Fast-Track Article

Using Advance Purchase Orders to Forecast New Product Sales

Wendy W. Moe • Peter S. Fader
University of Texas-Austin, 2100 Speedway, CBA 7.216, Austin, TX 78712
University of Pennsylvania-Wharton, Jon M. Huntsman Hall, Suite 700, Philadelphia, PA 19104
wendy.moe@bus.utexas.edu • faderp@wharton.upenn.edu

Marketers have long struggled with developing forecasts for new products before their launch. We focus on one data source—advance purchase orders—that has been available to retailers for many years but has rarely been tied together with postlaunch sales data. We put forth a duration model that incorporates the basic concepts of new product diffusion, using a mixture of two distributions: one representing the behavior of innovators (i.e., those who place advance orders) and one representing the behavior of followers (i.e., those who wait for the mass market to emerge). The resulting mixed-Weibull model specification can accommodate a wide variety of possible sales patterns. This flexibility is what makes the model well-suited for an experiential product category (e.g., movies, music, etc.) in which we frequently observe very different sales diffusion patterns, ranging from a rapid exponential decline (which is most typical) to a gradual buildup characteristic of "sleeper" products. We incorporate product-specific covariates and use hierarchical Bayes methods to link the two customer segments together while accommodating heterogeneity across products. We find that this model fits a variety of sales patterns far better than do a pair of benchmark models. More importantly, we demonstrate the ability to forecast new album sales before the actual launch of the album, based only on the pattern of advance orders.
(*Advance Selling; Diffusion; Forecasting; Entertainment Marketing; Hierarchical Bayes Analysis; Stochastic Models*)

Introduction

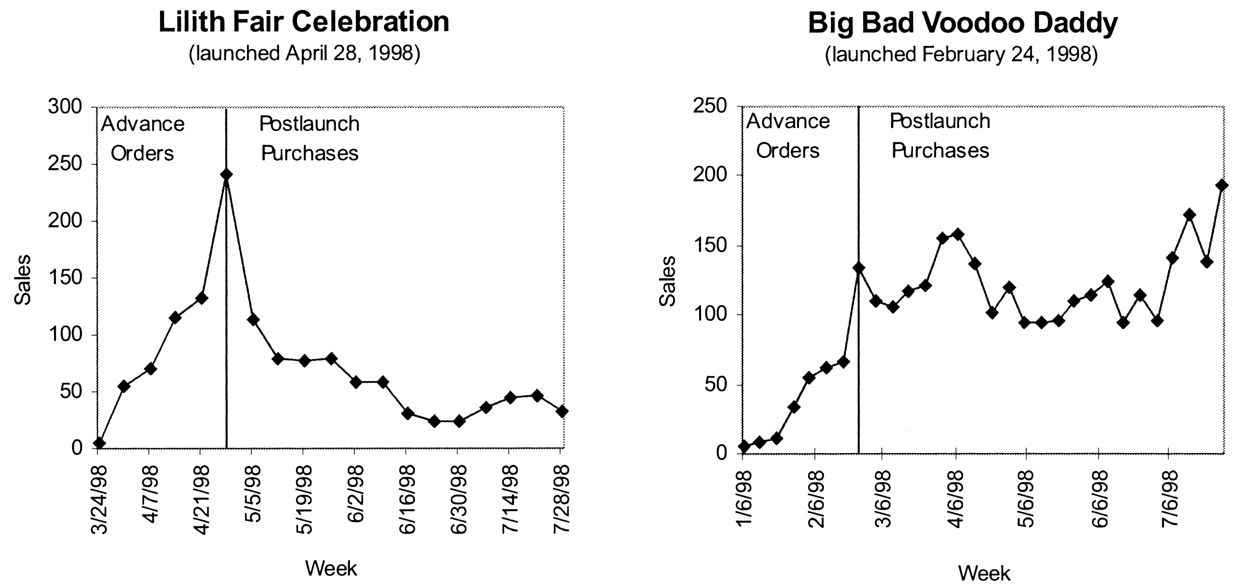
One of the key challenges in managing the launch of a new product is the attempt to obtain valid and reliable indicators about the product's likely future sales levels before the product is actually launched. A novel approach to address this issue arises in the form of *advance purchase orders.* For decades, retailers in many different industries have accepted advance orders from customers before the launch of a new product, but until recently it had been virtually impossible to track these orders systematically and link them directly with postlaunch sales data.

This has changed with the advent of the Internet and the associated progress in information technology. Online retailers such as Amazon.com and CDNOW routinely promote the ability to buy an item several weeks before it is actually available, and their internal data-warehousing systems provide

0732-2399/02/2103/347/$05.00
1526-548X electronic ISSN

MARKETING SCIENCE © 2002 INFORMS
Vol. 21, No. 3, Summer 2002, pp. 347–364

Figure 1 Sales Diffusion Patterns



a seamless link between their pre- and postlaunch sales records. Even retailers with a predominant offline presence are gaining similar capabilities, thanks to advances in data integration technologies (Swift 2001). As a result, advance orders may act as an early indicator of new product performance in many retail settings. In this paper, we develop a prelaunch forecasting model that uses the pattern of advance orders to predict future product sales and overall market potential.

Researchers have recognized the importance of advance orders to address managerial issues such as pricing and capacity constraints. For instance, Xie and Shugan (2001) present a normative model involving prelaunch announcements to identify conditions under which to allow advance orders, as well as determining optimal pricing schemes for these early sales. Their model and others (e.g., Desiraju and Shugan 1999, Shugan and Xie 2000) have uncovered a number of relevant managerial insights but have generally relied upon relatively simple assumptions about customer behavior and heterogeneity. None has been positioned as an empirical forecasting exercise.

With this background in mind, we present a new modeling approach to project the actual purchasing levels for a large set of new products (66 albums that were preannounced and sold at the CDNOW website), based only on the patterns of their prelaunch sales. Similar to the basic conceptual story associated with the classic diffusion model (Bass 1969), we assume that a new product's sales arise from two consumer groups, each with different adoption tendencies: an "innovator" group that tends to enter the market very early (and accounts for all of the advance-order sales) and a mass-market "follower" group that waits until the release of the product before making any purchase decisions. We capture this process using a mixed-Weibull distribution estimated via hierarchical Bayes methods, where the parameters governing the diffusion process for one group of consumers are correlated with those for the other group. Thus, the observed advance-order data (from the innovators) provides useful information about the nature (and number) of the followers. Our model will allow managers to estimate the overall potential market size for a given album based solely on the pattern of advance orders. The Bayesian methodology also helps control for differences across albums and, therefore, reveals a rich description

Table 1 Product Descriptions

Artist	Title	Street Date	Prelaunch Weeks	Postlaunch Weeks	Total Sales
Dave Matthews Band	Before These Crowded Streets	4/28/1998	4	14	6336
Madonna	Ray of Light	3/3/1998	5	22	6219
Sarah McLachlan	Surfacing	7/15/1997	5	55	5588
Eric Clapton	Pilgrim	3/10/1998	4	21	5269
Elton John	Something About the Way You . . .	9/23/1997	2	45	4601
Celine Dion	Let's Talk About Love	11/18/1997	4	37	4436
Bob Dylan	Time Out of Mind	9/30/1997	5	44	3938
Natalie Imbruglia	Left of the Middle	3/10/1998	4	21	3728
Beastie Boys	Hello Nasty	7/14/1998	5	3	3698
Pearl Jam	Yield	2/3/1998	5	26	3568
Garbage	Version 2.0	5/12/1998	5	12	3345
Tori Amos	From the Choirgirl Hotel	5/5/1998	5	13	3305
Natalie Merchant	Ophelia	5/19/1998	5	11	3254
Smashing Pumpkins	Adore	6/2/1998	5	9	3252
Enya	Paint the Sky With Stars-Best	11/11/1997	5	38	3135
Big Bad Voodoo Daddy	Big Bad Voodoo Daddy	2/24/1998	7	23	3087
Dave Matthews Band	Live at Red Rocks 8/15/95	10/28/1997	4	40	2732
Garth Brooks	Limited Series Box Set	5/5/1998	4	13	2625
Paul Simon	Songs From the Capeman	11/18/1997	5	37	2495
Metallica	Re-Load	11/18/1997	5	37	2378
Garth Brooks	Sevens	11/25/1997	3	36	2349
Bonnie Raitt	Fundamental	4/7/1998	5	17	2223
Ally McBeal	TV Soundtrack	5/5/1998	4	13	2196
Barenaked Ladies	Stunt	7/7/1998	5	4	2002
Van Halen	Van Halen 3	3/17/1998	5	20	1917
Janet Jackson	Velvet Rope	10/7/1997	5	43	1870
Page & Plant	Walking Into Clarksdale	4/21/1998	4	15	1793
Lucinda Williams	Car Wheels on a Gravel Road	6/30/1998	5	5	1727
Diana, Princess of Wales	Diana, Princess of Wales Tribute	12/2/1997	4	35	1679
Jimmy Buffett	Don't Stop the Carnival	4/28/1998	5	14	1598
Brian Setzer Orchestra	Dirty Boogie	6/23/1998	5	6	1549
Barbra Streisand	Higher Ground	11/11/1997	5	38	1540
Ani DiFranco	Little Plastic Castle	2/17/1998	5	24	1537
Massive Attack	Mezzanine	5/12/1998	5	12	1436
U2	Pop	3/4/1997	4	74	1407
Van Morrison	Philosopher's Stone	6/16/1998	5	7	1400
John Fogerty	Premonition	6/9/1998	5	8	1369
Propellerheads	Decksandrumsandrockandroll	3/24/1998	4	19	1347
Godzilla	Soundtrack	5/19/1998	4	11	1346
Lilith Fair	A Celebration of Women	4/28/1998	5	14	1330
Led Zeppelin	BBC Sessions	11/18/1997	6	37	1282
Radiohead	Airbag/How Am I Driving?	4/21/1998	5	15	1259
Bragg, Billy & Wilco	Mermaid Avenue	6/23/1998	5	6	1187
X-Files	Movie Soundtrack	6/2/1998	4	9	1179
Mavericks	Trampoline	3/10/1998	4	21	1038
Liquid Tension Experiment	Liquid Tension Experiment	3/10/1998	5	21	1009
Nanci Griffith	Other Voices Too (A Trip Back . . .)	7/21/1998	5	2	998
Frank Sinatra	In the Wee Small Hours	5/26/1998	5	10	938

