Marketing Science

Modeling DVD Preorder and Sales: An Optimal Stopping Approach

Sam K. Hui, Jehoshua Eliashberg, Edward I. George,

Please scroll down for article—it is on subsequent pages

Modeling DVD Preorder and Sales: An Optimal Stopping Approach

Sam K. Hui, Jehoshua Eliashberg, Edward I. George

The Wharton School of the University of Pennsylvania, Philadelphia, Pennsylvania 19104
{kchui@wharton.upenn.edu, eliashberg@wharton.upenn.edu, edgeorge@wharton.upenn.edu}

When a DVD title is announced prior to actual distribution, consumers can often preorder the title and receive it as soon as it is released. Alternatively, once a title becomes available (i.e., formally released), consumers can obtain it upon purchase with minimal delay. We propose an individual-level behavioral model that captures the aggregate preorder/postrelease sales of motion picture DVDs. Our model is based on an optimal stopping framework. Starting with the utility function of a forward-looking consumer, and allowing for consumer heterogeneity, we derive the aggregate preorder/postrelease sales distribution. Even under a parsimonious specification for the heterogeneity distribution, our model recovers the typically observed temporal pattern of DVD preorder and sales, a pattern which exhibits an exponentially increasing number of preorder units before the release, peaks at release, and drops exponentially afterward. Using data provided by a major Internet DVD retailer, we demonstrate a number of important managerial implications stemming from our model. We investigate the role of preorder timing through a policy experiment, estimate residual sales, and forecast post-release sales based only on preorder information. We show that our model has substantially better predictive validity than benchmark models.

Key words: optimal stopping; timing model; online retailing; motion picture
History: This paper was received March 3, 2007, and was with the authors 3 months for 2 revisions; processed by Gerard Tellis. Published online in *Articles in Advance* July 31, 2008.

1. Introduction

The home video market is an important ancillary market for the motion picture industry (Eliashberg et al. 2006). In 2005, it generated about $23.8 billion in total revenues (Standard and Poor's 2006), whereas theatrical ticket revenues only accounted for about $9.0 billion (http://www.mpaa.org). Despite its economic importance, the motion picture home video market has received less attention among marketing researchers than the theatrical market, where researchers have offered rich behavioral explanations (e.g., Jones and Ritz 1991, Sawhney and Eliashberg 1996, Zufryden 1996) as well as systematic explanations (e.g., Jedidi et al. 1998) for box office revenue patterns. The potential for modeling home video sales is suggested by remarkably similar temporal patterns of DVD sales across different titles. For example, the weekly preorder/sales pattern (from a major Internet retailer) for the title *24 Hour Party People* shown in Figure 1 is typical—(preorder) sales increase exponentially before the release week, reach their peak during the week of release, and then drop exponentially afterward.

In previous literature, researchers have used population-level statistical models to describe the temporal diffusion patterns of media products such

as in Figure 1, which usually differs from the conventional S-shaped product life cycle (e.g., Hauser et al. 2006). For instance, Lehmann and Weinberg (2000) modeled the temporal patterns of home video sales revenues after release using an exponential distribution and obtained a reasonably adequate fit. In a similar vein, Moe and Fader (2002) modeled the temporal pattern of music CD sales using a mixture-Weibull model. Their model is based on two segments of consumers: innovators, who may preorder CDs, and followers, whose actions are influenced by word-of-mouth information and so only purchase CDs after their release. The purchase timing decisions for each segment of consumers are then assumed to follow two separate Weibull distributions. With this specification, Moe and Fader (2002) reported that their model was able to capture many different types of preorder/sales patterns. They further demonstrated how retailers may use preorder information to predict the (post-release) CD sales.

In both of the above studies, the proposed models are not based on individual-level decision making, but instead on a prespecified functional form of the sales curve (e.g., the exponential distribution in Lehmann and Weinberg 2000) or on a population-level behavioral model (Moe and Fader 2002). In this paper, we take a different approach by developing a

Figure 1   Sales Pattern for the DVD Title *24 Hour Party People*
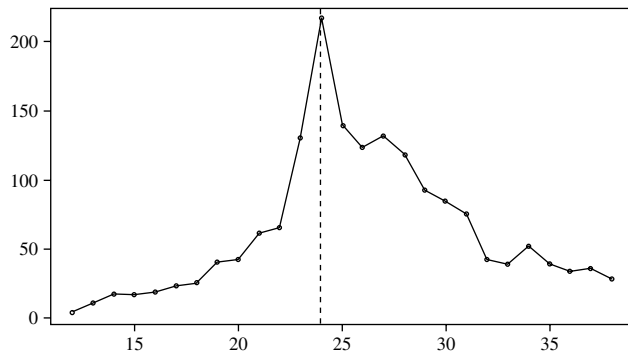
Notes. The dotted line denotes the release date of the DVD. The *y*-axis denotes weekly unit sales, whereas the *x*-axis is the number of weeks since the movie's release.

model based on individual-level "rational" behavior to explain the pattern shown in Figure 1. Instead of using an a priori functional form to "fit" the sales curve, we begin by specifying a utility function for an individual consumer. We then derive the utility-maximizing decision rule for each consumer (i.e., her optimal purchase timing), specify the degree of heterogeneity across consumers, and aggregate across their individual decisions to derive the aggregate sales pattern. Thus, the functional form that describes the aggregate sales pattern is an *outcome* of our model, rather than an a priori assumption. In this respect, our model is similar in spirit to recent NEIO models in marketing (e.g., Chintagunta et al. 2006), where empirical patterns are often explained as outcomes of consumer utility-maximizing decisions.

At the heart of our model is a forward-looking consumer (e.g., Song and Chintagunta 2003, Sun et al. 2003) interested in a DVD title, who visits a DVD ordering Web page at random time intervals, and who does not remember the DVD release date. On each visit, she decides whether to preorder/purchase the DVD instantly, or waits until her next visit to reconsider. Under any general interarrival distribution, the consumer's decision then corresponds to an optimal stopping problem (Chow et al. 1971), and the optimal solution is to follow a "threshold" rule (Ferguson 2000), i.e., to preorder the DVD if the consumer arrives within a certain time from the release date. Based on the optimal stopping rule, we obtain the distribution of an individual consumer's purchase timing. Allowing heterogeneity across consumers, we obtain the aggregate temporal sales pattern that recovers the qualitative characteristics of the observed pattern in Figure 1. In addition to this basic model, we further extend our framework to allow for a segment of consumers who do remember the DVD release date and buy directly at that time.

We calibrate our basic (DVD-I) and extended (DVD-II) models to a data set containing weekly

DVD preorder/sales information provided by a major Internet retailer. Both models are seen to substantially outperform benchmark models such as Moe and Fader (2002) and the Weibull-Gamma model (Jaggia and Thosar 1995). We further demonstrate some potential managerial implications from our study. We investigate the role of preorder timing through a policy experiment, estimate residual sales, and forecast post-release sales based only on preorder information.

The remainder of this paper is organized as follows. In §2 we develop our model of DVD purchase timing. Section 3 provides an overview of our data set, along with key summary statistics. In §4, we compare our model to other benchmark models to assess model performance, then interpret the parameters estimates. In §5, we demonstrate some potential managerial implications stemming from our model. Finally, §6 concludes and outlines directions for future research.

2.   Model
In this section, we describe our model of DVD purchase timing in detail. For clarity of exposition, we first focus on an individual consumer in §§2.1–2.4. We describe our model setup and notations in §2.1, and specify the consumer's utility function in §2.2. In §2.3, we show how the consumer's purchase timing decision can be viewed as an optimal stopping problem in which the solution takes the form of a "threshold" rule (Ferguson 2000). In §2.4, the distribution of an individual consumer's purchase timing is derived under the assumption of exponentially distributed interarrival times. In §2.5, we account for consumer heterogeneity and derive the sales distribution by aggregating across consumers. This gives us the DVD-I model and its associated likelihood function. In §2.6, we extend the DVD-I model to the DVD-II model which allows for a second segment of consumers who remember the release date and come back in that week to purchase the DVD.

2.1.   Model Setup
Figure 2 shows the timing of the important events in our model corresponding to a specific movie. Time $t = 0$ represents the theatrical release date. At time $t = l$, the Internet retailer allows consumers to begin preordering the DVD. The DVD is released (and shipped to fulfill preorders) at time $t = k$, which is called the "window" in the movie industry.[1]

---

[1] Note that we make the assumption of instantaneous order fulfillment. That is, we assume that if a DVD is ordered after its release, it will be immediately shipped and the consumer will receive it instantaneously. Similarly, if a DVD is ordered before its release, the order will be shipped and will arrive on the release date. This assumption is made for analytical convenience; similar results hold

Figure 2   Timeline of the Major Events in Our Model

Figure 3   Arrival Process of a Consumer to the DVD Ordering Webpage
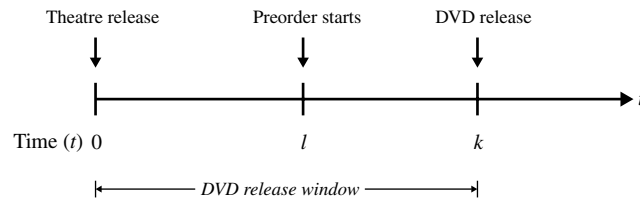
Next, we specify the arrival process by which a consumer visits the DVD ordering Web page. On each visit, the consumer decides whether to purchase the DVD, or to wait till her next visit to reconsider. As shown in Figure 3, we index the consumer's visits by $n = 0, 1, 2, \ldots$, where the nth visit takes place at time t_n. The interarrival time between the nth and the $(n+1)$th visit is denoted by w_n (i.e., $w_n = t_{n+1} - t_n$), which are assumed to be drawn i.i.d. from a general distribution (known both to the consumer and the researcher). To denote this distribution, we use $G(w)$ and $g(w)$ for cumulative distribution function and the probability distribution function, respectively.
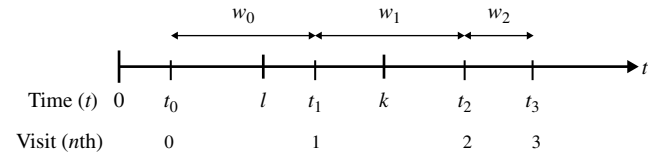
2.2. Specification of Consumer Utility

We use the following notation to specify the utility of the DVD to the focal consumer. Let $\tilde{u}(s)$ be the instantaneous utility from the DVD at time s after obtaining it. Let y be the price[2] of the DVD, and let $\tilde{x}_y(s)$ be the instantaneous utility from an outside option (obtainable at price y) at time s after obtaining it. This specification allows us to compare the utility of the DVD to the utility of other outside options that the consumer would forego by purchasing the DVD. For instance, the consumer could have used that money to buy a CD or a new book.

