8

A Survival Analysis of Albums on Ranking Charts

Sudip Bhattacharjee[1], Ram D. Gopal[2], James R. Marsden[3], Rahul Telang[4]

[1,2]*University of Connecticut*
[3,4]*Carnegie Mellon University*

Introduction

Maintaining security in the digital world continues to grow in complexity. Firms must protect operating hardware and sensitive data against increasingly innovative threats. With the emergence of digital goods comes a new security front where firms face the reproduction and rapid distribution of the digital goods themselves. Certainly, firms already have had to protect many of their goods from "knockoffs," but protecting digital goods represents a new level of challenge since the cost of copying and distributing such goods is virtually zero and can occur extensively within very short periods of time. The music industry has been the "poster industry" for facing such threats. The industry's goods are digital by nature. Further, the appearance of Peer-to-Peer networks offered the means to copy (download) the goods and distribute (share) them rapidly.

Through its industry association, the Recording Industry Association of America (RIAA), the music industry has continued to pursue mostly legal and technological strategies to eradicate the security threat of illegal copying and distribution. A recent study by Bhattacharjee et al. (2006a,b) provides evidence that, while individuals have tended to reduce their own sharing activity in response to RIAA's legal threats and actions, significant piracy opportunity remains. While individual firms may take steps to secure their digital goods, such constraints have two major drawbacks. First, such measures tend to impede the consumer's use of the digital good since they can restrict portability or require additional steps (e.g., security actions) that reduce consumer utility (Halderman 2002). Second, the measures have

E.M. Noam and L.M. Pupillo (eds.), *Peer-to-Peer Video*, doi: 10.1007/978-0-387-76450-4_8,

proven less than "foolproof" and rather easily beaten (Reuters 2002, Felten 2003). Sony BMG's recent use of a rootkit with the XCP technology (Reuters 2005, Bergstein 2005) provides a prominent illustration of how an attempted technological security constraint can backfire (Bradley 2005):

> Part of Sony's anti-pirating strategy is that some of its music will play only with media software included on the CD. When a user inserts the CD, he or she is asked to consent to an "end user licensing agreement" for a Digital Rights Management application. If the user agrees, the rootkit automatically installs and hides (or "cloaks") a suite of DRM software.

Unfortunately, the rootkit application created a possible secret backdoor for hackers which led Sony to "hastily" post a patch. However, the tool to remove the XCP application itself created new vulnerabilities (Russinovich 2005). The tale continues as California quickly filed suit under both unfair and deceptive trace acts and consumer protection acts, Texas filed suit for including "spyware" in its media player, and the Electronic Frontier Foundation filed suit seeking class-action status over its copy-protection software (Smith 2005). A posting (by concord (198387), 11/10/05, #13996982) in slashdot.org's bulletin board offers the following perspective on the Sony anti-piracy actions:

> Now for the first time it is actually safer to download and listen to pirated music then [sic] it is to purchase and use compact disks and DVDs. Piracy will become a matter of self-preservation.

In addition, security professionals have consistently noted that all CD and DVD encryption techniques that have been tried by the entertainment industry have been broken by savvy consumers (Schneier 2000, Craver et al. 2001, Patrizio 1999, Clarke 2005, Associated Press 2003). Given wide dissemination of the encrypted music product among users (factors that make breaking encryption easier), it is not unusual to observe such copy protection technologies being defeated by smart users (Bergstein 2005, Felten 2005).

Thus a firm considering possible actions to protect its digital product may find little return in costly technological and legal anti-piracy measures. But can the firms identify and respond to the changing market they face? The post 1998–1999 period is characterized by consumers who increasingly search and consume music products in digital formats. Here we focus attention on what significant changes have occurred in the landscape of music products and their market life cycle since the introduction of significant new technology, including Peer-to-Peer networks, and other market forces (including online music stores, higher penetration of broadband into homes, digital rights management (SDMI initiative), and evolving copyright laws (DMCA 1998, Sonny Bono Copyright Term Extension Act 1998). We first develop an analytical model of music album life cycle to provide a robust foundation to develop effective decision making tools for a music company to better manage its music products in the market place. The model demonstrates how the pattern of album life cycle has undergone a shift in the years following the introduction of new technologies and other market forces. Following the analytic model, we examine how some of the album specific and artist specific variables affect the album survival. Thus, incorporating key exogenous factors helps a decision maker to better predict and respond to market success of a digital good in a dynamic environment. A firm's ability to act with these decision tools, which combine product life cycle analytics with analysis of consumer actions on online computer networks, would provide greater market value protection for the firm's digital products than would technological and legal anti-piracy measures alone.