Table 1 Continued

Artist	Title	Street Date	Prelaunch Weeks	Postlaunch Weeks	Total Sales
Genesis	Vol. 1—Genesis Archives 1967–75	6/16/1998	3	7	910
Jeff Buckley	Sketches for My Sweetheart . . .	5/26/1998	5	10	904
Lenny Kravitz	5	5/12/1998	5	12	846
Yes	Keys to Ascension 2	11/11/1997	8	38	831
Jimi Hendrix	BBC Sessions	6/2/1998	4	9	798
Leann Rimes	Sittin' on Top of the World	5/5/1998	5	13	784
Frank Sinatra	Only the Lonely	5/26/1998	4	10	768
Marc Cohn	Burning the Daze	3/17/1998	5	20	767
Counting Crows	Across a Wire—Live in New York	7/14/1998	4	3	744
Maxwell	Embrya	6/30/1998	5	5	709
Brian Wilson	Imagination	6/16/1998	5	7	686
Various Artists	Where Have All the Flowers Gone?	3/17/1998	4	20	651
Tricky	Angels with Dirty Faces	6/2/1998	4	9	608
Xena: Warrior Princess	Bitter Suite—A Musical Odyssey	3/24/1998	4	19	586
Dwight Yoakam	Long Way Home	6/9/1998	5	8	578
Rod Stewart	When We Were the New Boys	6/2/1998	5	9	563
David Lee Roth Band	DLR Band	6/9/1998	3	8	540
Gillian Welch	Hell Among the Yearlings	7/28/1998	4	1	272

about the substantive nature of each of the underlying segments.

In the next section we describe the dataset that motivates our analysis. Next, we derive our model, and then we turn our attention to a discussion of the empirical results. After getting a good sense of the model parameters and their implications for each of the underlying duration model components, we will examine the model's ability to forecast album sales, using only prelaunch data.

Data

We obtained sales data from CDNOW, a leading online retailer of music albums, for 66 new albums. CDNOW accepts advance purchase orders for a number of albums that have yet to be released. As with ordinary albums, the Web pages for each of these prereleases include details about the album and artist, as well as information about the anticipated launch date. Advance purchase orders are fulfilled immediately, once the album becomes available for regular retail sale (these release dates are often tightly controlled by the record labels).

Since virtually all of CDNOW's business takes place online, there is relatively little operational distinction between an advance purchase and a traditional postlaunch sale. CDNOW's transaction servers handle both types of orders in the same manner, and their customer databases make no major distinctions between orders that arise before or after an album's actual launch data.

CDNOW provided us with data for all albums that were first made available for sale at their website between March 4, 1997 and August 4, 1998. We eliminated albums that lacked any advance-order information, as well as those that had extremely sparse postlaunch data (fewer than 100 units sold in their initial week). The length of the prelaunch period varied between two and eight weeks before the release date. Actual sales data were collected through the week ending August 3, 1999. Table 1 (see page 349) provides a brief summary of the 66 albums that we use for our analysis.

Figure 1 (see page 347) provides a graphical depiction of two of these albums. The left panel, representing *Lilith Fair: A Celebration of Women in Music*, a performance by various artists, shows a fairly typical order/sales pattern for many of the albums included in our database. Advance orders steadily ramp up during the five-week period before the album was formally released for sale, and the number of transactions peaks during the launch week (4/28/98). Following launch, sales decline in a (roughly) exponential manner, consistent with the patterns observed frequently for products such as albums (Moe and Fader 2001), movies (Krider and Weinberg 1998, Sawhney and Eliashberg 1996), and other consumer goods.

The right panel of Figure 1 shows a relatively unusual sales pattern, in this case for *Big Bad Voodoo Daddy,* by the band of the same name. The shape of the advance-order pattern appears to be quite similar to that of *Lilith Fair,* although it takes place over a longer period of time (seven versus five weeks) and involves a lower number of orders (243 versus 380 for *Lilith Fair*). After launch, however, the sales pattern becomes dramatically different. There is no evidence of an exponential sales decline; in fact, sales appear to be steadily increasing, even 23 weeks after the album first became available in retail stores. While few albums in our dataset show such steady sales performance over time, there are a number of other ''nonexponential'' exceptions to the relatively common postlaunch sales pattern exhibited by the *Lilith Fair* album.

A key issue to be addressed is whether (and how) the advance-order patterns shown in Figure 1 (and across the 64 other albums) can predict such a wide range of postlaunch sales curves. At first glance, this may seem to be a difficult task, given the seemingly inconsequential differences between the advance-order numbers for these two sample albums. However, as our model leverages a broader set of pre- and postlaunch purchase histories, some distinctive shapes and diagnostics will emerge. The casual glimpses provided by Table 1 and Figure 1 offer some initial insights about the appropriate modeling elements, but we now turn to a more formal development of the complete model.

Model Development

As mentioned earlier, the underlying premise of the model is similar to that of other diffusion models, in that we assume two general ''classes'' of consumers: innovators and followers (Bass 1969, Mahajan et al. 1995, Rogers 1983). Innovators tend to be driven by their innate desire to lead the market while followers are influenced by ''bandwagon'' effects that grow with the increasing number of adopters. In the standard diffusion modeling context, the follower segment emerges in parallel with the innovators and can start purchasing immediately after product launch, although many will wait until a critical mass of buyers has emerged.

Our current modeling situation is quite different. By definition, followers are unlikely to initiate orders for a product that no one yet possesses. Therefore, we assume that the followers will not begin making purchases until after the product is officially released. In a few rare instances there may be some followers jumping on the advance-order bandwagon, but in most cases it is unlikely that many followers will even be aware that the product exists during the prelaunch period.

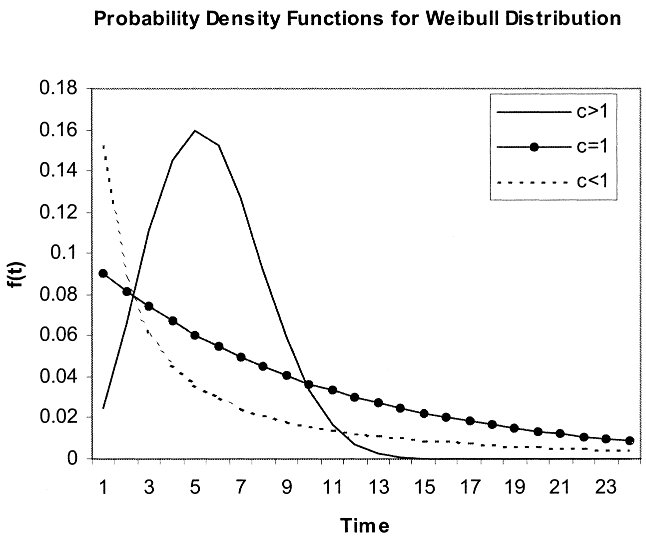
Therefore, during the prerelease period, we do not need to worry about any actual adoption behavior by the followers. However, there may still be some word-of-mouth effects influencing the adoption behavior among the (relatively homogeneous) innovators, which would allow sales to build up rapidly (or possibly fall off) during this period. Such a process is frequently modeled using a Weibull distribution with the following cumulative distribution function:

$$F_1(t) = 1 - e^{-\lambda_1 t^{c_1}}, \tag{1}$$

where λ_1 and c_1 are model parameters (taking on positive values) and t is the time (in weeks) since the album was first made available for advance purchase by the record label and on the CDNOW website.