We assume that consumers discount future utilities by a constant discounting factor, so that (relative to the time point $t = 0$) utilities obtained at time t are discounted by a factor $e^{-\gamma t}$, where $\gamma > 0$. This assumption is widely used in intertemporal choice models (e.g., Frederick et al. 2002, Lowenstein and Prelec 1992). The discount rate γ is taken as a constant and is assumed to be fixed across all titles and all consumers. We further assume that Internet search cost

is zero (e.g., Bakos 1997, Wu et al. 2004); i.e., the consumer does not incur any cost in visiting the website.

To simplify notation, we let

$$u = \int_0^\infty e^{-\gamma s} \tilde{u}(s)\, ds, \tag{1}$$

$$x = \int_0^\infty e^{-\gamma s} \tilde{x}_y(s)\, ds, \tag{2}$$

which correspond, respectively, to the "net present utility" of the DVD and of the outside option (relative to utility at time $t = 0$). We further assume that $u > x$ for our focal consumer, given his/her interest in the DVD title. Similar to the "participation constraint" in economics (Mas-Colell et al. 1995), this assumption ensures that the DVD is sufficiently attractive that the consumer will buy it—the question is when. Later in §2.5, we describe how we can estimate market potential, i.e., the number of consumers who have $u > x$, for each DVD in our data set.

We now specify the net utility (relative to time $t = 0$)[3] that the consumer receives if she preorders/purchases the DVD on her nth visit to the DVD Web page. The consumer may only preorder/purchase the DVD if $t_n \geq l$. Thus, if $t_n < l$, no decision needs to be made. In the discussion below, we separately consider the two remaining cases: (i) $l \leq t_n < k$ (prerelease), and (ii) $t_n > k$ (post-release).

First, consider $l \leq t_n < k$. In this case, as shown in Figure 4, the consumer preorders and pays for the DVD at time t_n, and receives the DVD at time $t = k$. Her net utility (relative to time $t = 0$), Y_n, is a sum of two parts: the forgone (discounted) utility of the outside option, and the (discounted) net utility from the DVD which will arrive at time $t = k$. Formally,

$$Y_n = -\int_{t_n}^\infty e^{-\gamma t} \tilde{x}_y(t - t_n)\, dt + \int_k^\infty e^{-\gamma t} \tilde{u}(t - k)\, dt. \tag{3}$$

---

if we allow for a constant shipping delay. (In the actual data, DVDs are typically shipped within 1–3 days, which is very short compared to the time unit of the data set which is weeks.) Further, we assume that there is no scarcity for the DVD, an assumption which we have verified with our data provider.

[2] Throughout this paper, we assume that the price of a DVD remains constant over the preorder period and afterward. We have verified this assumption empirically with our data provider. We further assume that consumers' expectation of prices are rational; i.e., they also expect prices to remain constant during their planning horizon.

---

[3] Note that to ensure that the utilities mentioned are in the same units and thus directly comparable, all the utilities in our model are discounted relative to the time $t = 0$, the theatrical release time. This specification is done without loss of generality, because discounting is a monotonic transformation and thus the relative order of utilities is preserved. As a concrete example, assume that at time t_1, a consumer is deciding between options A and B, with (instantaneous) utilities U_A and U_B, respectively. Because $U_A > U_B \Leftrightarrow U_A e^{-\gamma t_1} > U_B e^{-\gamma t_1}$, the consumer's decision is invariant to the time point relative to which utilities are discounted.

Figure 4   Consumer's Net Utility for Preordering the DVD
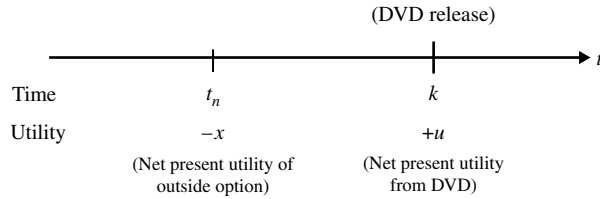
Changing variables in the integral (with $z = t - t_n$ and $v = t - k$), we obtain

$$Y_n = -e^{-\gamma t_n}x + e^{-\gamma k}u, \tag{4}$$

where u and x are defined in Equations (1) and (2).

In the other case, when $t_n > k$ (as shown in Figure 5), the consumer pays for and receives the DVD at the same time t_n. In this case, we have

$$Y_n = -\int_{t_n}^{\infty} e^{-\gamma t}\tilde{x}_y(t - t_n)\,dt + \int_{t_n}^{\infty} e^{-\gamma t}\tilde{u}(t - t_n)\,dt.$$

Changing variables in the integral (with $z = t - t_n$), we have

$$Y_n = -e^{-\gamma t_n}x + e^{-\gamma t_n}u = e^{-\gamma t_n}(u - x), \tag{5}$$

where, again, u and x are defined in Equations (1) and (2).

Together, Equations (4) and (5) give us a complete specification for the consumer's discounted net utility Y_n (with respect to $t = 0$):

$$Y_n = \begin{cases} -e^{-\gamma t_n}x + e^{-\gamma k}u & \text{for } l \le t_n < k, \\ -e^{-\gamma t_n}(x - u) & \text{for } t_n \ge k. \end{cases} \tag{6}$$

Note that our utility specification above is similar to other specifications based on discount utilities and has been employed to derive optimal stopping rules (Allaart 2004, Dubins and Teicher 1967). From the previous literature on optimal stopping theory, the existence of an optimal solution to our problem has been proven. A formal proof is available upon request.

2.3. Consumer Preorder/Purchase Timing Decisions

Under our model setup, the consumer faces an optimal stopping problem (Chow et al. 1971) when deciding when to place an order for a DVD. Intuitively,

Figure 5   Consumer's Net Utility for Purchasing the DVD After Its Release
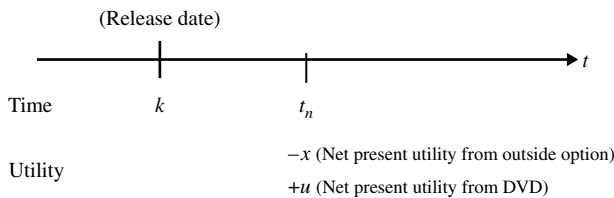
the consumer prefers to (pre-) order the DVD as close to the actual release date as possible. She derives the most utility if she purchases the DVD at the release date; i.e., there is no advantage for her to preorder a DVD because she has to pay in advance while only receiving the DVD on the release date. However, assuming her time of visit to the website is stochastic, if she decides to wait, she may "miss" the release date and have to purchase afterward (therefore receiving a lower utility due to discounting). As we show in the following derivation, the consumer's optimal decision rule is characterized by a "threshold rule" (e.g., Ferguson 2000); i.e., she will preorder the DVD if she happens to visit the website at time t_n, and the gap between time t_n and the release date k is *smaller* than a predetermined threshold.

We solve for the optimal stopping visit n^* by considering three cases: (i) $t_n < l$, (ii) $t_n \ge k$, and (iii) $l \le t_n < k$. In the first case ($t_n < l$), no decision needs to be made because preorder is not yet allowed. In the second case ($t_n \ge k$), there is no advantage for the consumer to wait any longer, and she should purchase the DVD immediately. Formally, for any $t_{n'} \ge t_n \ge k$,

$$Y_{n'} = -e^{-\gamma t_{n'}}(x - u) \le -e^{-\gamma t_n}(x - u) = Y_n. \tag{7}$$

The third case ($l \le t_n < k$) is most interesting because the consumer has to trade off between the opportunity cost of ordering now (and paying in advance) versus the risk of waiting longer than necessary to enjoy the DVD. This is similar to the idea of a "delay premium" (e.g., Lowenstein 1988); i.e., some consumers are willing to pay a premium to ensure that their enjoyment from the DVD will not be delayed. According to a manager at our data provider, our specification is consistent with their understanding of why consumers preorder DVDs in the absence of preorder discounts and stock-out possibilities, the environment under which our data set was collected.

Thus, the consumer makes a (pre-)order/wait decision each time she visits the DVD Web page. She buys the DVD at time t_n if her utility of buying exceeds the expected utility of "waiting." Because our model satisfies the property of "monotonicity" (see Appendix I), the 1-step look-ahead rule is equivalent to the globally optimal solution.[4] More precisely, the consumer buys the DVD at visit n (i.e., time t_n) if her discounted net utility of buying now exceeds the expected utility of buying at visit $n + 1$, i.e., if $Y_n \ge E(Y_{n+1} \mid \Gamma_n)$, where Γ_n denotes the probability measure with respect to the information available up to the nth visit.