The Landscape – Rankings and Survival Longevity

In a number of domains – including music, movies, books, university sports, and academics – rankings are the yardsticks to measure success. Appearance and longevity of survival on ranking charts are important for market success and job security. Rankings have limited slots (e.g., top 10, top 25, or top 100) and are reported on a periodic basis (ranging from weekly for music charts or in-season sports to annually for business school rankings).

High rankings and longevity on ranking charts would seem to have inherent links to the concept of "superstars," a phenomenon studied by Rosen (1981). Following Rosen's initial work, Adler (1985) suggested the existence of the superstar phenomenon in artistic industries where only a relatively small number of music artists and their products garner enormous success. Adler argued that consumers minimize the cost of search by simply choosing artists who are already popular among other consumers. Adler's "concentration of success" phenomenon has been empirically studied by several authors (Simon 1955, Yule 1924). Examples cover quite a range and include Albert's (1998) analysis of motion pictures, Cox and Chung's (1991) study of research output in academics, Simon's (1955) examination of the distribution of words in prose, and Levene et al.'s (2002) consideration of the growth of Internet websites. Approximately 30,000 albums are released annually by the major music labels alone (Goodley 2003). Given that a mere handful of successful albums can significantly affect the profitability of a music label, it is critically important for the labels to have an estimation of the potential life cycle of the albums released early in their release period. This would enable them to form informed decisions and channel limited marketing and promotional budgets towards potential winners.

But what happens if the landscape changes significantly and past business practices do not apply as well? What happens when advances occur in markets that make a consumer's search for information and product access far easier? Does ranking longevity, or life cycle on the chart, change dramatically? In fact, in the past few years, the music industry has seen such a technological and market revolution. Easier search for information and product sampling is an integral part of buying an experience product such as music. The advent of MP3 and online file-sharing technologies now allow consumers to access and exchange millions of digitized music files over Peer-to-Peer networks.[1,2]

We develop a stochastic model of the distribution of album longevity on the Billboard Top 100 Chart.[3] We estimate the model annually for periods before and after the major technological and market changes, that is, the introduction of MP3, broadband, and the Napster Peer-to-Peer online sharing technology that took place over the 1998–1999 period. What we find is that, despite declining numbers of new album releases after 1999 (Ziemann 2002), the probability of survival on the Billboard Chart had a major shift downward. We use this survival information and develop a regression model that incorporates consumer behavior on online file-sharing networks. This is used to estimate the continued success of albums on the Billboard Chart. We emphasize that the same stochastic model form yields similar useful fit results for the differing periods with, of course, different parametric estimation values. Continued refreshing of the model estimation can be utilized by firms as a benchmark to adjust their decision making on individual albums as the market continues to shift over time.

The Stochastic Model of Survival

Rankings and longevity on the charts is a key indicator of a music album's success, and is closely followed by music labels and music industry analysts each week. Since 1913, *Billboard* magazine has provided weekly summary chart information based on sales of music recordings (Bradlow and Fader 2001). The Billboard Top 200 Chart is based on "…a national sample of retail store sales reports collected, compiled, and provided by Nielsen Soundscan" (from Billboard website). We use the freely available list of the weekly top 100 albums in our analysis. Based on empirical observation, we assume that once an album drops off the Billboard Top 100, the album does not re-appear on the chart.[4] Thus each week, some albums drop off the ranking chart and an equal number of albums appear for the first time. At the end of a hypothetical "first" week of the chart, we would have 100 albums that have appeared on the chart for exactly one week.