Note that this Weibull distribution has no upper bound, so we allow for the fact that many innovators may not have ordered the album by the official

Figure 2 **Weibull Distributions**



launch week (denoted as t^*). As a result, the Weibull process continues for this group, even after the product is launched and the follower segment has emerged. In other words, after the album becomes available for actual purchase (at t^*), there is a mix of "late innovators" and followers that comprises the customer base for it.

When the followers begin to make purchases (at t^*), we assume that their behavior is governed by a separate (but correlated) Weibull distribution. However, because the follower segment does not become active until the time of product launch, its Weibull distribution is time-shifted by t^*:

$$F_2(t) = 1 - e^{-\lambda_2(t-t^*)^{c_2}} \quad \textit{for } t \geq t^*. \tag{2}$$

The Weibull distribution provides an extremely flexible way to capture the different behavioral tendencies for each of the two consumer groups. A closer look at the Weibull distribution (see Figure 2), using different values for the shape parameter, c (c_1 for the innovators and c_2 for the followers), provides some intuition about its suitability for our purposes. The case of $c > 1$ reveals a rapid buildup in sales over time. This seems to fit well with innovators' purchase patterns for both albums shown in Figure 1. In contrast, when $c < 1$, sales fall off quickly, which likely describes the followers' behavior for the *Lilith Fair* album. Finally, $c = 1$, with its constant hazard rate, corresponds to the case of the standard exponential distribution, which has often provided an adequate description of postlaunch sales for music albums.

In the postlaunch period, the model assumes that both the innovator and follower groups are active, thereby giving us a combination of two Weibull distributions. This blend of Weibull distributions can yield a wide variety of different sales patterns, including the steady rise seen for *Big Bad Voodoo Daddy* shown earlier. Looking at the two periods (pre- and postlaunch) together, it seems evident that a combination of two Weibull distributions is required to obtain a suitable fit to the data. It is unlikely that any ordinary diffusion model can capture the convex buildup in sales (i.e., $c_1 > 1$), followed immediately by a similarly convex decline (presumably with $c_2 < 1$) after the launch date. (We will use two benchmark models to test this assertion.) As noted earlier, this type of pattern occurs for many albums in our dataset, which is consistent with the model proposed here.

Therefore, to accommodate these unique patterns and to allow for two distinct classes of consumers, we use the *mixed-Weibull* model with cumulative distribution function (cdf):

$$F(t) = \begin{cases} \phi[1 - e^{-\lambda_1 t^{c_1}}] & \text{for t} < \text{t*} \\ \phi[1 - e^{-\lambda_1 t^{c_1}}] & \\ \quad +(1-\phi)[1 - e^{-\lambda_2(t-t^*)^{c_2}}] & \text{for t} \geq \text{t*}, \end{cases} \tag{3}$$

where ϕ ($0 \leq \phi \leq 1$) represents the fraction of buyers associated with the innovator segment and λ_1, c_1, λ_2, $c_2 > 0$. This type of model is used frequently in manufacturing/reliability settings (e.g., Jiang and Murthy 1998) because of its practical blend of parsimony and flexibility.

Estimation

To estimate the mixed-Weibull model, there are four implementation issues (discrete time, right truncation, hierarchical Bayes methods, and product-specific covariates) that we must address to allow us to properly deal with the dataset.

Discrete time refers to the fact that the observed data comes to us in weekly increments, although we assume that the underlying behavior occurs in continuous time. This is a common convention in duration models and simply requires us to model each week's sales as the difference between the cdf from the beginning to the end of each week:

$$F(t) = \begin{cases} \phi[e^{-\lambda_1(t-1)^{c_1}} - e^{-\lambda_1 t^{c_1}}] & \text{for t} < \text{t*} \\ \phi[e^{-\lambda_1(t-1)^{c_1}} - e^{-\lambda_1 t^{c_1}}] & \\ \quad + (1-\phi) & \\ \quad \times [e^{-\lambda_2(t-t^*-1)^{c_2}} - e^{-\lambda_2(t-t^*)^{c_2}}] & \text{for t} \geqslant \text{t*}. \end{cases} \tag{4}$$

Right truncation exists because we have no information about the number of buyers (or nonbuyers) beyond the end of the data period. Ordinarily, we would write the complete likelihood function as:

$$L = \prod_{t=1}^{T} P(t)^{Sales_t} \cdot R(T)^{N-\Sigma_{t=1}^{T} Sales_t}, \tag{5}$$

where T is the total time for which we observe sales, $R(T)$ is the reliability (or survival) function which equals $1 - F(T)$, and N is the total number of possible buyers (e.g., the size of a consumer panel). Instead of attempting to estimate N directly as a part of the model (as is done for the typical Bass model), we plainly acknowledge that our dataset is right truncated, and we rescale all of the model's probabilities accordingly. Thus, in place of $P(t)$ we use

$$P'(t) = \frac{P(t)}{\hat{F}(T)}, \tag{6}$$

where $\hat{F}(T)$ is the estimated fraction of eventual customers who have bought the album by the end of the observation period. Effectively, we treat the observed sales as arising from a multinomial distribution, where the multinomial probabilities are specified by Equation (7):

$$Sales_t \sim Multinomial\left(P'(1), P'(2), P'(3), \ldots, P'(T); \sum_{t=1}^{T} Sales_t\right). \tag{7}$$

This rescaling allows us to take full advantage of the *pattern* seen in week-to-week sales, rather than focusing on the overall level of sales observed. This procedure is a natural way to deal with right-truncated data, and it allows us to estimate the total market size from a limited amount of observed sales data. Using this rescaled probability function, we can rewrite the likelihood function and calculate a forecast of the total market size as follows:

$$L = \prod_{t=1}^{T} P'(t)^{Sales_t} \quad \text{and} \quad \hat{N} = \frac{\sum_{t=1}^{T} Sales_t}{\hat{F}(T)}. \tag{8}$$

Once we have obtained the parameter estimates, we can then impute the overall (potential) sales volume by recognizing that our best estimate for N, $\hat{N}$, is the total number of observed sales through time T, divided by the estimated fraction of an album's eventual sales that have occurred by time T. This is how we will generate market-size forecasts from the prelaunch data.

We use *hierarchical Bayes* methods to estimate our prelaunch model across a variety of new product introductions. In previous research, Lenk and Rao (1990) used a similar approach to estimate the Bass model across several new product categories. Neelamegham and Chintagunta (1999) used these methods to model sales diffusion across movie box office sales in a number of international markets.

In our model, the pre- and postlaunch sales patterns of each new album can be characterized by five parameters: λ_1, c_1, λ_2, c_2, and ϕ. There are two issues that we need to take into account when estimating the model across multiple new albums. First, we must allow for heterogeneity in these parameters across these different albums. Second, we accommodate the covariation between the parameters describing the innovator segment of consumers with the parameters describing the follower segment. Therefore, the parameters for each new album, j, are distributed multivariate normal as follows:

$$\begin{bmatrix} \log(\lambda_{1j}) \\ \log(c_{1j}) \\ \log(\lambda_{2j}) \\ \log(c_{2j}) \\ \text{logit}(\phi_j) \end{bmatrix} \sim MVN(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \quad \textit{for } j = 1, 2, \ldots, J, \quad (9)$$

where J is the total number of albums in the dataset, $\boldsymbol{\mu}$ is a vector of mean parameter values, and Σ is the covariance matrix among the parameters describing the sales patterns for each album. We use a log transformation to ensure that the Weibull parameters are positive, and we use a logit transformation for ϕ to constrain it to be between zero and one. Appropriate and diffuse priors were specified for $\boldsymbol{\mu}$ and Σ. Specifically, $\boldsymbol{\mu}$ is assumed to follow a multivariate normal distribution with a zero mean and low precision while Σ was assumed to follow a Wishart distribution with $df = 5$ to represent vague priors. If we define $\boldsymbol{\theta}_j = \{\log(\lambda_{1j}), \log(c_{1j}), \log(\lambda_{2j}), \log(c_{2j}), \text{logit}(\phi_j)\}$, the likelihood across all albums in the dataset can then be written as

$$L = \prod_{j=1}^{J} \int f(\text{data} \mid \boldsymbol{\theta}_j) \cdot f(\boldsymbol{\theta}_j \mid \boldsymbol{\mu}, \boldsymbol{\Sigma})\, d\boldsymbol{\theta}, \quad (10)$$

where $f(data|\boldsymbol{\theta}_j)$ is given by the likelihood function in (8) and $f(\boldsymbol{\theta}_j|\boldsymbol{\mu}, \Sigma)$ is multivariate normal, as given in (9).