[4] The proof one-step look-ahead rule is the optimal solution to the monotonic optimal stopping problem and can be found in Ferguson (2000).

Letting $\Delta_n = k - t_n$ (the gap between the current time and the release date), we have

$Y_{n+1} \mid$ buy at visit n

$$= \begin{cases} e^{-\gamma t_n}(e^{-\gamma \Delta_n} u - e^{-\gamma w_n} x) & \text{for } w_n \leq \Delta_n, \\ e^{-\gamma t_n}(e^{-\gamma w_n} u - e^{-\gamma w_n} x) & \text{for } w_n > \Delta_n, \end{cases} \quad (8)$$

$$\Rightarrow E(Y_{n+1} \mid \Gamma_n) = e^{-\gamma t_n} \left[\int_0^{\Delta_n} (e^{-\gamma \Delta_n} u - e^{-\gamma w_n} x) g(w_n) dw_n \right.$$
$$\left. + \int_{\Delta_n}^{\infty} e^{-\gamma w_n}(u - x) g(w_n) dw_n \right], \quad (9)$$

where $g(.)$ is the probability density function of the interarrival time defined in §2.2. Equations (8) and (9) precisely describe the consumer's trade-off discussed above. Note that two implicit assumptions are made here. First, we assume that consumers know k. This assumption is generally valid given that the release date is displayed prominently on the DVD page. The second assumption is that the consumers do not remember the release date. We later relax this assumption in the DVD-II model of §2.6, where we allow for a segment of consumers who do remember the DVD release date, and thus purchase during the week of the release.

After some algebraic manipulations shown in Appendix II, we obtain

$$Y_n \geq E(Y_{n+1} \mid \Gamma_n)$$
$$\Leftrightarrow \frac{x}{u}[1 - E(e^{-\gamma w_n})] \leq e^{-\gamma \Delta_n}[1 - G(\Delta_n)]$$
$$- \int_{\Delta_n}^{\infty} e^{-\gamma w_n} g(w_n) dw_n \equiv H(\Delta_n), \quad (10)$$

where the second term of the inequality (10) is denoted by $H(\Delta_n)$. Because the left side of inequality (10) is independent of Δ_n, we only need to show that $H(\Delta_n)$ is a decreasing function of Δ_n to prove that the optimal solution takes the form of a threshold rule. Differentiating $H(\Delta_n)$ with respect to Δ_n,
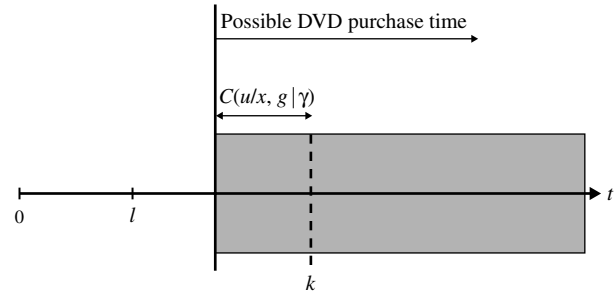
$$\frac{dH(\Delta_n)}{d\Delta_n} = e^{-\gamma \Delta_n}(-g(\Delta_n))$$
$$+ (1 - G(\Delta_n))(e^{-\gamma \Delta_n})(-\gamma) + e^{-\gamma \Delta_n} g(\Delta_n)$$
$$= -\gamma e^{-\gamma \Delta_n}(1 - G(\Delta_n)) < 0. \quad (11)$$

Thus, the optimal stopping rule is to buy the DVD whenever $\Delta_n = k - t_n$ is below a threshold that depends on the ratio (u/x), γ, and the interarrival time distribution g. We write

$$n^* = \min\{n \mid \Delta_n \leq C((u/x), g \mid \gamma)\}, \quad (12)$$

where $C((u/x), g \mid \gamma)$ captures the dependence of the threshold on the ratio (u/x) and on the interarrival

Figure 6 Illustration of the Threshold Rule

Note. The shaded area indicates the possible time of DVD preorder/purchase for a specific consumer.

time distribution g.[5] We denote the corresponding purchase time as t_{n^*}.

Combining the above three cases, the consumer preorders a DVD if she visits the DVD Web page at time $t \geq d$, where $d = \max\{l, k - C((u/x), g \mid \gamma)\}$ is the larger of the time when preorder is available and the threshold in Equation (12). This result is depicted graphically in Figure 6; the shaded area represents the possible DVD purchase time for our focal consumer.

2.4. Individual-Level Distribution of Purchase Timing Under Exponential Interarrival Times

Following previous literature on Internet visit behavior (Moe and Fader 2004, Park and Fader 2004), we restrict our attention, in this subsection, to the case where the consumer's interarrival times follow an exponential distribution with parameter λ, and derive the corresponding distribution of her purchase time. The assumption of exponentially distributed interarrival times has been widely used in marketing and operations research in models of consumers' interarrival/interpurchase times (e.g., Morrison and Schmittlein 1988) and other duration models (e.g., Morrison and Schmittlein 1980).

Substituting the exponential density, $g(w_n) = \lambda e^{-\lambda w_n}$ into Equation (10) and performing the algebraic simplifications outlined in Appendix III, the following threshold decision rule is obtained. The consumer will preorder the DVD if

$$\Delta_n \leq \frac{1}{\lambda + \gamma} \ln\left(\frac{u}{x}\right) = C. \quad (13)$$

The form of the threshold rule above leads to two immediate, intuitive insights. First, the threshold is an increasing function of u/x. Thus, the more net discounted utility a consumer derives from the DVD (relative to her outside option), the more likely she is to preorder. Second, the threshold is a decreasing

---

[5] Dependence of the threshold on the discount rate γ is not denoted explicitly because it is treated as a constant in our model, as we discussed earlier.

function of λ; the higher a consumer's frequency of visits to the DVD website, the less likely she is to preorder. Intuitively, because more frequent visitors have more chances to visit the DVD Web page, they can afford to wait a little longer in the hope of preordering closer to the actual release date. Their risk of missing the release date is lower.

Considering the cumulative distribution function (CDF) of t_{n^*} (the purchase time) conditional on t_{n^*-1} (the time of the last visit before purchase), we have

$$F_{t_{n^*}}(t \mid t_{n^*-1}) = 1 - P(t_{n^*} > t \mid t_{n^*-1})$$
$$= 1 - P(w_{n^*-1} > t - t_{n^*-1} \mid w_{n^*-1} > d - t_{n^*-1})$$
$$= 1 - e^{-\lambda[t-d]} \quad \text{for } t > d, \quad (14)$$

where $d = \max(l, k - C)$ as defined earlier. Note that in the second step above, we utilize the "memoryless" property of the exponential distribution. Because the expression in Equation (14) is independent of t_{n^*-1}, we also have the unconditional CDF of t_{n^*}, $F_{t_{n^*}}(t) = 1 - e^{-\lambda[t-d]}$. Thus, the purchase time t_{n^*} follows a shifted exponential distribution with parameter λ; i.e., $t_{n^*} + d$ follows an exponential distribution with parameter λ. The corresponding probability density associated with Equation (14) is the (shifted) exponential density, $f_{t_{n^*}}(t) = \lambda e^{-\lambda(t-d)}$.

2.5. Consumer Heterogeneity: DVD-I Model

We now show that under a simple and reasonable specification of heterogeneity across consumers, namely that $(u/x)_i$ (where i indexes consumer) follows a Pareto distribution with parameter v, i.e., $f((u/x)_i) = v((u/x)_i)^{-v-1}$ for $(u/x)_i \in [1, \infty],$[6] our model can generate the observed temporal pattern of DVD preorder and sales. Using a change of variable, $\ln((u/x)_i)$ follows an exponential distribution with rate parameter v.[7] The Pareto distribution has been found to be a good approximation in various economic contexts involving heterogeneity between units, e.g., wealth distributions (Wold and Whittle 1957) and income distributions (Clementi and Gallegati 2005).

Although we allow $(u/x)_i$ to vary across the population of DVD buyers, we keep the parameter λ fixed. We make this simplifying assumption for two reasons. First, models with a similar homogeneity assumption have been successfully applied to model box office

sales patterns (e.g., Sawhney and Eliashberg 1996). Second, assuming a homogeneous λ maintains analytical tractability when deriving the marginal distribution of purchase times across the population.