Let p_i denote the probability that an album will remain on the chart for one more week after having been on the chart for exactly i weeks. In week 2, the expected number of albums that would remain on the chart is $100p_1$. The expected number of albums that drop out of the chart after the first week is $100(1\text{-}p_1)$ which is also the same number of albums expected to enter the chart for the second week since there must be a total of 100 active albums on the Billboard Top 100 chart in any given week. That is, we model a stochastic process with one absorptive state that might be termed "falling off the chart." Table 8.1 details the stochastic process for the first three periods.

More formally, let $C_{k,i}$ indicate the number of albums that appear on the kth week's Billboard Chart and have appeared for i weeks ($i = 1,\ldots,k$). $C_{12,5}$ would be the number of albums on the 12th week's chart that had appeared for 5 weeks (charts 8 through 12). Let $D_{k\text{-}1,w}$ be the number of albums that appeared on week k-1's chart, do not appear on week k's chart, and which were on the charts for w weeks. $D_{21,4}$ would be the number of albums that met the following criteria: appeared on chart 21, did not appear on chart 22, and appeared on the chart for 4 weeks (from weeks 18 to 21).

Table 8.1 Illustration of Stochastic Process

	Expected number of albums that have been and are currently on the chart for:				Expected number of albums that had dropped out of the chart after:			
	1 week	2 weeks	3 weeks	4 weeks	1 week	2 weeks	3 weeks	4 weeks
Week1	100	0	0	0	0	0	0	0
Week 2	$100(1\text{-}p_1)$	$100p_1$	0	0	$100(1\text{-}p_1)$	0	0	0
Week 3	$100p_1(1\text{-}p_2)$ + $100(1\text{-}p_1)^2$	100 $(1\text{-}p_1)p_1$	$100p_1p_2$	0	$100(1\text{-}p_1)$ + $100(1\text{-}p_1)^2$	$100p_1$ $(1\text{-}p_2)$	0	0

The following summarize the stochastic process (for convenience, we ignore expected value signs and use a general "*n*" rather than the 100 total for our Billboard Chart):

$$\sum_{i=1}^{k} C_{k,i} = n$$

(In each week, *k*, there must be $n = 100$ albums on the chart so summing across various weeks on the chart, from 1 week to *k* weeks, must yield 100 albums.)

$$TD_{k-1,w} = \sum_{j=1}^{k-1} D_{j,w}$$

(At the end of the *k*-1 chart, this is the total number of albums that were on the chart exactly *w* weeks before falling off.)

$$\sum_{w=1}^{k-1} D_{k-1,w} = C_{k,1}$$

(The number of new albums coming onto the *k*th chart must be equal to the number of albums that fell off the charts after week *k*-1; we sum across those that were on for 1 week, 2 weeks, up to *k*-1 weeks to get the total number that dropped off after the *k*-1 chart.)

$$C_{k,i} = p_{i-1} C_{k-1,i-1}$$

(*Note expected value operators are not shown.) (Expected number of albums that appear on chart *k* and have then survived for *i* weeks. Value is obtained multiplying the eligible albums (those which have been on the chart *i*-1 weeks) and the probability of remaining on the chart for an *i*th week given album was on the chart *i*-1 weeks.)

$$D_{k-1,w} = (1 - p_w) C_{k-1,w}$$

(Expected number of albums that drop off after week *k*-1 having been on the charts for *w* weeks is obtained by multiplying the probability of dropping off given that the album has been on the charts *w* weeks times the number of eligible albums, those that have been on the charts *w* weeks in the *k*-1 chart.)

Let T_k be the total number of music albums that had appeared on the chart at the end of week k. The steady state ($TD_{k,w}$ / T_k) for this stochastic model can be shown to be (see Appendix):

$$\lim_{k\to\infty}\frac{TD_{k,w}}{T_k} = (1-p_w)\prod_{j=1}^{w-1} p_j$$

Note that when $p_i = p$ for all i, the steady state is that of the geometric distribution (see similar distributions, e.g., Chung and Cox (1994), going back to Simon (1955) and Yule (1924)).