We fit the proposed model to our data using winBUGS, a statistical software package that uses Gibbs sampling and Monte Carlo Markov chain (MCMC) methods. To generate posterior estimates of the model parameters, winBUGS samples each model parameter from its prior distribution conditional on the data and the current values of all other parameters. We simulated 15,000 MCMC iterations and monitored the traces for convergence. The model appeared to converge steadily, with no problems or unusual patterns observed across the series of iterations. We discarded the first 5000 as an initial burn-in, and the posterior expectations for each of the parameters were calculated using the remaining 10,000 iterations.

We use *product-specific covariates* to explain some of the differences that exist across albums in our dataset. One such covariate is the timing of the album's prelaunch announcement. Bayus et al. (2001) showed that the timing of a new product's announcement relative to the actual release often serves as a strategic signal and can have a strong effect on how the market will respond to their product. Additionally, the amount of time between the product's announcement and release creates an opportunity for several prelaunch activities that may affect the market's response to the product, both in terms of prelaunch advanced orders and postlaunch sales. First, earlier announcements provide the marketer more opportunities to advertise the new product in an effort to increase awareness. This increased awareness may affect both the innovator segment and the prelaunch orders, as well as the followers and postlaunch sales. Secondly, a longer time before the official release of the product also creates greater opportunities for word-of-mouth effects to diffuse within and across the two consumer segments.

Therefore, we include, as a covariate in our model, the number of weeks in advance of the official release that the album was announced and available for advanced orders (w_j). We incorporate this measure into the model through the mean of the multivariate normal distribution that determines each album's parameter vector, $\boldsymbol{\theta}_j$.

$$\boldsymbol{\theta}_j \sim MVN(\boldsymbol{\mu}_j, \boldsymbol{\Sigma}) \quad \textit{where } \boldsymbol{\mu}_j = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 w_j. \quad (11)$$

In future applications it may be useful to introduce other covariates (such as artist history and music genre) into the model to further explain some of the additional variability that exists across albums. For this initial analysis, however, we want to focus on the validity of the mixed-Weibull process to determine if our basic story about pre- and postlaunch sales diffusion offers an adequate description of the data at hand.

Benchmark Models

We estimate two families of models as benchmarks: (1) Bass diffusion and (2) Weibull-gamma mixture. The continuous-time Bass model is a natural benchmark to compare against any model of new product sales:

$$F(t) = \frac{1 - e^{-(p+q)t}}{1 + \frac{q}{p} e^{-(p+q)t}}, \tag{12}$$

where t_1 represents the first week in which advance purchase orders are taken. We allow for heterogeneity in the p and q parameters across albums by using hierarchical Bayes methods where $[\log(p_j), \log(q_j)] \sim$ multivariate normal ($\boldsymbol{\mu}$, Σ). We infer market size (analogous to the m parameter in the standard discrete-time Bass model) using the same technique proposed for our mixed-Weibull model in Equation (8). In addition, we estimate another Bass model with product-specific covariates where the mean of the multivariate normal distribution, $\boldsymbol{\mu}$, is a function of the number of advance-order weeks, w_j (see Equation (11)). Overall, both versions of the Bass model fit significantly worse than does our mixed-Weibull model both in terms of log-likelihood as well as in-sample MAPE (median absolute percent error), as shown in Table 2. The general shape of the Bass model allows for a gradual buildup in sales in the early stage of the diffusion process, much as we see in the prelaunch phase. However, it is unable to accommodate the rapid changeover to the exponential decline in sales that is often observed right after product launch, as seen earlier in Figure 1.

An alternate benchmark model is a Weibull-gamma mixture. Our proposed model conceptualizes sales as coming from two discrete segments of customers, where the sales generated by each segment can be modeled as a Weibull process. The

Table 2 Comparison to Benchmark Models

	Log-Likelihood	MAPE
Proposed model, no covariates	−10925	10.6%
Proposed model, with covariates	−9095	9.8%
Bass model, no covariates	−22525	126.9%*
Bass model, with covariates	−22525	114.7%*
Weibull-gamma, no covariates	−12150	11.6%
Weibull-gamma, with covariates	−12150	11.6%

*Elton John's album is an extreme outlier with the Bass model. If we omit it from the calculations, MAPE = 43.1% without covariates and 43.0% with covariates.

Weibull-gamma, in contrast, assumes a continuous heterogeneity distribution where each customer's Weibull rate parameter (λ) varies in accordance with a gamma distribution:

$$F(t) = \int_0^\infty F(t \mid \lambda, c) \cdot g(\lambda)\, d\lambda = 1 - \left(\frac{a}{a + t^c}\right)^r, \tag{13}$$

where $F(t|\lambda, c)$ is the cumulative Weibull distribution with parameters λ and c and $g(\lambda)$ is the gamma distribution representing heterogeneity in λ. Again, we allow for heterogeneity across albums by assuming $\boldsymbol{\theta}_j \sim$ MVN($\boldsymbol{\mu}$, Σ), where $\boldsymbol{\theta}$ is the matrix of the log-transformed model parameters, a_j, r_j, and c_j and we incorporate the covariate effect of w_j through the mean of the $\boldsymbol{\theta}$ distribution, $\boldsymbol{\mu}$. We find that the continuous mixture of Weibull distributions fits far better than the Bass model, but still much worse than the discrete, time-shifted mixture of Weibull distributions in our proposed model (see Table 2).

It is interesting to note that neither of these benchmark models makes significant use of the number of advance-order weeks, w_j, as a covariate to explain differences across each album's parameters. In contrast, Table 2 shows that this covariate offers a significant boost to the performance of the proposed model. Furthermore, as we will discuss in the next section, the w_j factor also offers a great deal of useful insight about the nature of (and interplay between) the innovator and follower customer segments.

Table 3 Model Parameter Estimates

	Intercepts			Slopes	
λ_1	β_{01}	−0.5636		β_{11}	−0.9783
c_1	β_{02}	0.6278		β_{12}	0.0648
λ_2	β_{03}	−0.0266		β_{13}	−0.8220
c_2	β_{04}	0.1345		β_{14}	−0.0384
ϕ	β_{05}	0.8147		β_{15}	−0.2945
	0.4797	0.7039	−0.0714	−0.1452	0.1471
	0.7039	4.7100	−0.2714	−1.0190	0.5093
$\Sigma =$	−0.0714	−0.2714	0.3642	0.1804	−0.1232
	−0.1452	−1.0190	0.1804	1.6830	−0.3027
	0.1471	0.5093	−0.1232	−0.3027	0.7153

Results

Summary of Parameter Estimates

Table 3 reports the posterior estimates for the model parameters across the 10,000 MCMC draws. The parameters reported in the table represent the log-transforms (and the logit transform in the case of ϕ). The λ parameters reflect the rate of adoption within each segment (λ_1 for innovators and λ_2 for followers): The larger the λ, the more quickly the segment members will buy a given album. The c parameter for each segment (c_1 for innovators and c_2 for followers) describes the buildup (or slowdown) pattern of the adoption process. For $c < 1$, we expect declining sales throughout the life of the album (for that segment of customers), while $c > 1$ results in a (roughly) bell-shaped curve with a buildup in sales in the first few weeks after that segment has been exposed to the new album.

Figure 3 (see p. 357) plots the changes in the expected values (and 95% confidence bounds, based on the 10,000 MCMC draws) of model parameters as we vary the number of prelaunch weeks. For each segment, the rate of purchase (λ_1 and λ_2) decreases as the number of prelaunch weeks increases. More interesting is the effect on the shape parameter for each segment. For the innovators, the value of c_1 increases quite substantially with more prelaunch weeks and is always greater than one. This indicates that the sales pattern among innovators is more or less bell shaped and becomes more so with additional prelaunch weeks. Coupled with the effect on λ_1, the sales diffusion among the innovators becomes more diffuse for albums with more prelaunch weeks (Figure 4).