Under these specifications,

$$(u/x)_i \sim \text{Pareto}(v) \qquad \ln((u/x)_i) \sim \exp(v), \quad (15)$$

the distribution of the threshold across consumers (C_i denotes the threshold for the ith consumer) is

$$C_i = \frac{1}{\lambda + \gamma} \ln(u_i/x) \;\Rightarrow\; C_i \sim \exp(v(\lambda + \gamma))$$
$$\Rightarrow\; C_i \sim \exp(\tau), \quad (16)$$

where $\tau = v(\lambda + \gamma)$. The marginal distribution of purchase timing across the population, $f(t)$, is then obtained by integrating the distribution function associated with Equation (14) across all consumers (indexed by i), under the mixture distribution specified by Equation (16):

$$f(t) = \int_{-\infty}^{\infty} f_i(t \mid C_i) \pi(C_i) \, dC_i$$
$$= \int_0^{\infty} \lambda e^{-\lambda(t-d)} 1_{\{t>d\}} \tau e^{-\tau C_i} \, dC_i$$
$$= \lambda \tau e^{-\lambda t} \int_0^{\infty} e^{\lambda d - \tau C_i} 1_{\{t>d\}} \, dC_i. \quad (17)$$

From this expression we obtain the following closed-form expressions for the marginal distribution $f(t)$ and the corresponding cumulative distribution function $F(t)$ of the DVD purchase time (see Appendix IV):

$$f(t) = \begin{cases} \left[\dfrac{\lambda \tau}{\lambda + \tau} e^{\lambda k} (1 - e^{-(\lambda + \tau)(k-l)}) + \lambda e^{\lambda l - \tau(k-l)} \right] e^{-\lambda t} \\ \qquad\qquad\qquad\qquad\qquad t \geq k, \\[2mm] \dfrac{\lambda \tau}{\lambda + \tau} e^{-\tau(k-t)} + \dfrac{\lambda^2}{\lambda + \tau} e^{-\lambda(t-l) - \tau(k-l)} \\ \qquad\qquad\qquad\qquad\qquad l \leq t < k, \end{cases} \quad (18)$$

$$F(t) = \frac{\lambda}{\lambda + \tau} e^{-\tau k} \left[e^{\tau t} - e^{\tau l} + e^{(\tau + \lambda)l}(e^{-\lambda l} - e^{-\lambda t}) \right]$$
$$\text{for } t \leq k,$$
$$= F(k) + \frac{A}{\lambda}(e^{-\lambda k} - e^{-\lambda t}) \quad \text{for } t > k, \quad (19)$$

where $A = (\lambda \tau / (\lambda + \tau)) e^{\lambda k} (1 - e^{-(\lambda + \tau)(k-l)}) + \lambda e^{\lambda l - \tau(k-l)}$.

The properties of $f(t)$ are interesting. First, the density is strictly exponentially decreasing in the region $t \geq k$. This is consistent with the observation in Lehmann and Weinberg (2000) that post-release revenues are exponentially decreasing. For $l \leq t < k$, $f(t)$ is a sum of two terms. The first term gives us the exponentially increasing preorder pattern that we observe in the data. The second term acts as a "correction term" due to the fact that preorders are available

---

[6] The range for $(u/x)_i$ here is consistent with our assumption that (among DVD buyers) the ratio of the utility from the DVD compared to the outside option, $(u/x)_i$, is larger than 1.

[7] In the Technical Appendix at http://mktsci.pubs.informs.org, we consider relaxing this assumption by allowing $\ln(u/x)_i$ to follow a gamma distribution. We find that it only provides a marginally better fit at the cost of the analytic tractability and computational efficiency of a closed-form solution. Thus, we focus on using the exponential distribution assumption in this paper.

Figure 7   **Distribution of Aggregate Sales Derived from Our Model (Equation (18)); $\lambda = 0.3$, $\tau = 0.4$)**
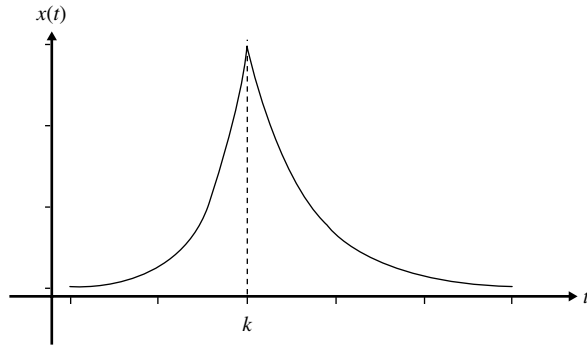
only from $t = l$ onward; its magnitude decreases if preorders are allowed earlier, and vanishes if preorders are always available (i.e., in the limit $l \to -\infty$).

The density $f(t)$ in Equation (18) is plotted in Figure 7. As can be seen, the pattern in Figure 7 closely resembles that of Figure 1; in particular, (preorder) sales increase exponentially before the release date and decrease exponentially after it. This provides some evidence that our model, based on the individual-level utility maximization and forward-looking behavior, provides an explanation for the pattern of the observed aggregate sales data.

By varying λ and τ and plotting the resulting density function, we can study our model properties more closely and further understand the sources of identification for our model parameters. Figures 8 and 9 plot $f(t)$ as λ and τ are varied, respectively. From the figures, we see that λ primarily controls the rate of decline, whereas τ mainly controls the rate of ascent. More specifically, larger values of λ and τ correspond to faster rates of decline and ascent, respectively. Thus, the rate of decline after release provides the source of identification for the parameter λ, whereas the rate of ascent prerelease provides the source of identification for the parameter τ.

Because our proposed model is defined in continuous time while actual preorder/sales data are usually recorded in discrete time units (e.g., week or day), we need to first discretize our model before

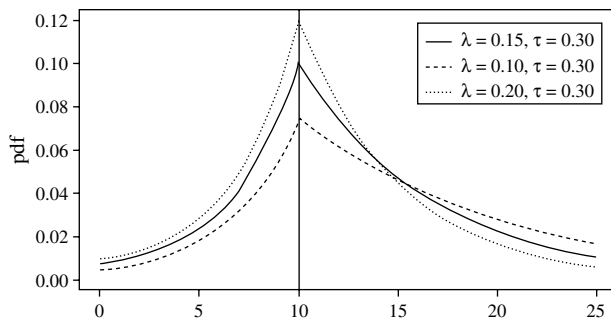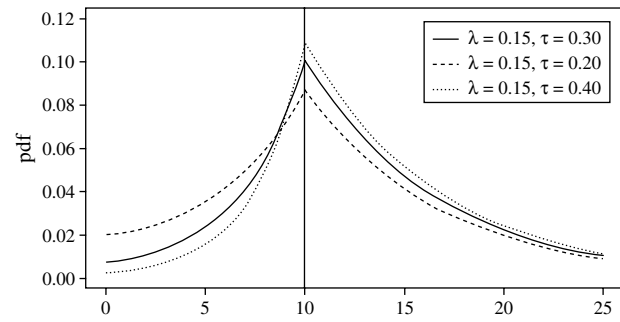**Figure 8**   **Plots of pdf by Varying λ**

Figure 9   **Plots of pdf by Varying τ**

calibrating our model parameters on actual data. This is the same issue faced by Moe and Fader (2002); to tackle this problem, we use an approach proposed by Schmittlein and Mahajan (1982) and applied in the paper by Moe and Fader (2002). First, we let M denote the market potential, i.e., the total number of consumers who have $u > x$, for a DVD. We let p_j denote the probability that an individual consumer preorders/purchases the DVD in week j so that

$$p_j = F(j+1) - F(j), \tag{20}$$

where $F(.)$ denotes the cumulative distribution function of the purchase timing (Equation (19)).

Finally, we derive the likelihood function of the data given model parameters (M, λ, τ). Let $\vec{y}_i = (y_{i,l}, y_{i,l+1}, \ldots, y_{i,T_i})$ be the vector of sales of a DVD, where $y_{i,j}$ denotes preorders/sales of the ith DVD at (calendar) week j, and T_i denotes the last week of the data collection window for the ith DVD. We have

DVD-I $l(\vec{y}_i \mid M, \lambda, \tau)$

$$= \frac{M!}{\prod_{j=l}^{T} y_{i,j}! \left(M - \sum_{j=l}^{T} y_{i,j}\right)!} \left[\prod_{j=l}^{T} (p_j)^{y_{i,j}} \right]$$

$$\cdot [1 - F(T+1)]^{M - \sum_{j=l}^{T} y_{i,j}}. \tag{21}$$

The parameters M, λ, τ for each DVD can then be estimated by maximizing Equation (21) given the weekly preorder/sales data. We refer to Equation (21) as the DVD-I model.

2.6. Model Extension: DVD-II
The DVD-I model makes the assumption that consumers do not remember the release date of a DVD. However, some consumers do in fact remember the release date, perhaps from their visit to the DVD page, or from prerelease media advertising. To account for this segment of consumers, we extend the DVD-I model to what we call the DVD-II model.

Similar to the specification used in the paper by Moe and Fader (2002), we denote the proportion of consumers who remember the DVD release date by ϕ. We assume that this particular segment of consumers will purchase the DVD during the release week; i.e., $t = k$. Under the DVD-II model, we obtain a mixture

specification (e.g., Muthen and Masyn 2005):

$$p'_j = [1 - \phi][F(j+1) - F(j)] \quad \text{for } j \neq k,$$
$$p'_k = [1 - \phi][F(k+1) - F(k)] + \phi, \tag{22}$$

where p'_j denotes the probability that an individual consumer preorders/purchases on week j. Finally, we derive the likelihood function associated with the DVD-II model:

$$\text{DVD-II} \quad l(\vec{y}_i \mid M, \lambda, \tau)$$
$$= \frac{M!}{\prod_{j=l}^{T} y_{i,j}! (M - \sum_{j=l}^{T} y_{i,j})!} \left[\prod_{j=l}^{T} (p'_j)^{y_{i,j}} \right]$$
$$\cdot [1 - F(T+1)]^{M - \sum_{j=l}^{T} y_{i,j}}. \tag{23}$$

The DVD-II model parameters M, ϕ, λ, τ for each DVD can then be estimated by maximizing Equation (23) given the preorder/sales data.

3. Data

We obtained our data set from a major Internet retailer for books, DVDs, music CDs, electronics, and other household products. Our data set contains weekly preorder and sales figures, along with the release date, for a sample of movie DVDs from February 2002 to November 2004. For each DVD, we obtained the complete preorder record from the week when preorder first became available, together with weekly postrelease sales up to the 15th week after release. Similar to the data-preparation procedure described in Moe and Fader (2002), we eliminated DVDs with incomplete preorder information or very sparse sales (i.e., less than 100 units sold during the release week, the criterion used in Moe and Fader 2002). Our final data set contains a total of 251 titles, which we use in all our subsequent analyses.