For $k > i$, by expansion, we have $C_{k,i} = p_{i-1}, p_{i-2}, \ldots, p_2, p_1, C_{k-i,1}$. At one extreme, it is possible that all p_i's are equal, that is, that the probability of an album remaining on the Billboard Chart is independent of the number of weeks the album has already been on the chart. At the other extreme, all p_i values could be different. From empirical observations, we choose a step function for the p_i values as explained below. It is consistently the case (see below) that the largest "falling off the chart" occurs for albums that have been on the chart just one week. In addition, there appears to be at least one clear "shift" point. After albums have been on the chart for some number of weeks, the probability of remaining on the chart shifts upward. As an example, for three shift points (four "p"s), our model would utilize p_i values as follows:

$$p_1 < p_2 = p_3 = \ldots = p_a < p_{a+1} = p_{a+2} = \ldots = p_b < p_{b+1} = p_{b+2} = \ldots = p_{k-1}.$$

Data and Stochastic Model Estimation

Our Billboard Chart data includes all weekly data over the periods 1995–1997 and 2000–2002, the pre- and post-change periods in the markets. We investigate whether the market landscape has shifted and, if so, what the implications are for music firms. We note that the data observations are not a random sample and, in reality, are the entire populations for the two periods studied. We view them as all realizations from a stochastic process for the selected periods.

Preliminary evaluation of the data and discussions with individuals knowledgeable about the industry suggest that the album "chart drop-off process" is quite rapid. During the years studied, while one album did in fact remain on the chart for 151 weeks, the vast majority of albums had much shorter chart life spans. Table 8.2 summarizes the number and percentage of albums that debuted in a given year and the number of weeks they remained on the chart before departing.

Table 8.2 Album Dropoff Behavior on Charts

	Year of Debut					
	1995	1996	1997	2000	2001	2002
Number of albums debuting on the Billboard 100 during the year	323	339	361	341	366	383
Total number of albums dropping off after 1 week	43 (13.3%)	41 (12.1%)	55 (15.2%)	86 (25.2%)	91 (24.9%)	91 (23.8%)
Total number of albums dropping off after weeks 1 through 4	122 (37.8%)	119 (35.1%)	120 (33.2%)	190 (55.7%)	189 (51.6%)	197 (51.4%)
Total number of albums dropping off after weeks 1 through 8	162 (50.2%)	169 (49.9%)	189 (52.4%)	252 (73.9%)	262 (71.6%)	282 (73.6%)
Total number of albums dropping off after weeks 1 through 13 (3 months)	205 (63.5%)	222 (65.5%)	234 (64.8%)	284 (83.3%)	310 (84.7%)	331 (86.4%)
Total number of albums dropping off after weeks 1 through 20	247 (76.5%)	267 (78.7%)	282 (78.1%)	290 (85.0%)	324 (88.5%)	344 (89.8%)

Since the majority of albums dropped off the chart within the first three months, we decided to focus on modeling and estimating a stochastic process of that length.[5] As outlined in the previous section, the family of stochastic processes we are utilizing includes an array of shift points from 1 to 13. That is, one case would be where the probability of falling off the chart remains the same no matter the number of weeks the album has appeared on the chart. The other extreme would be 13 shift points where the probability is different for each of the 13 possible weeks an album could have remained on the chart.

We used a brute force solution process beginning with one p value. Table 8.3 summarizes the outcomes for the single p stochastic process.

Table 8.3 Single p Stochastic Process Estimates

	1995	1996	1997	2000	2001	2002
$\hat{p}$	0.83	0.85	0.84	0.76	0.79	0.80
Computed χ^2	23.47	16.62	29.29	13.89	28.14	32.93
α*	0.0240	0.1646	0.0036	0.3077	0.0053	0.0010

Table 8.4 Multiple p Stochastic Process Estimates

	1995	1996	1997	2000	2001	2002
$\hat{p}_1$	0.86	0.89	0.86	0.72	0.74	0.76
$\hat{p}_2$	0.92	0.90	0.91	0.82	0.85	0.85
Weeks: $\hat{p}_1$	2	3	2	1	1	1
Computed χ^2	11.16	7.99	17.60	7.52	8.16	14.46
α*	0.4301	0.7138	0.0913	0.7556	0.6993	0.2087

We used the χ^2 goodness of fit test for the null hypothesis that the stochastic model is appropriate for the observed process. The α* values (normally indicated as p values, but we use α* here to avoid any confusion) indicate that level of significance at which we would begin rejecting the null hypothesis. That is, we only reject the null hypothesis for a level of significance greater than α*. Thus, at a 0.05 level of significance we would reject the proposed stochastic model for years 1995, 1997, 2001, and 2002. We would accept the null hypothesis (the proposed model) for years 1996 and 2000.

