 1994.

Einav, Liran. Gross Seasonality and Underlying Seasonality: Evidence from the U.S. Motion Picture Industry. Working Paper, Stanford Economics Department, 2001.

Elberse, Anita and Jehoshua Eliashberg. Demand and Supply Dynamics for Sequentially Released Products in International Markets: The Case of Motion Pictures. *Marketing Science*, 22:3, 329-354, 2003.

Eliashberg, Jehoshua and Mohanbir Sahwney Modeling Goes to Hollywood: Predicting Individual Differences in Movie Enjoyment. *Management Science*, 40:9, 1151-1173, 1994.

Eliashberg, Jehoshua, Jedid-Jah Jonker, Mohanbir S. Sawhney, and Berend Wierenga MOVIEMOD: An Implementable Decision Support System for Pre-Release Market Evaluation of Motion Pictures. *Marketing Science*, 19:3, 226-243, 2000.

Eliashberg, Jehoshua, Sanjeev Swami, Charles Weinberg, and Berend Wierenga Implementing and Evaluating SILVERSCREENER: A Marketing Management Support System for Movie Exhibitors. *Interfaces*, 31:3, S108-S127, 2001.

Erdem, Tülin, and Michael P. Keane Decision-making Under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets. *Marketing Science*, 15:1, 1-20, 1996.

Fabrikant, Geraldine. "Sale of DVD's is Challenging Movie Rental Business." *The New York Times*, April 16, 2001.

Gelman, Andrew. Prior Distributions for Variance Parameters in Hierarchical Models. *Working Paper*, Columbia University, 2004.

Gelman, Andrew and Donald B. Rubin. A single sequence from the Gibbs sampler gives a false sense of security. Bayesian Statistics 4, J.M. Bernardo et. al. eds. New York: Oxford University Press, 1992a.

Gelman, Andrew and Donald B. Rubin. Inference from iterative simulation using multiple sequences (with discussion). *Statistical Science*, 7, 457-511, 1992b.

Gelfand, A. E. and A. F. M. Smith. Sampling Based Approaches to Calculating Marginal Densities. *Journal of the American Statistical Association*, 85, 398-409, 1990.

Gertner, Jon. "Box Office in a Box." *The New York Times*, November 14, 2004.

Gradshteyn, I. S. and I. M. Ryzhik. Table of Integrals, Series, and Products, 5^{th} *ed.* Academic Press: New York, 1994.

Gonul, Fusun, and Kanaan Srinivasan. Estimating the Impact of Coupon Expectations on Purchase Behavior: A Dynamic Structural Model. *Marketing Science*, 15:3, 262-79, 1996.

Hu, Ye, Jehoshua Eliashberg, and Jagmohan S. Raju. Strategic Release of Home Videos. Working Paper, Department of Marketing, the Wharton School, 2004.

Huberman, Bernardo, Peter Pirolli, James Pitkow, and Rajan Lukose. Strong Regularities in World Wide Web Surfing. *Science*, 280: 95-97, 3 April 1998.

Jones, J. Morgan and Christopher J. Ritz. Incorporating Distribution into New Product Diffusion Models. *International Journal of Research in Marketing*, 8, 91-112, 1991.

Karlin, Samuel and Howard Taylor. A First Course In Stochastic Processes, 2nd ed. Academic Press, New York: 1975.

Kirkpatrick, David. "Action-Hungry DVD Fans Sway Hollywood." *The New York Times*, August 17, 2003.

Kridel, Donald J., Dale E. Lehman, and Dennis L. Weisman. Option Value, Telecommunication Demand, and Policy. *Information Economics and Policy*, 5, 125-44, 1993.

Krider, Robert E. and Charles B. Weinberg Competitve Dynamics and the Introduction of New Products. *Journal of Marketing Research*, 35, 1-15, 1998.

Lambrecht, Anja and Bernd Skiera. Paying Too Much and Being Happy About it: Causes and Consequences of Tariff Choice-Biases Working Paper, Johann Wolfgang Goethe-University Frankfurt am Main, 2004.

Lehmann, Donald R., and Charles B. Weinberg Sales Through Sequential Distribution Channels: An Application to Movies and Videos. *Journal of Marketing*, 64, 18-33, 2000.

Litman, Barry R. and Hoekyun Ahn. Predicting Financial Success of Motion Pictures. in B. R. Litman, ed., The Motion Picture Mega-Industry, Allyn & Bacon, Needham Heights, MA: 1998.

Lukose, Rajan and Bernardo Huberman. Surfing as a Real Option. Working Paper, 1998.

Medved, Michael. Hollywood vs. America: Popular Culture and the War on Traditional Values. HarperCollins, New York: 1992.