Table 1 presents key summary statistics of our data set. The median total DVD sales (from the start of the preorder period until the 15th week after release) is 3,084 units. Because our data set contains DVDs of blockbuster movies as well as those from independent distributors, total sales vary significantly across different titles, ranging from just 526 units to over 200,000 units sold. Averaged across titles, 15.7% of the preorders/sales occurred during the week of release throughout our data collection window.

In our data set, a DVD is available for preorder, on average, about 11.6 weeks before its release. The median window between theatrical and DVD release is about 22 weeks. This is roughly consistent with MPAA (2006), which reported that the average window between the theatrical and DVD release is about five months across all studios.

To study the empirical temporal pattern of preorders/sales in more detail, we plot the density of

Table 1 Summary Statistics from Our Data Set

|	Mean	Median	St. dev.	Min	Max
Total sales[a]	10,952.7	3,084.0	27,435.5	526.0	221,229.0
Number of preorder weeks	11.6	11.0	2.6	6.0	25.0
% of total sales during the release week	15.7	15.1	4.3	3.5	38.1
Window between theatrical and DVD release (weeks)	27.7	22.0	30.9	5.0	289.0

[a]Total sales represent the total number of units sold within the time period considered.

preorders/sales versus the number of weeks since a title's release date, aggregated across all titles. More specifically, we define y_{is} as the number of sales for the ith DVD during the sth week after its release. Thus, $s = 1$ denotes one week after the ith DVD's release, $s = 0$ denotes the week of release, and $s = -1$ denotes one week before the release. We define r_{is} as the proportion of sales for the ith DVD that occurred during the sth week after release; i.e., $r_{is} = y_{is} / \sum_s y_{is}$. We denote $r_s = (\sum_i r_{is})/N$ as the mean of r_{is} across all DVDs. A plot of r_s versus s is shown in Figure 10. From the figure, we see that across all DVDs, preorders/sales are generally exponentially increasing before release, and exponentially decreasing after release. This roughly replicates the preorder/sales pattern of the title *24 Hour Party People*, shown in Figure 1.

4. Results

In this section, we describe the results of applying our two models to our data set. In §4.1, we demonstrate the fit of both DVD-I and DVD-II models and compare them against two benchmarks—Moe and Fader (2002) and the Weibull-Gamma model. After

Figure 10 Temporal Density of Preorders/Sales Across All DVDs
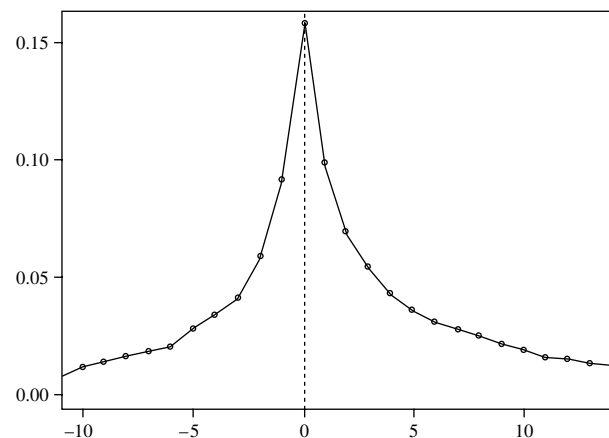
Notes. The x-axis denotes the number of weeks before/after release (the dotted line at $t = 0$ indicates the week of the DVD's release); the y-axis plots the density of preorders/sales.

validating the model's performance, we proceed to discuss parameter estimates in §4.2.

4.1. Model Comparison

As discussed in §2, the Moe and Fader (2002) model is a six-parameter model based on a mixture of two Weibull distributions: one representing the behavior of innovators, and the other representing the behavior of followers. The cumulative distribution function that it yields can be written as

$$F(t) = \begin{cases} \phi[1 - e^{-\lambda_1 t^{c_1}}] & \text{for } t < k, \\ \phi[1 - e^{-\lambda_1 t^{c_1}}] + (1 - \phi)[1 - e^{-\lambda_2 (t-k)^{c_2}}] & \\ & \text{for } t \geq k, \end{cases} \quad (24)$$

where ϕ is the fraction of buyers associated with the innovator segment; λ_1, c_1 are the Weibull parameters associated with the innovators segment; and λ_2, c_2 are the Weibull parameters associated with the followers segment.

We also considered the benchmark model used by Moe and Fader (2002), namely the Weibull-Gamma model. This model is based on the assumption that consumers preorder/purchase based on a Weibull distribution with rate parameters that vary according to a Gamma distribution. The cumulative distribution can be derived as (see Moe and Fader 2002):

$$F(t) = 1 - \left(\frac{a}{a + t^c}\right)^r \quad \text{for } t \geq l. \quad (25)$$

We compare the four candidate models (DVD-I, DVD-II, Moe and Fader 2002, Weibull-Gamma) with three common model comparison criteria, log-likelihood (LL), Bayesian Information Criteria (BIC) (Schwarz 1978), mean absolute percentage error (MAPE), and a fourth criterion which we call mean relative absolute error (MRAE). We define MRAE to be the average over all DVDs of

$$\text{RAE}_i = \frac{\sum_t |y_{it} - \hat{y}_{it}|}{\sum_t y_{it}}, \quad (26)$$

where y_{it} denotes the actual sales for the ith DVD during week t, and $\hat{y}_{it}$ denotes the corresponding predicted value. As opposed to MAPE, which is very sensitive to weeks with very few sales,[8] each RAE_i measures total absolute deviations as a percentage of total sales. Thus MRAE, together with MAPE, allows us to obtain a more complete assessment and comparison of model fit.

The values of the four model comparison criteria, LL, BIC, MAPE, and MRAE, are presented in Table 2 for the four models DVD-I, DVD-II, Moe and Fader (2002), and the Weibull-Gamma. We find that the
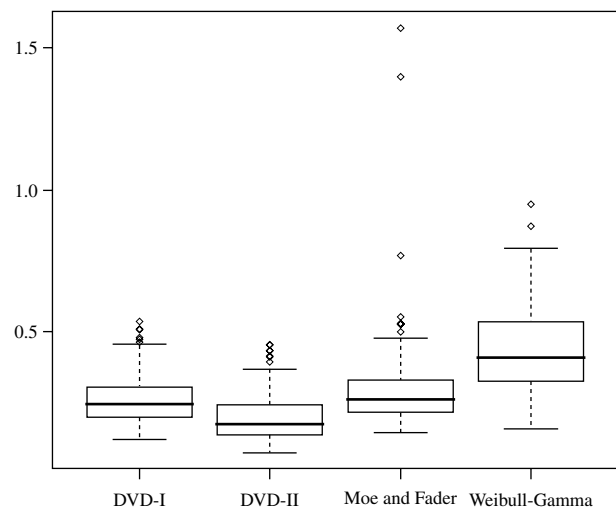
[8] For example, a prediction of 10 units of sales on a week with actual sales of 5 units would generate an absolute percentage error of $(10-5)/5 = 100\%$.

Table 2　Comparison with Benchmark Models

	Log-likelihood	BIC	MAPE (%)	MRAE (%)
DVD-II	−184,273.1	379,641.3	34.7	19.8
DVD-I	−224,003.4	456,328.1	37.4	26.5
Moe and Fader (2002)	−357,864.7	732,372.1	64.0	29.5
Weibull-Gamma	−499,907.6	1,010,910.4	101.2	43.7

DVD-II model performs better on all measures than the DVD-I model, which in turn outperforms both the Moe and Fader (2002) and the Weibull-Gamma model. The DVD-II model obtains the highest log-likelihood, the lowest MAPE, and the lowest MRAE in comparison to all models. Compared to DVD-I, it has a lower BIC, which indicates that the two-segment model captures the actual data more closely than the single-segment model assumed in DVD-I. The overall superiority of DVD-II is further revealed by boxplots of the RAE for each DVD across the four models, shown in Figure 11. In particular, DVD-II appears to yield uniformly lower RAE quantiles than the other three models.

The DVD-II model (or even the single segment DVD-I model) performs considerably better than Moe and Fader's (2002) model, even with a smaller number of parameters per DVD (for each title, DVD-I uses three parameters, DVD-II has four parameters, whereas the Moe and Fader 2002 model uses six parameters). We believe that our models perform better because the behavioral premise behind them is more consistent with the DVD context. Specifically, Moe and Fader's (2002) model is based on the assumption that there are two segments of consumers: innovators who try new CDs, and followers who imitate innovators (presumably because of positive word-of-mouth). This dimension of word-of-mouth effects is generally less relevant for the

Figure 11　Boxplot of RAE Across DVDs for Each of the Four Models (DVD-I, DVD-II, Moe and Fader 2002, Weibull-Gamma)

Table 3   **Summary of Parameter Estimates**

	Mean	St. dev.	Median
M	11,643.6	28,488.1	3,499.4
ϕ	0.062	0.036	0.057
λ	0.165	0.063	0.158
τ	0.273	0.094	0.262

DVD market, where sequential release is more prevalent across distribution windows. Before a DVD is released, the movie has already been shown in a theatre (typically for 5 months), and extensive word-of-mouth information, either from other consumers who watched the movie or from movie critics, has already been widely available long before the DVD is released. Thus, it is reasonable to expect that imitation effects, the key mechanism behind the model of Moe and Fader (2002), is less relevant in the DVD setting than in the music CD context that they focused on. Finally, the Weibull-Gamma model performed worst, which is not surprising because it does not take into account differences between the preorder and post-release periods.