We then repeated the brute force solution process for a model with one shift point (two "p"s). The results are summarized in Table 8.4. This time, using the χ^2 goodness of fit test and a 0.05 level of significance, we would accept the null hypothesis of model appropriateness for all years. The α* levels ranged were quite high: 0.4301 (1995), 0.7138 (1996), 0.0913 (1997), 0.7556 (2000), 0.6993 (2001), and 0.2087 (2002). Repeating the process for two shift points (three "p"s), we found little improvement and thus focus on p_1, p_2 stochastic model with parameter estimates as indicated in Table 8.4 and illustrated in Fig. 8.1.

Consider the model specification differences in the periods before (1995–1997) and after (2000–2002). We note the following:

1. the shift period occurs earlier (the $\hat{p}_1$ estimate has only a 1-week duration in each of the 2000–2002 years compared to 2- or 3-week duration in each of the 1995–1997 years); and

2. in every year during the 2000–2002 period, the $\hat{p}_1$ and $\hat{p}_2$ values are less than the corresponding $\hat{p}_1$ and $\hat{p}_2$ values for each year in the 1995–1997 period.

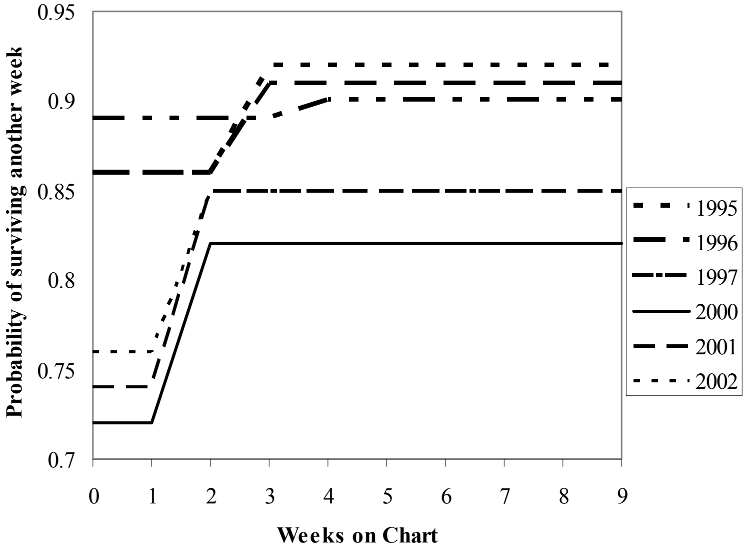
The 2000–2002 values of $\hat{p}_1$ are 0.72, 0.74, and 0.76, respectively, compared to $\hat{p}_1$ values for 1995–1997 of 0.86, 0.89, and 0.88. The 2000–2002 values of $\hat{p}_2$ are 0.82, 0.85, and 0.85, respectively, compared to $\hat{p}_2$ values for 1995–1997 of 0.92, 0.90, and 0.91. These outcomes suggest quite different parameters for our stochastic model before and after the 1998–1999 market shift. The probability of remaining on the chart after 1 week fell by an average of 0.1 (0.84 before and 0.74 after). Further, the probability of remaining on the chart (after the process shift) was, on average, 0.07 lower (0.91 before and 0.84 after). Table 8.5 provides the estimated probabilities of remaining on the chart for a set number of weeks for each of the years.

Table 8.5 Estimated Probability of Survival on Chart

Weeks	1995	1996	1997	2000	2001	2002
2	0.86	0.89	0.86	0.72	0.74	0.76
3	0.74	0.79	0.74	0.59	0.63	0.65
4	0.68	0.70	0.67	0.49	0.53	0.55
5	0.63	0.63	0.61	0.40	0.45	0.47
6	0.58	0.57	0.56	0.33	0.39	0.40
7	0.53	0.51	0.51	0.27	0.33	0.34
8	0.49	0.46	0.46	0.22	0.28	0.29
9	0.45	0.42	0.42	0.18	0.24	0.24
10	0.41	0.37	0.38	0.15	0.20	0.21
11	0.38	0.34	0.35	0.12	0.17	0.18
12	0.35	0.30	0.32	0.10	0.15	0.15
13	0.32	0.27	0.29	0.08	0.12	0.13