Miravete, Eugenio. Choosing the Wrong Calling Plan? Ignorance and Learning. *American Economic Review*, 93, 297-310, 2002.

Mortimer, Julie. Vertical Contracts in the Video Retail Industry. Working Paper, Department of Economics, Harvard University, 2004a.

Mortimer, Julie. Price Discrimination and Copyright Law: Evidence from the Introduction of DVDs. Working Paper, Department of Economics, Harvard University, 2004b.

Moul, Charles C. Measuring Market Power in the Theatrical Distribution of Movies Working Paper, Department of Economics, Washington University, 2004.

Nunes, Joseph C. A Cognitive Model of People's Usage Estimations. *Journal of Marketing Research*, 37: 397-409, 2000.

Neelamegham, Ramya and Pradeep Chintagunta. A Bayesian Model to Forecast New Product Performance in Domestic and International Markets. *Marketing Science*, 18:2 115-36, 1999.

Neelamegham, Ramya and Dipak Jain. Consumer Choice Process for Experience Goods: An Econometric Model and Analysis. *Journal of Marketing Research*, 36: 373-86, 1999.

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

Orbach, Barak Y. and Liran Einav Uniform Prices for Differentiated Goods: The Case of the Movie-Theater Industry Harvard Olin Discussion Paper No. 337, October, 2001.

Prosser, Elise K. How Early Can Video Revenue Be Accurately Predicted? *Journal of Advertising Research*, March-April: 47-55, 2002.

Radas, Sonja and Steven M. Shugan Seasonal Marketing and Timing New Product Introductions. *Journal of Marketing Research*, 35: 296-315, 1998.

Ravid, S. Abraham. Information, Blockbusters, and Stars: A Study of the Film Industry. *Journal of Business*, 72: 483-92, 1999.

Reinan, John. "Movie Rentals Rewind." *Star Tribune*, February 28, 2004.

Rossi, Peter E., Robert E. McCulloch, and Greg M. Allenby. The Value of Purchase Histroy Data in Target Marketing *Marketing Science*, 15, 321-340, 1996.

Shaked, Avner and John Sutton Relaxing Price Competition Through Product Differentiation. *Review of Economic Studies*, 49, 3-13, 1982.

Spiegelhalter, D., N. Best, B. Carlin, A. van der Linde. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society, Series B*, 64, 1-34, 2002.

Sawhney, Mohanbir S. and Jehoshua Eliashberg A Parsimonious Model for Forecasting Gross Box-Office Revenues of Motion Pictures. *Marketing Science*, 15, 113-131, 1996.

Varian, Hal. Buying, Sharing and Renting Information Goods. *Journal of Industrial Economics*, 48, 4:473-88, 2000.

Appendices

A Equivalence of Equations 1 and 2

The discrete time process of equation 1 leads to normally distributed differences over time of the following form:

$$X_{\tau+s} - X_s = \sum_{i=1}^{\tau} \epsilon_i \sim \mathrm{N}(-\mu\,\tau, \tau\sigma^2)$$

For every pair of disjoint time intervals, the $[\tau_1,\, \tau_2]$ and $[\tau_3,\, \tau_4]$ the differences $X_{\tau_4} - X_{\tau_3}$ and $X_{\tau_2} - X_{\tau_1}$ are independent normally distributed random variables.

The Brownian motion of equation 2 properties come from Karlin and Taylor (1975):

Definition 1 *(Karlin and Taylor, 1975) Brownian motion with drift is a stochastic process $\{X(s); t \geq 0\}$ with the following properties:*

- *Every increment $X(\tau+s) - X(s)$ is normally distributed with mean $-\mu\,\tau$ and variance $\tau\,\sigma^2$.*
- *For every pair of disjoint intervals $[t_1, t_2]$ and $[t_3, t_4]$ the increments $X(\tau_4) - X(\tau_3)$ and $X(\tau_2) - X(\tau_1)$ are independent random variables with distributions given above, and similarly for n disjoint time intervals, where n is an arbitrary positive integer.*
- *$X(0) = 0$ and $X(t)$ is continuous at $t = 0$*

The first point mirrors the property of our discrete time continuous state stochastic process. The second point is true as well for both processes. The third point is the only difference. Here we can add an arbitrary number to the process, so that the process starts not at 0 but at $X(0) = X_0$. The continuity is different than the discrete time process, but that property is precisely what makes the expression easier to manipulate.