4.2. Parameter Estimation and Interpretation

We now turn to interpreting the parameter estimates for the DVD-II model which are summarized in Table 3. The estimated market potential (ultimate market size) M has a mean of 11,643.6 across DVDs. It exhibits huge variation across titles that, to a large extent, tracks the number of units already sold. Indeed, across titles M shows a strong positive correlation with the cumulative number of sales observed from the start of the preorder period up until the 15th week after release ($r = 0.9986$, $p < 0.001$). This provides some support for the validity of our model. Overall, the ratio of mean total observed sales (i.e., cumulative market penetration) from Table 2 to the mean market potential is about 91%. However, despite the strong correlation of total sales with market potential, the variation across titles of the ratio of total observed sales to market potential leads to some interesting managerial implications. We return to this issue in §5.

We find that the mean proportion of consumers who "remember" the release date of the DVD is about 6.2%. Because about 15.7% of total sales occurred during the week of release (see Table 1), this suggests that roughly one-half to one-third of release-week sales are accounted for by this segment of consumers. This information may be relevant to home video distributors as they make prerelease advertising decisions. Finally, the mean rate of visiting the DVD website is about 0.165, which suggests that consumers' median interarrival time (except for the segment of consumers who remember the release date) is about four weeks.

5. Managerial Implications

Three managerial issues, which are of particular interest to the retailer that provided our data, are (i) whether/how the length of the preorder period affects preorder/sales patterns; (ii) estimation of the residual sales for each DVD after the 15th week; and (iii) forecasts of post-release sales based only on preorders. In §§5.1–5.3 below we discuss how implications from our model can be used to address these issues. We should add that such implications are only a first step toward generating managerial insights from our model. As we work more closely with the Internet retailer and obtain additional data, more implications may be feasible. We discuss this again in §6.

5.1. Policy Experiment: Preorder Timing

Because l (the calendar time when preorder is first available) is a retailer's strategic variable, our model allows retailers to adjust l, using a policy experiment, to examine how preorder/sales patterns will be affected.

We study the effect of changing the length of the preorder period by differentiating the sales pattern $f(t)$ (Equation (18)) with respect to l. After algebraic manipulations, we have

$$\frac{\partial f(t)}{\partial l} = \begin{cases} \lambda(\lambda + 2\tau)e^{-\lambda t - \tau k + (\lambda + \tau)l} > 0 & t \geq k, \\ \lambda^2 e^{-\lambda t - \tau k + (\lambda + \tau)l} > 0 & l \leq t < k. \end{cases} \quad (27)$$

Equation (27) indicates that the partial derivative $\partial f(t)/\partial l$ is always positive in both the region $t \geq k$ and the region $l \leq t < k$. Thus, for a fixed k (release date), as l increases (i.e., the preorder period is shortened), the region $l \leq t < k$ will become smaller, while preorders and sales of the other periods will increase, with the magnitude of the predicted changes governed jointly by the other model parameters λ, τ, and the window k.

As a concrete numerical example, we take the title *24 Hour Party People* (considered in §1) to illustrate our approach. Originally, the title was available for preorder 12 weeks before its release date. We analyze the preorder/sales patterns of the DVD under two different scenarios: (i) if preorder is allowed only 3 weeks before the release date, and (ii) if the preorder period is extended to 24 weeks before the release date. The patterns under the two scenarios are shown in Figure 12. The solid line is the preorder/sales pattern under $l = 12$ (the original value); the dashed line shows the predicted preorder/sales pattern under $l = 3$, whereas the dotted line shows the predicted pattern under $l = 24$.

As can be seen, adjusting the number of preorder weeks has only a modest effect on the preorder/sales pattern. Shortening the number of preorder weeks to three weeks pushes some of the sales to the other preorder post-release periods; as a result, the dashed

Figure 12 Policy Experiment: Preorder/Sales Pattern with Different Number of Preorder Weeks
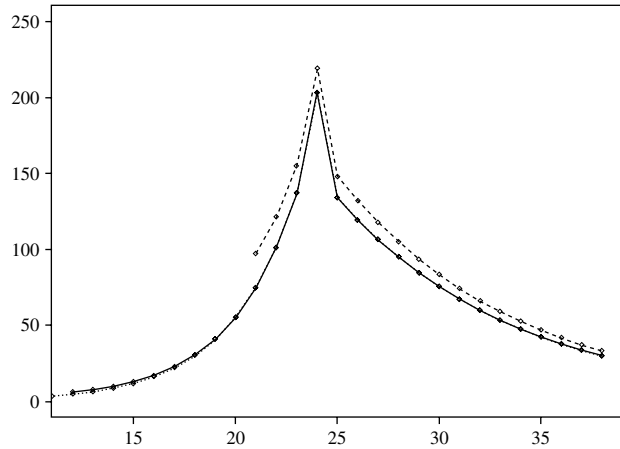
Note. Solid line: 12 weeks of preorder period; dashed line: 3 weeks of preorder period; dotted line: 24 weeks of preorder period.

Figure 13 Histogram of the Proportion of Market Potential Realized by the End of the 15th Week Since Release Across the 223 Titles in Our Data Set


line is always above the solid line, in line with the analysis provided by Equation (27). On the other hand, doubling the number of preorder weeks from 12 to 24 has a negligible effect on the preorder/sales pattern (as shown by the dotted line). This is because very few consumers are willing to preorder a DVD more than three months before the release date, as is suggested by our model estimates.

5.2. Estimating Residual Sales
Because our model estimates the overall market potential of each DVD, retailers may also use our model to estimate the amount of residual sales for each DVD beyond a certain date. By comparing market potential with total sales reached by the end of the 15th week (the horizon that our retailer is interested in), managers can determine which DVDs have exhausted their market potential, and which ones still have "legs" to continue. This may have important implications for inventory management and/or the allocation of promotion efforts.

The quantity of interest to retailers, therefore, is the ratio $\pi_i = (\sum_t y_{it})/M_i$, i.e., the market penetration that has already been realized by the end of the data collection window (i.e., by the 15th week). A histogram of π_i is shown in Figure 13. The mean percentage penetration by week 15 is about 91%, with a fair amount of variation among the different titles. The 10 titles with highest π_i are shown in the top panel of Table 4; among these, more than 99.5% of the total market potential has already been realized. In contrast, the 10 titles with lowest π_i are shown in the bottom panel of Table 4; the percentage penetration among these titles varies from 50% to 75%, indicating that substantial residual sales potential is still available for these titles.
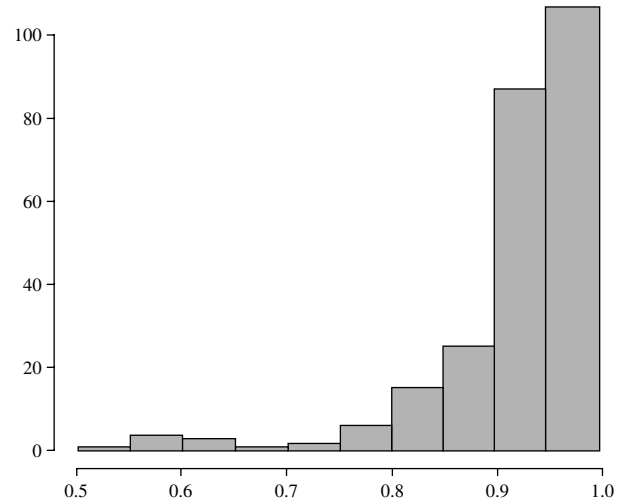
5.3. Forecasting Post-Release Sales from Preorders
To demonstrate the out-of-sample forecasting capability of our model, we consider predictive performance on a holdout sample. More precisely, by randomly dividing our data into a training set of 201 DVDs and a holdout sample of 50 DVDs, we consider how well a model, calibrated on the 201 DVDs, can predict post-release sales of the 50 DVDs using only preorder information (up to one week prior to release date); this is the same forecasting problem considered by Moe and Fader (2002). In the discussion below, we compare our DVD-II model, which demonstrated the best in-sample fit, against the state-of-the-art model, i.e., the model developed by Moe and Fader (2002).

In order to use preorder information to forecast post-release sales, we treat the parameters for each DVD as realizations from a hyper-distribution, in effect embedding our model within a hierarchical model; this is also the approach taken by Moe and Fader (2002). Specifically, we assume that the (transformed) parameters for each DVD are drawn from a common multivariate normal distribution with mean $\vec{\mu}$ and covariance matrix Σ, as follows:

$$(\log(M_i), \text{logit}(\phi_i), \log(\lambda_i), \log(\tau_i))' \sim MVN(\vec{\mu}, \Sigma), \quad (28)$$

where log- and logit-transformations are taken to ensure that the transformed parameters can take on any real values. The parameter correlation structure implicit in the covariance matrix Σ allows us to use a new DVD's preorder information to infer the posterior distribution of all its parameters.