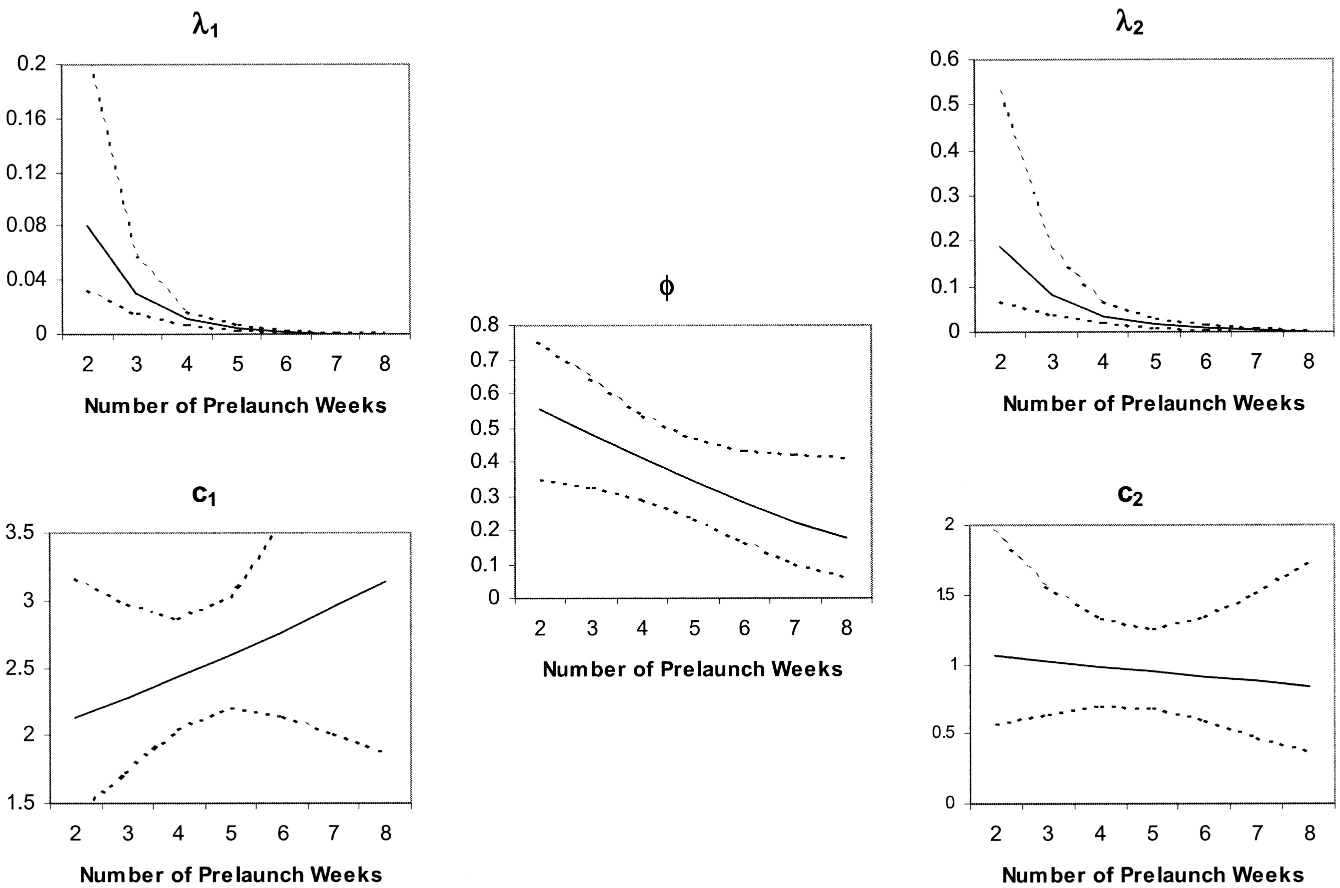
For the follower segment, the number of prelaunch weeks has a different effect on the c_2 parameter, i.e., a slight decline when more prelaunch weeks are present. Yet, note that we cannot reject the hypothesis that c_2 equals one regardless of the length of the advanced order period. In other words, the followers' behavior appears to be highly consistent with an exponential sales pattern.

The effect of the number of prelaunch weeks on the follower's purchase rate is much like that of the innovator group, with λ_2 decreasing as the number of prelaunch weeks increase. When these dynamics are combined, we see (in Figure 4) that the overall Weibull distribution for the followers is similar to that of the innovators in the sense that a larger number of prelaunch weeks is associated with a more diffuse sales pattern.

The final panel of Figure 3, for ϕ (the segment-size parameter), shows that a longer prelaunch period is associated with a follower segment that is larger (on a relative basis) compared to the innovators. This makes intuitive sense, since a longer gestation period for word-of-mouth effects will help create a larger set of potential followers.

Rather than examining each component of sales diffusion in isolation, we bring together the effects of all of these parameters across both segments and see how the number of prelaunch weeks impacts the overall pattern of sales diffusion (third panel of Figure 4) (see p. 358). A new album released with only two weeks of prelaunch activity is likely to experience dramatically higher sales in its launch week than an album that has a longer prelaunch period. This would seem to suggest that more hype in the prelaunch period (i.e., a longer time before launch) does not necessarily translate into a bigger launch week. In fact, despite a short prelaunch period, albums with fewer prelaunch weeks capture a higher percentage of their overall sales in this early period, all else being equal. For

Figure 3 Effect of the Number of Prelaunch Weeks



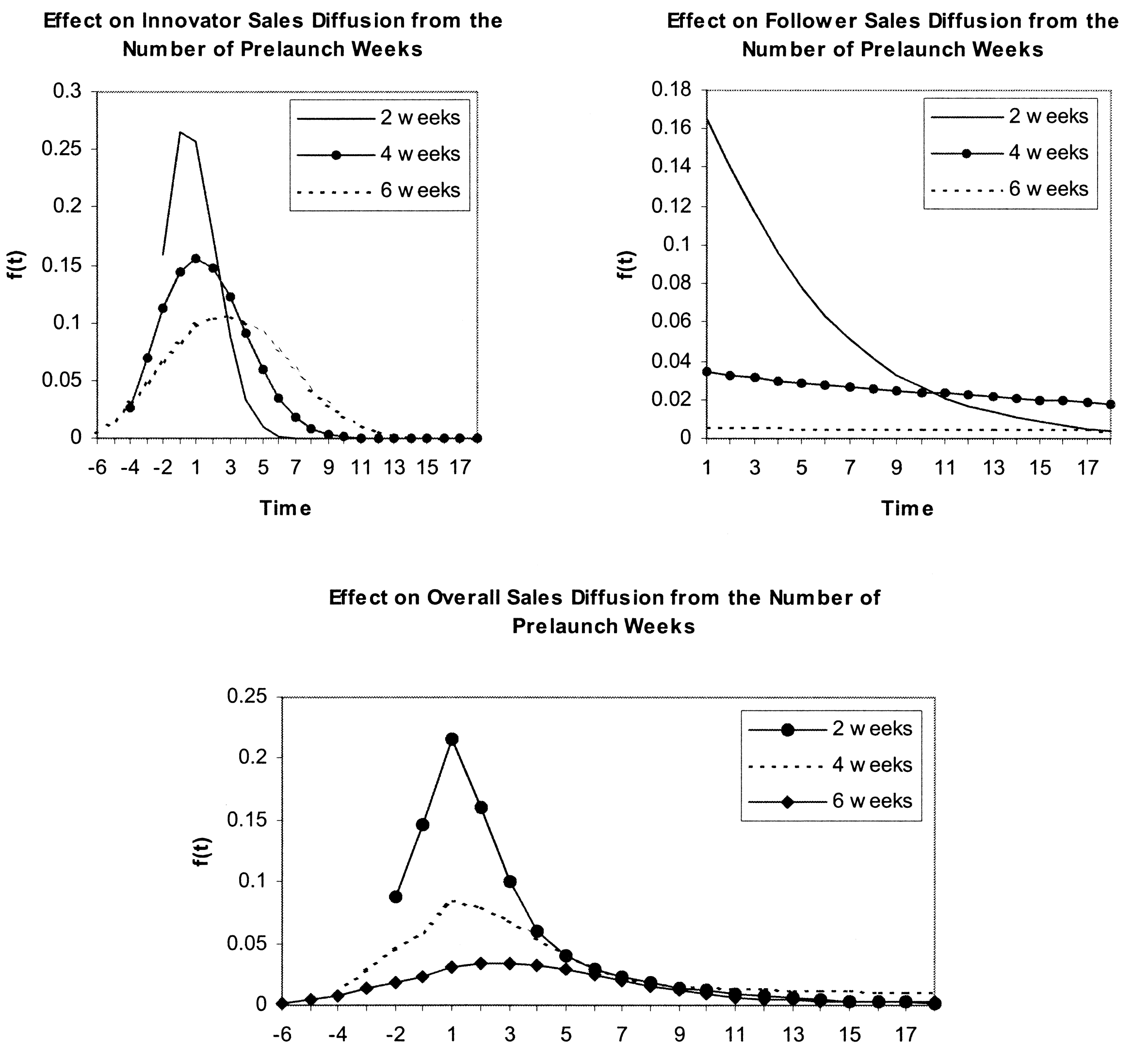
example, a two-week prelaunch would, on average, capture 45% of its total sales in the prelaunch period and launch week combined (compared to only 23% with a four-week prelaunch and 10% with a six-week prelaunch).