B Derivation of $\bar{X}$: Equation 4

In our model, the state variable is X, the control is $u \epsilon \{0, 1\}$, and the terminal payoff for stopping is 0. Let $F(X)$ denote the current value of the video plus the value of the option to keep watching it in the future. Then the Bellman equation is:

$$F(X) = \max_u \left[0,\ X + \frac{1}{1 + rdt}\ \mathsf{E}[F(X') \mid X] \right] \tag{20}$$

where X' is the state variable at some point in the future. In the continuation region, the second term on the right hand side is the larger of the two:

$$F(X) = X + \frac{1}{1 + rdt}\ \mathsf{E}[F(X') \mid X] \tag{21}$$

Using standard analysis techniques for solving continuous time stopping problems (see Dixit and Pindyck 1994, pg 110), we proceed by transforming the value function into an ordinary differential equation. Firstly, note the simple identity for a small change in X over the value function: namely that $F(X') = F(X + dX) = F(X) + dF$. Then taking expectations, we have $\mathsf{E}(F(X + dX) \mid X) = \mathsf{E}(dF) + F(X)$. We can expand the expected continuation formula using Ito's Lemma:

$$\mathsf{E}(F(X + dX)) = F(X)\ - [\mu\, F'(X) - \frac{1}{2}\,\sigma^2\, F''(X)]\, dt + o(dt) \tag{22}$$

where $o(dt)$ represent terms that go to zero faster than dt and $F'(X)$ represents the derivative with respect to X. Dividing both sides by dt, one can show that the value function $F(X)$ is equal to the second term which contains the first and second derivatives of the value function. Then we have the following ordinary differential equation (ODE) to solve:

$$r\, F(X) = X - \mu\, F'(X) + \frac{1}{2}\,\sigma^2\, F''(X) \tag{23}$$

This second order ODE has one particular and two general solutions, with one positive and one negative root:

$$F(X) = C_1\, e^{b_1\, X} + C_2\, e^{b_2\, X} + \frac{rX - \mu}{r^2} \tag{24}$$

where

$$b_1 = \frac{\mu + \sqrt{\mu^2 + 2r\sigma^2}}{\sigma^2} > 0 \tag{25}$$

and

$$b_2 = \frac{\mu - \sqrt{\mu^2 + 2r\sigma^2}}{\sigma^2} < 0 \tag{26}$$

Define $\bar{X}$ as the optimal stopping point. We have the following three boundary conditions for the two constants and $\bar{X}$. This gives us three equations for three unknowns.

$$\begin{aligned} F(\bar{X}) &= 0 \\ F'(\bar{X}) &= 0 \\ \lim_{X\to\infty} F'(X) &< \infty \end{aligned} \tag{27}$$

The first condition simply states that the value function is zero at the optimal threshold $\bar{X}$. Utilities above this threshold are by definition in the continuation region. The second condition is known in the economics literature as the "smooth pasting" condition, which states that if $\bar{X}$ is optimal, then the derivative of the value function at that point should be zero (the value function "smoothly" descends to zero). The last condition ensures that the derivative of the value function is bounded.

The last condition implies that $C_1 = 0$. To see this, note that the derivative of the value function is:

$$F'(X) = b_1 C_1\, e^{b_1 X} + b_2 C_2\, e^{b_2 X} + \frac{1}{r} \tag{28}$$

since $b_1 > 0$ and $b_2 < 0$ the first term approaches infinity and the second term approaches zero as $U \to \infty$. The first two conditions imply that:

$$\bar{X} = \frac{b_2 \mu + r}{r b_2} \tag{29}$$

We can substitute terms for b_2 and this yields the same term that appears in equation 4.

Stopping Threshold $\bar{X}$ is Non-positive

Proposition 1 *For $\sigma^2 > 0$, $r > 0$, and $0 < \mu < \infty$, the optimal threshold $\bar{X} < 0$. When μ approaches infinity, the threshold goes to zero:* $\lim_{\mu\to\infty} \bar{X} = 0$.

Proof We need to show that:

$$\frac{b_2 \mu + r}{r b_2} < 0$$

Since $b_2 < 0$, we need to show that:

$$b_2 \mu > -r$$

Re-writing b_2 as in equation 26 and some simple algebra shows that is true only if:

$$\mu^2 + r\sigma^2 > \mu\sqrt{\mu^2 + 2r\sigma^2}$$

Squaring both sides, we see that $\bar{X}$ is negative only if:

$$r^2\sigma^4 > 0$$

which is always the case given our starting assumptions. If μ approaches infinity, the $r\sigma^2$ terms are overcome in the above equations, and the inequality becomes an equality: hence, $\bar{X}$ goes to zero.

$$\lim_{\mu\to\infty} \mu^2 + r\sigma^2 = \lim_{\mu\to\infty} \mu\sqrt{\mu^2 + 2r\sigma^2}$$