We use a similar specification for the Moe and Fader (2002) model. In particular, the (transformed) parameters in Equation (24) are assumed to be

Table 4 Top Panel: 10 Titles with Highest % of Market Potential Realized by the 15th Week (Since the Week of Release); Bottom Panel: 10 Titles with Lowest % of Market Potential Realized by the 15th Week (Since the Week of Release)

Title	Market potential	% reached by 15th week after release
Ten titles with highest cumulative % penetration by the 15th week after release		
Seabiscuit	3,514.7	99.9
Rugrats Go Wild	1,420.3	99.9
The Powerpuff Girls—The Movie	2,633.4	99.9
Spirit—Stallion of the Cimarron	17,019.0	99.8
Ice Age	53,348.3	99.8
Dumb and Dumberer: When Harry Met Lloyd	1,278.0	99.8
Like Mike	3,711.1	99.8
The Crocodile Hunter—Collision Course	2,019.1	99.8
The Country Bears	1,160.0	99.7
Pirates of the Caribbean—The Curse of the Black Pearl	222,045.3	99.6
Ten titles with lowest cumulative % penetration by the 15th week after release		
Swimming Pool	23,593.8	72.1
Anger Management	11,359.3	67.4
The Italian Job	28,748.1	63.4
Daddy Day Care	6,686.1	61.4
The Big Lebowski	5,964.3	61.2
From Justin to Kelly	1,573.2	59.6
Shaolin Soccer	3,814.8	59.6
The Count of Monte Cristo	16,783.9	57.0
40 Days and 40 Nights	5,161.7	55.6
The Trip	8,262.9	52.6

drawn from a common multivariate normal distribution with mean $\vec{\mu}_{MF}$ and covariance matrix Σ_{MF}, the same hyperprior specification used in Moe and Fader (2002):

$$\big(\log(M_i), \text{logit}(\phi_i), \log(\lambda_{1i}), \log(c_{1i}),$$
$$\log(\lambda_{2i}), \log(c_{2i})\big)' \sim MVN(\vec{\mu}_{MF}, \Sigma_{MF}). \quad (29)$$

We proceed as follows for both of the above procedures. Using an MCMC procedure, we simulate from the posterior distribution of the parameters for any DVD in the holdout sample, conditional only on its preorders up to one week before the release date. We then compute the posterior mean of its parameters, and use these parameter estimates to forecast the remaining sales for that DVD. These estimates are then compared against the actual sales in the holdout sample to compute MAPE and MRAE metrics.

The MAPE for the DVD-II model is 39.1%, and MRAE is 35.1%, whereas the MAPE for Moe and Fader's (2002) model is 72.3% and MRAE is 49.9%. Thus, the predictive performance of our DVD-II model is substantially better than the model of Moe and Fader's (2002). This improvement, which again

may be due to the behavioral premise underlying the DVD-II model, may have important managerial implications for improving inventory control and promotion decisions.

6. Conclusion and Future Research

In this paper, we developed a behaviorally motivated model of aggregate DVD preorder/sales patterns. We modeled the purchase timing decision of an individual consumer using an optimal stopping framework that explicitly captures her forward-looking behavior. We allowed consumer utilities to be heterogeneous, and derived the aggregate preorder/sales curve (DVD-I model). We then extended our model to handle a segment of consumers who remember the release date, resulting in the DVD-II model. We calibrated both models on data and showed that they outperformed state-of-the-art benchmark models, such as the Moe and Fader (2002) model and the Weibull-Gamma model. Finally, we demonstrated how our model can generate managerial insights through policy experiment, estimation of residual sales, and forecasting post-release sales based only on preorders.

To the best of our knowledge, the model proposed here is the first attempt to explain the temporal preorder/sales pattern of DVDs using an individual-level modeling framework. While our model is a first step toward understanding the home video market, it can be further extended in a number of different directions. We briefly note some of these possibilities below.

(i) *Incorporating nonstationarity*: We assumed that the distribution of consumer visits to the DVD website is stationary, and that the market potential is fixed. In reality, these can change over time due to advertising activities, seasonality (Einav 2007, Radas and Shugan 1998) and other reasons. To allow for such nonstationarities, one can specify our parameters as a function of time that depends, for example, on advertising intensity; e.g., $\lambda(t) = \lambda_0 + \beta A(t)$, where $A(t)$ denotes the intensity of advertising over time. Similarly, the market potential M can also be modeled as a function of advertising intensity. Further, seasonality effects can be handled using the methodology developed in Radas and Shugan (1998).

(ii) *Dynamic pricing*: We assume that the price of a DVD is constant over time. Although this assumption may be reasonable for movie DVDs (as verified empirically with our data provider), it may not hold for other categories, for example, books, video games, music CDs, where promotions tend to be offered early after their release. Researchers may extend our optimal stopping framework to handle dynamic pricing, by allowing consumers to act strategically based

on their expectation of future price changes (e.g., Nair 2007, Sun 2006).

(iii) *Movie characteristics and cross-category analysis*: It may be interesting to investigate how characteristics of a movie such as genre, story (Eliashberg et al. 2007), star involvement (e.g., Wei 2006) are related to the sales pattern of its DVD. In addition, preordering has become prevalent in many different categories (e.g., books, video games, music CDs). Variants of our model can be applied to each category to study how category characteristics affect model parameters. In other categories, the preorder/sales pattern may be different from the pattern in Figure 1; a concave-up pattern may also be possible, in which case the model extension discussed in Appendix V may be useful.

(iv) *Joint modeling of box office and DVD revenue*: Our framework can be potentially extended to an integrated model of a consumer's decision of when/whether to watch a movie in a theatre and whether/when to purchase the corresponding DVD. This extension allows one to understand not only the cannibalization/synergy effects between movie and DVD, but also the role of the "window," in a more structural manner. Currently, our policy experiment only addresses changes in sales when l (the length of the preorder period) is varied; with the above extension, one can then conduct policy experiments by varying k (the window) as well. The optimal window can then be derived, a topic of interest to many other marketing researchers (e.g., Eliashberg et al. 2006, Lehmann and Weinberg 2000, Prasad et al. 2004). Already, some progress has been made in this area (e.g., Luan 2005); we believe that more managerial insights can be gained with more research attention.

Appendix

I. Proof of Monotonicity of Our Model Setup

A monotone optimal stopping problem is defined as follows (Ferguson 2000):

Let A_n be the event $\{Y_n \geq E(Y_{n+1} \mid \Gamma_n)\}$. A problem is *monotone* if $A_0 \subset A_1 \subset A_2 \cdots$ almost surely. From Equations (9) and (10), we show that $Y_n \geq E(Y_{n+1} \mid \Gamma_n) \Leftrightarrow \Delta_n \leq C$ for some constant C. Consider Y_{n+1}. If $t_{n+1} > k$, then the consumer will buy at time t_{n+1} based on Equation (7). If $t_{n+1} \leq k$, then we have

$$\Delta_{n+1} = k - t_{n+1} \leq k - t_n \leq C \quad \Leftrightarrow \quad Y_{n+1} \geq E(Y_{n+2} \mid \Gamma_{n+1}).$$

Thus, the condition of monotonicity is satisfied.

II. Derivation of the Threshold Rule for General Interarrival Time

$$E(Y_{n+1} \mid \Gamma_n) = e^{-\gamma t_n} \left[\int_0^{\Delta_n} (e^{-\gamma \Delta_n} u - e^{-\gamma w_n} x) g(w_n) \, dw_n \right.$$

$$\left. + (u - x) \int_{\Delta_n}^{\infty} e^{-\gamma w_n} g(w_n) \, dw_n \right]$$

$$= e^{-\gamma t_n} \left[e^{-\gamma \Delta_n} u G(\Delta_n) - x \int_0^{\Delta_n} e^{-\gamma w_n} g(w_n) \, dw_n \right.$$

$$\left. + (u - x) \int_{\Delta_n}^{\infty} e^{-\gamma w_n} g(w_n) \, dw_n \right]$$

$$= e^{-\gamma t_n} \left[e^{-\gamma \Delta_n} u G(\Delta_n) \right.$$

$$\left. + u \int_{\Delta_n}^{\infty} e^{-\gamma w_n} g(w_n) \, dw_n - x E(e^{-\gamma w_n}) \right].$$

Thus,

$$Y_n \geq E(Y_{n+1} \mid \Gamma_n)$$

$$\Leftrightarrow \quad -x + e^{-\gamma \Delta_n} u \geq e^{-\gamma \Delta_n} u G(\Delta_n)$$

$$+ u \int_{\Delta_n}^{\infty} e^{-\gamma w_n} f(w_n) \, dw_n - x E(e^{-\gamma w_n})$$

$$\Leftrightarrow \quad e^{-\gamma \Delta_n} [1 - G(\Delta_n)] - \int_{\Delta_n}^{\infty} e^{-\gamma w_n} g(w_n) \, dw_n \geq \frac{x}{u} [1 - E(e^{-\gamma w_n})].$$

III. Derivation of the Threshold Rule for Exponentially Distributed Interarrival Time

Start from Equation (10),

$$\frac{x}{u} [1 - E(e^{-\gamma w_n})] \leq e^{-\gamma \Delta_n} [1 - G(\Delta_n)] - \int_{\Delta_n}^{\infty} e^{-\gamma w_n} g(w_n) \, dw_n.$$

Because $w_n \sim \exp(\lambda)$, we have $g(w_n) = \lambda e^{-\lambda w_n}$ and $G(\Delta_n) = 1 - e^{-\lambda w_n}$.

$$E(e^{-\gamma w_n}) = \int_0^{\infty} e^{-\gamma w_n} \lambda e^{-\lambda w_n} \, dw_n$$

$$= \lambda \int_0^{\infty} e^{-(\lambda + \gamma) w_n} \, dw_n = \frac{\lambda}{\lambda + \gamma},$$

$$\int_{\Delta_n}^{\infty} e^{-\gamma w_n} g(w_n) \, dw_n = \int_{\Delta_n}^{\infty} e^{-\gamma w_n} \lambda e^{-\lambda w_n} \, dw_n$$

$$= \frac{\lambda}{\lambda + \gamma} e^{-(\lambda + \gamma) \Delta_n}.$$

Thus, we have

$$\frac{x}{u} \left[1 - \frac{\lambda}{\lambda + \gamma} \right] \leq e^{-\gamma \Delta_n} \left[1 - (1 - e^{-\lambda \Delta_n}) \right] - \frac{\lambda}{\lambda + \gamma} e^{-(\lambda + \gamma) \Delta_n}$$

$$\Leftrightarrow \quad \frac{x}{u} \left[\frac{\gamma}{\lambda + \gamma} \right] \leq e^{-(\lambda + \gamma) \Delta_n} - \frac{\lambda}{\lambda + \gamma} e^{-(\lambda + \gamma) \Delta_n}$$

$$\Leftrightarrow \quad \frac{x}{u} \leq e^{-(\lambda + \gamma) \Delta_n}$$

$$\Leftrightarrow \quad \Delta_n \leq \frac{1}{\lambda + \gamma} \ln \left(\frac{u}{x} \right).$$