Model Fit

Figure 5 (see p. 359) plots the model fit for the two albums described earlier. Expected sales track the actual sales extremely well, with MAPEs of 5.6% for *Lilith Fair* and 9.3% for *Big Bad Voodoo Daddy.*

When we compare the segment-level decompositions in Figure 6 (see p. 360), we see an interesting contrast in the patterns underlying the overall sales for each of the albums. The innovator segment, which begins purchasing before the launch of the album, shows a buildup in buying probabilities (i.e., $c_1 > 1$) for both albums in the prelaunch phase. According to the parameter estimates shown in Table 4 (see pp. 361 and 362), *Lilith Fair* has a much higher rate parameter (λ_1) for this segment, so the innovators for this album rapidly take themselves out of the market within a few weeks of launch. Virtually all of the sales after mid-May are due to the followers, who seem to exhibit a conventional exponential-like decay pattern ($c_2 < 1$). On the other hand, *Big Bad Voodoo Daddy* has a much slower innovator rate, so this segment continues to dominate the buyer base for this album for months after the launch date. This difference across the two albums is also accentuated by the fact that the innovator segment is twice as large (on a relative basis) for *Big Bad Voodoo Daddy* as it is for *Lilith Fair* (as seen in the values of the ϕ parameter). Ironically, the follower segments show

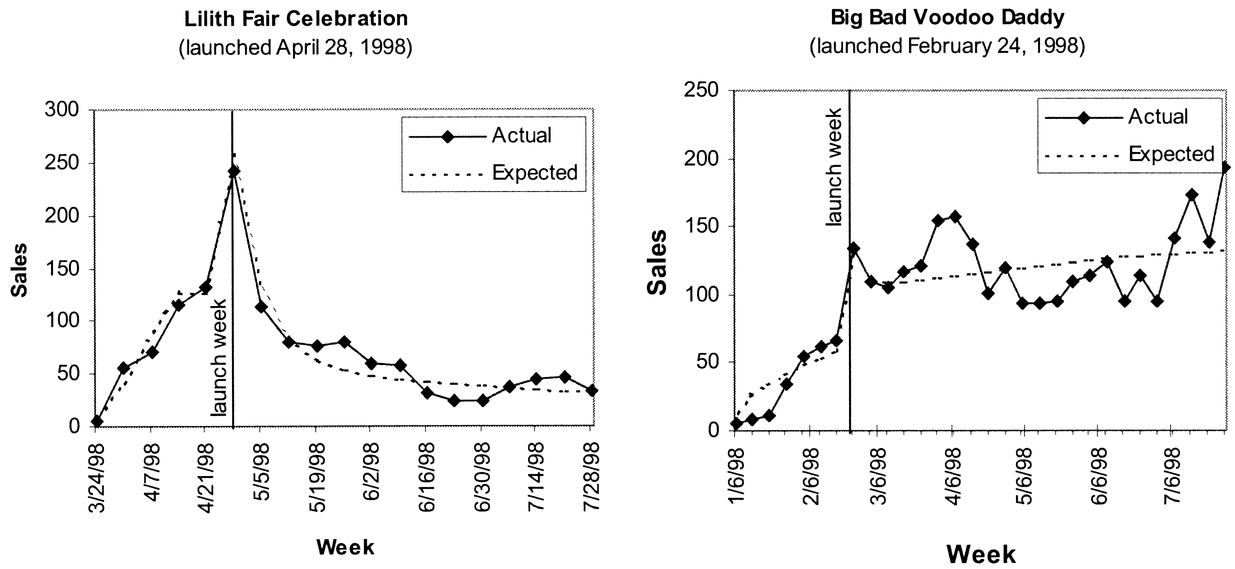
Figure 4 **Changes in Sales Patterns due to Variation in Prelaunch Weeks**



the same basic sales pattern over time for the two albums, but this similarity is largely inconsequential in light of the differences described above. The bottom line here is that the *Big Bad Voodoo Daddy* album has much greater future sales potential than does *Lilith Fair*. By the end of the observation period, *Big Bad Voodoo Daddy* has attained approximately 5% of its cumulative sales, whereas *Lilith Fair* album has already achieved a larger fraction of its total lifetime sales (30%).

Table 4 provides a similar analysis for each of the albums in our data set. The first five columns of the table report the album-specific parameter estimates. The next column shows $\hat{F}(T)$, the fraction of eventual sales achieved by each album by the end of the observation period. Coupled with the actual cumulative sales for each album at the end of the observation period, one can estimate the total expected market size for any given album over its lifetime, using Equation (8).

Figure 5 Model Fit



Correlation

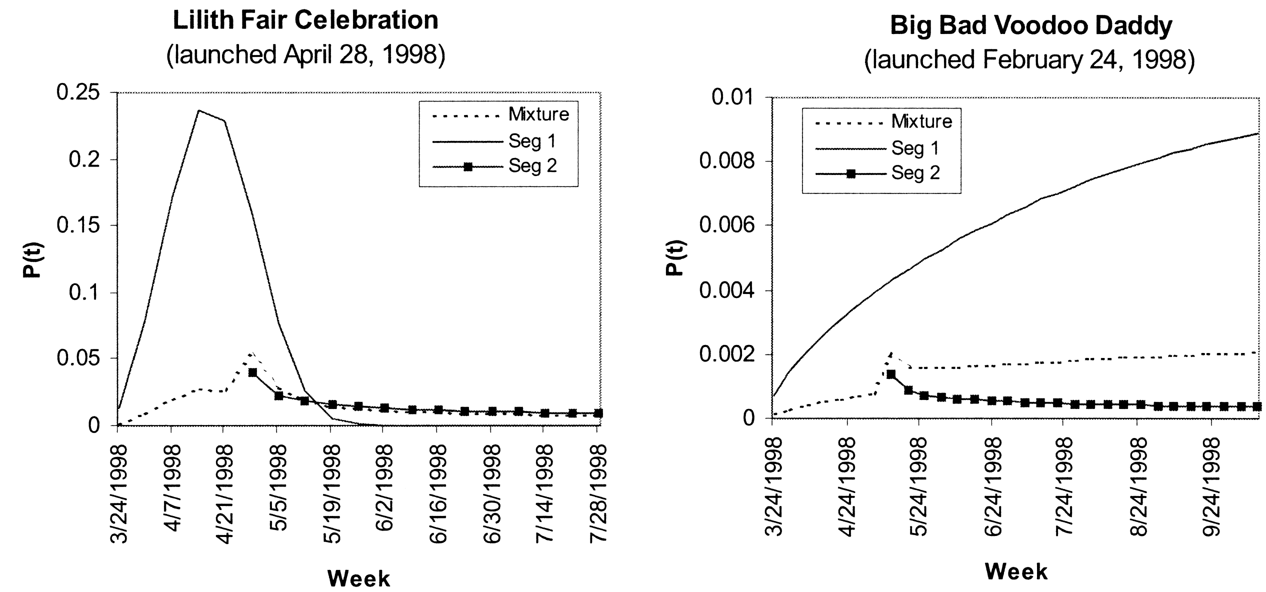
At this point, we turn to the correlations among the model parameters, which is not only an effective way to summarize much of the information shown in Table 4 but also provides useful insights about how the characteristics of the innovator segment associate with the likely behavior of the followers. The correlation matrix for the model with album-specific covariates is as follows

$$\begin{array}{c} \qquad\; \log(\lambda_1) \quad \log(c_1) \quad \log(\lambda_2) \quad \log(c_2) \quad \mathrm{logit}(\phi) \\ \rho = \begin{bmatrix} 1.0000 & & & & \\ 0.4683 & 1.0000 & & & \\ -0.1708 & -0.2072 & 1.0000 & & \\ -0.1616 & -0.3619 & 0.2304 & 1.0000 & \\ 0.2511 & 0.2775 & -0.2414 & -0.2759 & 1.0000 \end{bmatrix}. \end{array}$$

Overall, there appears to be some positive correlation between the rate and shape parameters within each segment, although it is not dramatic. Additionally, each matched pair of the Weibull parameters shows a negative correlation across the two segments ($\rho_{\lambda_1,\lambda_2} = -0.1709$, $\rho_{c_1,c_2} = -0.3619$), suggesting that more rapid diffusion among one segment is associated with less rapid diffusion among the other. We also see some moderately high correlations involving the segment size parameter, ϕ. This means that faster diffusion among the innovators is associated with a greater proportion of sales from that segment relative to other albums. Equivalently, a faster rate of diffusion among followers is associated with a smaller relative size for the follower segment.