□

C Establishing Equation 6

The expected utility to the buyer from equation 6 is

$$\mathsf{E}[TU_B \mid X_0] = X_0 + \left(\frac{X_0}{r} - \frac{\mu}{r^2}\right)(1-\phi) + \frac{\mu}{r}\int_0^\infty t\, e^{-rt} f(t)dt - K \tag{30}$$

The third term in equation 30 can be written as a modified Bessel function of the second kind (Gradshteyn and Ryzhik 1994, equation 3.471.9):

$$\int_0^\infty t\, e^{-rt} f(t)dt = \frac{2\,(X_0 - \bar{X})}{\sigma^2\sqrt{2\pi}}\, e^{\frac{-\mu\,(X_0 - \bar{X})}{\sigma^2}} \left(\frac{(X_0 - \bar{X})^2}{\mu^2 + 2\sigma^2 r}\right)^{\frac{1}{4}} K_{\frac{1}{2}}\left(\frac{(X_0 - \bar{X})}{\sigma^2}\sqrt{\mu^2 + \sigma^2 r}\right)$$

The Bessel function can be re-written as (Gradshteyn and Ryzhik 1994, equation 8.469.3):

$$K_{\frac{1}{2}}(z) = \sqrt{\frac{\pi}{2}}\,\frac{e^z}{\sqrt{z}}$$

We can then substitute the actual expression and further simplify:

$$\int_0^\infty t\, e^{-rt} f(t)dt = \phi\,\frac{(X_0 - \bar{X})}{\sqrt{\mu^2 + 2\sigma^2 r}}$$

The whole expression appears in equation 7.

D Proof of Equation 13

The probability of buying from equation 13 is:

$$P(\text{Buy}_{ij}) = \sum_{x \geq \tau} p_{ij} \, (1 - p_{ij})^x$$

If we expand the sum, and multiply the numerator and denominator by $(1 - p_{ij})^\tau$, we have

$$p_{ij}((1 - p_{ij})^\tau + (1 - p_{ij})^{\tau+1} + \ldots) = p_{ij}(1 - p_{ij})^\tau (1 + (1 - p_{ij}) + (1 - p_{ij})^2 + \ldots)$$

This is the same as a normal infinite geometric series. We can then simplify this to:

$$p_{ij}(1 - p_{ij})^\tau \left(\frac{1}{1 - (1 - p_{ij})} \right) = (1 - p_{ij})^\tau$$

which is what appears on the right hand side of equation 13.

E List of Movies Used in Analysis

2 FAST 2 FURIOUS	HOUSE OF 1,000 CORPSES
25TH HOUR	HOW TO LOSE A GUY IN 10 DAYS
ABOUT SCHMIDT	HUNTED
ADAPTATION	IDENTITY
ANALYZE THAT	IRREVERSIBLE
ANGER MANAGEMENT	JERRY SEINFELD COMEDIAN
ANIMATRIX	JUST MARRIED
ANTWONE FISHER	KID STAYS IN THE PICTURE
BASIC	LAUREL CANYON
BEND IT LIKE BECKHAM	LIFE OF DAVID GALE
BOAT TRIP	LORD OF THE RINGS: TWO TOWERS
BOWLING FOR COLUMBINE	LOST IN LA MANCHA
BRINGING DOWN THE HOUSE	LOVE LIZA
BULLETPROOF MONK	MALIBU'S MOST WANTED
CATCH ME IF YOU CAN	MAN APART, A
CHICAGO	MIGHTY WIND
CONFESSIONS OF A DANGEROUS MIND	NARC
CONFIDENCE	OLD SCHOOL
CORE	PHONEBOOTH
DADDY DAY CARE	PIANIST
DANCER UPSTAIRS	PUNCH-DRUNK LOVE
DAREDEVIL	QUIET AMERICAN
DARK BLUE	RABBIT-PROOF FENCE
DELIVER US FROM EVA	RAISING VICTOR VARGAS
DIE ANOTHER DAY	REAL WOMEN HAVE CURVES
DREAMCATCHER	RECRUIT
DYSFUNKTIONAL FAMILY	RUSSIAN ARK
FINAL DESTINATION 2	SECRETARY
FRIDA	SHANGHAI KNIGHTS
GANGS OF NEW YORK	SHAPE OF THINGS
GODS AND GENERALS	SOLARIS
GOOD THIEF	SPIDER
GURU	SPIRITED AWAY
HE LOVES ME/HE LOVES ME NOT	SPUN
HEAD OF STATE	TALK TO HER
HEAVEN	TEARS OF THE SUN
HOLES	TWO WEEKS NOTICE
HOURS	VIEW FROM THE TOP