The results provided in Tables 8.2, 8.4, and 8.5 and Graph 8.1 indicate a shift in chart life cycles of music albums following the technological and market innovations of 1998–1999. In the 1995–1997 period, 50% of the albums that appeared on the chart would be expected to last at least 7 weeks and at least 40% would be expected to last 9 weeks. In the 2000–2002 period, less than 50% would be expected to last 5 weeks. Less than 40% would be expected to make it to the sixth week. It is important to note that although the probability of remaining on the chart has changed before and after the 1998–1999 period, the structural robustness of the model with one shift point (two "*p*"s) is maintained in both periods.

The nature of the estimated models indicates a shift after the technological innovations of MP3 and online file-sharing that occurred over the 1998–1999 period. The 2000–2002 period is characterized by a much shorter life cycle. In our stochastic life cycle analysis, we utilized the entire set of outcomes (Billboard Top 100 rankings) for the comparison periods: 1995–1997 and 2000–2002. We find that music as a digital good has been significantly impacted by market changes brought about by easier information dissemination and product access to potential consumers. While overall album survival has decreased in the 2000–2002 period, the chances of survival increase dramatically after an album has survived beyond the first week during this period. This indicates a pattern that the "good" albums survive more. This also suggests that the new environment brought on by technological and other market innovations is not conducive to lower quality music albums. Easier sampling and information dissemination hurts the lower quality albums. In general, the life cycle of lower quality products will tend to diminish faster under this new environment. The analysis also suggests that albums face a shorter life cycle overall.



Graph 8.1 Multiple p Stochastic Process

Our analysis in this phase also emphasizes the robustness of our stochastic life cycle model and consistency of results within each of the 3-year periods, together with the differences between the two periods. Even with significant "churn" in the music market and related environment between the 1995–1997 and 2000–2002 periods, our simple yet robust model effectively captures the stochastic component of the chart life cycle process.

Duration Model of Album Survival

The stochastic model above does not easily yield to analyzing various exogenous factors that affect the survival probability of an album. For that, we turn to duration models in economics which allow for "regression"-like approach. The underlying process of survival is a weekly stochastic process that governs whether an album exits the charts (hazard of an album exiting the charts). Thus, in a hazard framework, one assesses the impact of an explanatory variable on the hazard of exiting the chart rather than on the length of survival time. A proportional hazard model (PHM) would allow for including various exogenous regressors and estimate the probability of an album exiting the chart given a certain number of weeks already on the chart. Thus PHM specification is of the form

$$h_1(t_i, X) = h_0(t_i)\exp(X_i\,\beta) \tag{1}$$

$h_0(t)$ is the baseline hazard function. A hazard function is simply $f(t)/(1\text{-}F(t))$ where $f(t)$ is the probability distribution and function and $F(t)$ is the cumulative distribution function. $(1\text{-}F(t))$ is also referred to as survival function. While a Cox PHM specification uses non-parametric form for baseline hazard function (Cox 1972), a Weibull PHM specification employs a parametric form to estimate this hazard. X is a set of covariates which shift the hazard function proportionally and β are the parameters to be estimated. A more "regression"-like framework is Accelerated Failure time (AFT) model. In AFT one can write the survival duration on Billboard chart as a regression model

$$\ln(T_i) = X_i\gamma + \varepsilon_i \tag{2}$$

The difference being that in (1), β is interpreted as affecting the hazard rate, while in (2) γ is interpreted as affecting the log of duration. Weibull yields to both PHM and AFT specifications, while Cox only admits the PHM specification. In case of Weibull, the error term $\varepsilon_i = \ln(\sigma u_i)$. This leads to ε_i having an extreme value distribution whose variance σ is to be estimated.

PHM models also allow for controls for unobserved heterogeneity (similar to random effect models in regression). In particular, in continuous time PHM models, data could be dispersed and not controlling for such heterogeneity may produce incorrect estimates. To incorporate unobserved heterogeneity, we modify (1) such that

$$h_1(t_i, X) = h_0(t_i)\exp(X_i\beta + \nu) \tag{3}$$

where v is gamma distributed with mean 0 and variance σ^2 which can be estimated.