These correlations are consistent with the patterns shown earlier in Figure 3. The full set of these cross-parameter associations conveys a number of plausible relationships: (1) The length of the prelaunch period is negatively correlated with the innovators' rate of adoption, i.e., a longer prelaunch period is often accompanied by less urgency for the innovators to act quickly; (2) the length of the prelaunch period is positively correlated with the relative size of the follower population, i.e., a longer prelaunch period creates greater awareness and word-of-mouth opportunities, thereby attracting more followers; and (3) the relative size of followers is negatively associated with the rate of diffusion among the follower group, i.e., a broader set of followers will include a larger number of relative laggards, slowing down the diffusion process for the group as a whole.

Figure 6 **Mixed-Weibull Distributions**



Forecasting

One application of the proposed model is to provide a tool with which to forecast future product sales based solely on prelaunch advance orders. To test the forecasting abilities of this model, we randomly assigned the 66 albums into two separate groups. We used one group of 33 albums to estimate the model parameters. We then used these estimates in conjunction with actual observed prelaunch sales for the remaining 33 albums to obtain album-specific posterior estimates for each holdout album, $E[\boldsymbol{\theta}_j|\boldsymbol{\beta}_0, \boldsymbol{\beta}_1, \Sigma, \text{Sales}_t]$. This was accomplished using MCMC to simulate 10,000 different parameter vectors and, hence, sales patterns for each album. Because of the covariance structure among the parameters, we were able to estimate the expected value of the follower parameters and the segment weights, even though we observed only prelaunch sales. We then reversed the process, using the originally held-out albums for model estimation and the other half of the albums for forecasting. As a result, we generated forecasts for all 66 albums based solely on each one's prelaunch sales patterns.

We first test how well the future *pattern* of sales can be forecasted. To do this, we assume that total sales by the end of the observation period are known. In this case, we calculate the posterior parameter estimates to generate $P'(t)$ and test the fit of the multinomial distribution given in (7). For each album, we calculate the MAPE for each iteration and report the median MAPE for each album, as well as across albums. Table 5 illustrates the accuracy of the forecasts by reporting the number of albums with forecasts that fall within a given range. For over half of the albums, the model's holdout forecasting accuracy falls within 20% of actual sales. This includes both of our example albums, *Lilith Fair* and *Big Bad Voodoo Daddy*, with MAPEs of 12.1% and 12.6%, respectively. Again, this illustrates the ability of the model to accommodate very different sales patterns that may exist.

Although some of the forecasts shown in Table 5 (see p. 362) are less accurate than others, it is important to note that these forecasts are based only on prelaunch advance orders. For some albums, this can be as little as two weeks of data. Forecast accuracy will only improve as more sales data is observed. Furthermore, many albums are also greatly

Table 4 Estimated Parameters for Each Album

Artist	Title	λ_1	c_1	λ_2	c_2	ϕ	$\hat{F}(T)$*
Dave Matthews Band	Before These Crowded Streets	0.049	2.317	0.058	1.147	0.464	0.84
Madonna	Ray of Light	0.002	3.194	0.002	1.933	0.472	0.79
Sarah McLachlan	Surfacing	0.000	1.364	0.001	0.100	0.617	0.03
Eric Clapton	Pilgrim	0.016	2.545	0.040	1.286	0.551	0.94
Elton John	Something About the Way You . . .	0.025	3.912	0.054	1.079	0.802	0.99
Celine Dion	Let's Talk About Love	0.010	1.469	0.062	0.479	0.663	0.70
Bob Dylan	Time Out of Mind	0.006	1.576	0.012	0.165	0.452	0.43
Natalie Imbruglia	Left of the Middle	0.009	2.789	0.027	1.113	0.152	0.62
Beastie Boys	Hello Nasty	0.009	2.526	0.023	0.445	0.400	0.35
Pearl Jam	Yield	0.009	2.883	0.064	0.913	0.272	0.79
Garbage	Version 2.0	0.016	2.253	0.001	1.502	0.147	0.19
Tori Amos	From the Choirgirl Hotel	0.009	2.555	0.004	2.401	0.752	0.95
Natalie Merchant	Ophelia	0.005	2.857	0.001	2.100	0.371	0.49
Smashing Pumpkins	Adore	0.007	2.863	0.001	3.480	0.702	0.98
Enya	Paint the Sky With Stars-Best	0.004	2.570	0.004	1.704	0.389	0.89
Big Bad Voodoo Daddy	Big Bad Voodoo Daddy	0.001	1.658	0.001	0.697	0.201	0.05
Dave Matthews Band	Live at Red Rocks 8/15/95	0.012	0.684	0.005	0.096	0.646	0.10
Garth Brooks	Limited Series Box Set	0.017	2.379	0.051	0.547	0.098	0.27
Paul Simon	Songs From the Capeman	0.002	3.141	0.019	1.560	0.622	1.00
Metallica	Re-Load	0.013	2.398	0.001	1.520	0.210	0.41
Garth Brooks	Sevens	0.022	2.468	0.003	1.898	0.508	0.96
Bonnie Raitt	Fundamental	0.014	3.249	0.168	0.922	0.273	0.93
Ally McBeal	TV Soundtrack	0.016	2.429	0.023	0.786	0.102	0.24
Barenaked Ladies	Stunt	0.002	2.050	0.015	0.777	0.299	0.09
Van Halen	Van Halen 3	0.002	3.747	0.123	0.862	0.732	0.95
Janet Jackson	Velvet Rope	0.003	1.239	0.019	0.371	0.455	0.19
Page & Plant	Walking Into Clarksdale	0.013	2.986	0.050	1.093	0.560	0.83
Lucinda Williams	Car Wheels on a Gravel Road	0.006	2.301	0.161	1.602	0.648	0.76
Diana, Princess of Wales	Diana, Princess of Wales Tribute	0.026	2.758	0.274	0.609	0.489	0.95
Jimmy Buffett	Don't Stop the Carnival	0.044	2.257	0.017	0.601	0.101	0.17
Brian Setzer Orchestra	Dirty Boogie	0.000	3.078	0.005	3.992	0.862	0.55
Barbra Streisand	Higher Ground	0.000	3.670	0.033	1.169	0.626	0.96
Ani DiFranco	Little Plastic Castle	0.021	2.241	0.051	1.165	0.388	0.92
Massive Attack	Mezzanine	0.006	2.686	0.005	1.281	0.152	0.25
U2	Pop	0.110	0.585	0.854	0.884	0.917	0.78
Van Morrison	Philosopher's Stone	0.002	3.445	0.009	2.204	0.633	0.81
John Fogerty	Premonition	0.002	3.757	0.045	1.137	0.303	0.57
Propellerheads	Decksandrumsandrockandroll	0.016	2.168	0.035	0.885	0.231	0.52
Godzilla	Soundtrack	0.001	4.295	0.018	2.087	0.262	0.95
Lilith Fair	A Celebration of Women	0.014	2.831	0.041	0.664	0.116	0.30
Led Zeppelin	BBC Sessions	0.002	3.091	0.001	2.099	0.430	0.96
Radiohead	Airbag/How Am I Driving?	0.012	2.452	0.000	2.334	0.234	0.39
Bragg, Billy & Wilco	Mermaid Avenue	0.001	5.079	0.106	1.609	0.183	0.88
X-Files	Movie Soundtrack	0.016	2.129	0.314	0.937	0.503	0.94
Mavericks	Trampoline	0.014	2.599	0.017	1.099	0.194	0.50
Liquid Tension Experiment	Liquid Tension Experiment	0.001	3.991	0.010	1.139	0.112	0.35
Nanci Griffith	Other Voices Too (A Trip Back . . .)	0.004	2.896	0.027	0.357	0.360	0.27
Frank Sinatra	In the Wee Small Hours	0.000	8.437	0.091	0.760	0.361	0.26
Genesis	Vol. 1—Genesis Archives 1967–75	0.061	1.576	0.007	3.443	0.846	0.91

Table 4 continues

Table 4 Continued

Artist	Title	λ_1	c_1	λ_2	c_2	ϕ	$\hat{F}(T)$*
Jeff Buckley	Sketches for My Sweetheart . . .	0.009	2.657	0.014	1.174	0.256	0.39
Lenny Kravitz	5	0.008	3.066	0.027	1.133	0.165	0.46
Yes	Keys to Ascension 2	0.019	1.988	0.017	0.457	0.095	0.17
Jimi Hendrix	BBC Sessions	0.014	2.907	0.045	0.814	0.065	0.28
Leann Rimes	Sittin' on Top of the World	0.006	2.722	0.021	0.950	0.176	0.35
Frank Sinatra	Only the Lonely	0.000	6.676	0.290	0.888	0.482	0.94
Marc Cohn	Burning the Daze	0.019	2.610	0.095	0.875	0.219	0.79
Counting Crows	Across a Wire—Live in New York	0.020	2.269	0.801	2.190	0.778	0.85
Maxwell	Embrya	0.003	2.984	0.006	0.980	0.465	0.46
Brian Wilson	Imagination	0.012	2.307	0.008	0.810	0.295	0.31
Various Artists	Where Have All the Flowers Gone?	0.008	3.334	0.079	0.797	0.102	0.62
Tricky	Angels With Dirty Faces	0.003	3.774	0.066	1.021	0.289	0.62
Xena: Warrior Princess	Bitter Suite—A Musical Odyssey	0.238	0.859	0.864	0.732	0.772	0.98
Dwight Yoakam	Long Way Home	0.013	2.015	0.036	0.266	0.276	0.29
Rod Stewart	When We Were the New Boys	0.001	4.446	0.022	1.302	0.340	0.55
David Lee Roth Band	DLR Band	0.035	1.518	0.762	0.278	0.728	0.73
Gillian Welch	Hell Among the Yearlings	0.005	2.797	0.018	1.149	0.477	0.17

*Cumulative distribution function achieved by the end of the observation period.