IV. Distribution of Purchase Timing with Consumer Heterogeneity

First, consider the case $t \geq k$. Clearly, it is also true that $t > d$ and hence $1_{\{t > d\}} = 1$. Thus,

$$f(t \mid t \geq k) = \lambda \tau e^{-\lambda t} \int_0^{\infty} e^{\lambda d - \tau C_i} \, dC_i$$

$$= \lambda \tau e^{-\lambda t} \left[\int_0^{k-l} e^{\lambda (k - C_i) - \tau C_i} \, dC_i + \int_{k-l}^{\infty} e^{\lambda l - \tau C_i} \, dC_i \right]$$

$$= \lambda \tau e^{-\lambda t} \left[e^{\lambda k} \int_0^{k-l} e^{-(\lambda + \tau) C_i} \, dC_i + e^{\lambda l} \int_{k-l}^{\infty} e^{-\tau C_i} \, dC_i \right]$$

$$= \lambda\tau e^{-\lambda t}\left[e^{\lambda k}\left(\frac{e^{-(\lambda+k)(k-l)}}{-(\lambda+\tau)} + \frac{1}{\lambda+\tau}\right) + e^{\lambda l}\left(\frac{e^{-\tau(k-l)}}{\tau}\right)\right]$$

$$= \left[\frac{\lambda\tau}{\lambda+\tau}e^{\lambda k}(1 - e^{-(\lambda+\tau)(k-l)}) + \lambda e^{\lambda l - \tau(k-l)}\right]e^{-\lambda t}.$$

Second, for the case $l \leq t < k$,

$$f(t \mid l \leq t < k)$$

$$= \lambda\tau e^{-\lambda t}\left[\int_{k-t}^{k-l} e^{\lambda(k-C_i)-\tau C_i}\, dC_i + \int_{k-l}^{\infty} e^{\lambda l - \tau C_i}\, dC_i\right]$$

$$= \lambda\tau e^{-\lambda t}\left[e^{\lambda k}\int_{k-t}^{k-l} e^{-(\lambda+\tau)C_i}\, dC_i + e^{\lambda l}\int_{k-l}^{\infty} e^{-\tau C_i}\, dC_i\right]$$

$$= \lambda\tau e^{-\lambda t}\left[\frac{1}{\lambda+\tau}e^{\lambda k}(e^{-(\lambda+\tau)(k-t)} - e^{-(\lambda+\tau)(k-l)}) + \frac{1}{\tau}e^{\lambda l}(e^{-\tau(k-l)})\right]$$

$$= \frac{\lambda\tau}{\lambda+\tau}[e^{-\tau k + \tau t} - e^{-\lambda t + \lambda l - \tau k + \tau l}] + \lambda e^{-\lambda t + \lambda l - \tau k + \tau l}$$

$$= \frac{\lambda\tau}{\lambda+\tau}e^{-\tau(k-t)} + \frac{\lambda^2}{\lambda+\tau}e^{-\lambda(t-l)-\tau(k-l)}.$$

References

Allaart, P. 2004. Optimal stopping rules for correlated random walks with a discount. *J. Appl. Probab.* **41**(2) 483–496.

Bakos, J. Y. 1997. Reducing buyer search costs: Implications for electronic marketplaces. *Management Sci.* **43**(12) 1676–1692.

Chintagunta, P., T. Erdem, P. Rossi, M. Wedel. 2006. Structural modeling in marketing: Review and assessment. *Marketing Sci.* **25** 604–616.

Chow, Y. S., H. Robbins, D. Siegmund. 1971. *Great Expectations: The Theory of Optimal Stopping*. Houghton Mifflin Company, Boston.

Clementi, F., M. Gallegati. 2005. Pareto's law of income distribution: Evidence for Germany, the United Kingdom, and the United States. A. Chaterjee, S. Yarlagadda, B. K. Chakrabarti, eds. *Econophysics of Wealth Distributions*. Springer, Milan, 3–14.

Dubins, L. E., H. Teicher. 1967. Optimal stopping when the future is discounted. *Ann. Math. Statist.* **38**(2) 601–605.

Einav, L. 2007. Seasonality in U.S. motion picture industry. *RAND J. Econom.* **38**(1) 128–146.

Eliashberg, J., A. Elberse, M. A. A. M. Leenders. 2006. The motion picture industry: Critical issues in practice, current research, and new research directions. *Marketing Sci.* **25**(6) 638–661.

Eliashberg, J., S. Hui, J. Zhang. 2007. From story line to box office: A new approach for green-lighting movie scripts. *Management Sci.* **53**(6) 881–893.

Ferguson, T. S. 2000. Optimal stopping and applications. http://www.math.ucla.edu/~tom/Stopping/Contents.html.

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

Hauser, J., G. J. Tellis, A. Griffin. 2006. Research on innovation: A review and agenda for marketing science. *Marketing Sci.* **25**(6) 687–717.

Jaggia, S., S. Thosar. 1995. Contested tender offers: An estimate of the hazard function. *J. Bus. Econom. Statist.* **13**(1) 113–119.

Jedidi, K., R. E. Krider, C. B. Weinberg. 1998. Clustering at the movies. *Marketing Lett.* **9**(4) 393–405.

Jones, J. M., C. J. Ritz. 1991. Incorporating distribution into new product diffusion models. *Internat. J. Res. Marketing* **8** 91–112.

Lehmann, D. R., C. B. Weinberg. 2000. Sales through sequential distribution channels: An application to movies and videos. *J. Marketing* **64**(July) 18–33.

Lowenstein, G. F. 1988. Frames of mind in intertemporal choice. *Management Sci.* **34**(2) 200–214.

Lowenstein, G., D. Prelec. 1992. Anomalies in intertemporal choice: Evidence and an interpretation. *Quart. J. Econom.* **107**(2) 573–597.

Luan, J. Y. 2005. Optimal inter-release timing for sequential releases. Working paper, Yale University, New Haven, CT.

Mas-Colell, A., M. D. Whinston, J. R. Green. 1995. *Microeconomic Theory*. Oxford University Press, New York.

Moe, W. W., P. S. Fader. 2002. Using advance purchase orders to forecast new product sales. *Marketing Sci.* **21**(3) 347–364.

Moe, W. W., P. S. Fader. 2004. Dynamic conversion behavior at e-commerce sites. *Management Sci.* **50**(3) 326–335.

Morrison, D. G., D. C. Schmittlein. 1980. Jobs, strikes and wars: Probability models for duration. *Organ. Behav. Human Performance* **25** 224–251.

Morrison, D. G., D. C. Schmittlein. 1988. Generalizing the NBD model for customer purchase: What are the implications and is it worth the effort? (with discussion and rejoinder). *J. Bus. Econom. Statist.* **6**(2) 145–166.

Motion Picture Association of American (MPAA). 2006. *U.S. Entertainment Industry: Market Statistics*. http://www.mpaa.org.

Muthen, B., K. Masyn. 2005. Discrete-time survival mixture analysis. *J. Educ. Behavioral Statist.* **30**(1) 27–58.

Nair, H. 2007. Intertemporal price discrimination with forward-looking consumers. Application to the U.S. market console for video games. *Quart. Marketing Econom.* **5**(3) 239–292.

Park, Y.-H., P. S. Fader. 2004. Modeling browsing behavior at multiple web sites. *Marketing Sci.* **23**(3) 280–303.

Prasad, A., V. Mahajan, B. J. Bronnenberg. 2004. Product entry timing in dual distribution channels: The case of the movie industry. *Rev. Marketing Sci.* **2**, Article 4.

Radas, S., S. M. Shugan. 1998. Seasonal marketing and timing new product introductions. *J. Marketing Res.* **35**(3) 296–315.

Sawhney, M. S., J. Eliashberg. 1996. A parsimonious model for forecasting gross box-office revenues of motion pictures. *Marketing Sci.* **15**(2) 113–131.

Schmittlein, D., V. Mahajan. 1982. Maximum likelihood estimation for an innovation diffusion model of new product acceptance. *Marketing Sci.* **1**(1) 57–78.

Schwarz, G. 1978. Estimating the dimension of a model. *Ann. Statist.* **6**(2) 461–464.

Song, I., P. K. Chintagunta. 2003. A micromodel of new product adoption with heterogeneous and forward-looking consumers: Application to the digital camera category. *Quant. Marketing Econom.* **1**(4) 371–407.

Standard and Poor's. 2006. *Industry Surveys: Movies and Home Entertainment*. McGraw-Hill, New York.

Sun, B. 2006. Technology innovation and implications for customer relationship management. *Marketing Sci.* **25**(6) 594–597.

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

Wei, L. 2006. Making sense of these million-dollar babies—Rationale behind superstar profit participation contracts. *Marketing Sci.* **25**(6) 678–680.

Wold, H., P. Whittle. 1957. A model explaining the Pareto distribution of wealth. *Econometrica* **25** 591–595.

Wu, D., G. Ray, X. Geng, A. Whinston. 2004. Implications for reduced search cost and free riding in e-commerce. *Marketing Sci.* **23**(2) 255–262.

Zufryden, F. S. 1996. Linking advertising to box office performance of new film releases: A marketing planning model. *J. Advertising Res.* **36**(4) 29–41.