Table 5 Tracking Performance in Holdout Period

MAPE Is Within	No. of Products
10%	24
20%	18
30%	7
40%	4
50%	3
>50%	10

affected by exogenous factors such as radio exposure, concert tours, Grammy nominations, and so on. Including covariates to capture some of these time-varying influences (as well as purely cross-sectional effects, such as artist history) may offer a significant improvement over the basic model discussed here. Nevertheless, our basic model performs extremely well, considering that these effects have been completely ignored.

In addition to the ability to accurately forecast the pattern of postlaunch sales, we also wish to forecast the overall volume of postlaunch sales, based only on prelaunch orders. We use the following four-step procedure to estimate the potential market size, $\hat{N}$, as well as the sales levels for each of the postlaunch weeks:

1. Determine $F(t^* - 1)$, i.e., the fraction of orders expected to have been placed by the week before album launch, which is given by (3).
2. Estimate the potential market size, $\hat{N}$, using this value of $F(t^* - 1)$ and the total observed prelaunch sales, as shown in (8).
3. Project weekly postlaunch probabilities, $P(t)$, for each week using (7).
4. Total sales during the observation period can be forecasted as $P(T)$ from Step 3 multiplied by $\hat{N}$ from Step 2.

Using this procedure, we obtain an empirical distribution of expected postlaunch sales using the full set of 10,000 MCMC simulations. Then for each album we can determine where the actual number of sales falls within this distribution of predictions, so as to determine how well the simulated draws frame this observed value. To test the accuracy of our predictions, we create a two-tailed "p-value" for each album by counting the fraction of simulated draws that are at least as far away from the mean as is the actual value. A p-value of 1.0 represents a perfect forecast, but a p-value of less than, say, 0.05 is way off the mark. Of the 66 albums, 50 have p-values greater than 0.20, suggesting that the forecasts are fairly accurate. Wallis (1942) provided a formal test to evaluate the overall significance of a collection of p-values, which was introduced to the marketing literature by Dutka (1984). Specifically, under the null hypothesis that the model is true, the test statistic, $2\Sigma\log(p)$, is distributed χ^2 with $2n$ degrees of freedom (where n is the number of p-values being examined). In our case, $\chi^2 = 62.7$ with 132 degrees of freedom, providing an overall p-value of well above 0.5. This suggests that our prelaunch model provides accurate forecasts for actual postlaunch sales.

It is worth emphasizing that this test does not only evaluate the accuracy of the point estimate (i.e., expected value) of the forecast for each album, but it also takes into account the variability (based on the model) around each of these estimates. In this sense, we are assessing both the mean and variance of these forecasts, so it is a doubly meaningful accomplishment to pass this hurdle. This test is well suited for other types of hierarchical Bayes models, and we encourage its use more broadly.

Discussion

Marketers have long struggled with developing forecasts for new products before their launch. We focus on a data source—advance purchase orders—that has been available to retailers for many years but has rarely been tied together with postlaunch sales data. We put forth a duration model that incorporates the basic concepts of new product diffusion, using a mixture of two distributions: One represents the behavior of innovators (i.e., those who place advance orders) and one represents the behavior of followers (i.e., those who wait for the mass market to emerge). The resulting mixed-Weibull model specification can accommodate a wide variety of possible sales patterns. This flexibility is what makes the model well-suited for an experiential product category (e.g., movies, music, etc.) in which we frequently observe very different sales diffusion patterns, ranging from a rapid exponential decline (which is most typical) to a gradual-buildup characteristic of "sleeper" products. We incorporate product-specific covariates and use hierarchical Bayes methods to link the two customer segments together while accommodating heterogeneity across products.

We find that this model fits a variety of albums very accurately. The model also provides the ability to forecast new album sales before the actual launch of the album, based on observable behavior. However, in a few cases, unspecified exogenous factors may have caused actual sales to diverge from what was anticipated. Future research may look to incorporate a more complete set of covariates (e.g., radio airplay, concert dates, award nominations, etc.) to obtain more accurate (and managerially useful) forecasts. To go a step farther, researchers can also explore the possibility of predicting some of these events based on prelaunch behavior.

There may be some important policy implications associated with this modeling framework and our empirical results. Some of the research cited earlier (Bayus et al. 2001, Desiraju and Shugan 1999, Shugan and Xie 2000, Xie and Shugan 2001) address questions concerning optimal timing and pricing policies at the prelaunch phase for a new product, but the underlying behavioral models are not nearly as rich (or realistic) as the specification we developed here. However, having established the various behavioral patterns that our model helped uncover, it now makes sense to connect these two streams of research. This will likely require additional datasets with more covariates (e.g., price) and more variation in some of the key elements (e.g., earlier prelaunch announcements) to fully understand how these key elements interact with one another.

In summary, our model and empirical analysis can open up many new avenues to expand the existing literature (both descriptive and normative) on preannouncement behavior for new products. We hope that the framework presented here provides a useful step for future researchers who wish to address some of these important issues.

Acknowledgments

The authors thank Eric Bradlow and Christophe Van den Bulte for their very helpful comments on earlier drafts of this paper.

References

Bass, F. M. 1969. A new product growth model for consumer durables. *Management Sci.* **15** (January) 215–227.

Bayus, Barry L., Sanjay Jain, Ambar G. Rao. 2001. Truth or consequences: An analysis of vaporware and new product announcements. *J. Marketing Res.* **38** (February), 3–13.

Desiraju, Ramarao, Steven M. Shugan. 1999. Strategic service pricing and yield management. *J. Marketing* **63** (1) 44–56.

Dutka, Solomon. 1984. Combining tests of significance in test marketing research experiments. *J. Marketing Res.* **21** (February) 118–119.

Jiang, R., D. N. P. Murthy. 1998. Mixture of Weibull distributions—Parametric characterization of failure rate function. *Appl. Stochastic Models Data Anal.* **14** 47–65.

Krider, Robert E., Charles B. Weinberg. 1998. Competitive dynamics and the introduction of new products: The motion picture timing game. *J. Marketing Res.* **35** (February) 1–15.

Lenk, P. J., A. G. Rao. 1990. New models from old: Forecasting product adoption by hierarchical Bayes procedures. *Marketing Sci.* **9** (1) 42–57.

Mahajan, Vijay, Eitan Muller, Frank M. Bass. 1995. Diffusion of new products: Empirical generalizations and managerial uses. *Marketing Sci.* **14** (3) G79–G88.

Moe, Wendy W., Peter S. Fader. 2001. Modeling hedonic portfolio products: A joint segmentation analysis of music CD sales. *J. Marketing Res.* **38** (3) 376–385.

Neelamegham, Ramya, Pradeep Chintagunta. 1999. A Bayesian model to forecast new product performance in domestic and international markets. *Marketing Sci.* **18** (2) 115–136.

Rogers, E. M. 1983. *Diffusion of Innovations*, 3rd ed. The Free Press, New York.

Sawhney, Mohanbir S., Jehoshua Eliashberg. 1996. A parsimonious model for forecasting gross box office revenues. *Marketing Sci.* **15** (2) 113–131.

Shugan, Steven M., Jinhong Xie. 2000. Advance pricing of services and other implications of separating purchase and consumption. *J. Service Res.* **2** (February) 227–239.

Swift, Ronald S. 2001. *Accelerating Customer Relationships.* Prentice Hall, New York.

Wallis, W. Allen. 1942. Compounding probabilities from independent statistical tests. *Econometrica* **10** (3/4) 229–248.

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

This paper was received March 22, 2002, and was with the authors 1 month for 1 revision; processed by Pradeep Chintagunta.