Data Set

We used the same data as described earlier. Thus our Billboard Chart data includes all weekly data over the periods 1995–1997 and 2000–2002, the pre- and post-change periods in the markets. In addition, we collected data on the survival model explanatory variables (X_i) which we operationalize as follows.

Survival: Number of weeks an album appears on the Billboard top 100 charts. On occasion, an album may drop off for some weeks and reappear again on the chart. Each album is continuously tracked till its final drop-off. As noted earlier, our data does not suffer from left or right data censoring issues, as we track each album from its chart debut (birth) until its final drop off (death) from the charts. Note that the drop-off may occur well beyond the 34 weeks of each time segment.

Debut rank: The rank at which an album debuts on the Billboard top 100 chart. Numerically higher ranked albums are less popular.

Debut post-TS: This is an indicator variable which is 0 for albums that debut from 1995 to 1997 and 1 for albums that debut from 2000 to 2002. This dummy captures the effect of technological changes on album survival.

Superstar: A binary variable denoting the reputation of the artist. If a given album's artist has previously appeared on the Billboard top 100 charts for at least 100 weeks (on or after January 1, 1991) prior to the current album's debut, then the variable is set to 1, otherwise 0.

Minor label: A binary variable that is set to 0 if the distributing label for a given album is one of Universal Music, EMI, Warner, or SONY-BMG. A value of 1 denotes independent and smaller music labels.

Solo male: A binary variable that denotes if an album's artist is a solo male (e.g., Eric Clapton).

Solo female: A binary variable that denotes if an album's artist is a solo female (e.g., Britney Spears).

Group: A binary variable that denotes if an album's artist is a group (male or female) (e.g., U2, The Bangles).

Holiday month debut: To control for the holiday effect (or “Christmas effect”), we include an indicator variable for December, which is 1 if album debuted in that month and 0 otherwise.

Number of albums released per year: Album survival may depend on the number of albums released per year.

Estimation and Results

We estimate model (3) using Weibull distribution and gamma frailty. The results are consistent with Cox Models as well.

Table 8.6 Estimates for Album Survival

Parameter	Weibull PHM (Hazard Ratio)
Debut rank	1.09** (14.8)
Debut post-TS	2.80** (5.7)
No of albums released	1.00 (1.3)
Superstar	0.20** (7.81)
Minor label	1.75** (3.03)
Solo male	3.07** (4.45)
Group	4.80** (6.22)
Holiday month debut	0.52** (2.83)
Frailty variance σ	3.52** (14.6)
Weibull shape parameter	3.62** (21.3)
	LL = -2014

We report hazard ratios that are easier to interpret. A hazard ratio >1 indicates that the variable increases the hazard rate and vice-versa. Thus a hazard ratio of 1.09 for debut rank means that each increase in debut rank, on average, increases the hazard rate by 9%. Thus an album debuting at higher numerical rank will exit the charts faster than the album debuting at lower numerical rank. Except the no_of_albums variable, all other variables are significant and in expected direction. Notice that in the post period, the hazard rate has gone up by as much as 180%. We can interpret the hazard ratios to affect the total duration as well (recall that Weibull model yields to both PHM and AFT specifications). Thus, we can calculate that in the post period, albums’ survival, on average, has decreased by 42%. Thus they survive only 5.8 weeks now if they survived for 10 weeks earlier.

Superstar effect is also quite strong. Controlling for debut, superstars tend to survive longer on charts than otherwise. In particular, at any time, non-superstars’ hazard of exiting the charts is 80% higher than superstars.

Minor labels do worse than major labels and surprisingly, both males and groups perform worse than females on Billboard 100 charts (at least in terms of survival). Albums debuting in the month of December tend to survive longer.

High frailty variance (variance of gamma distribution) indicates the importance of including unobserved heterogeneity.

We plot the hazard rate and survival rate for albums in post and pre period. Similarly, we plot the hazard rate and survival rate for albums with superstar to get more insight.

In Fig. 8.1a, we plot estimated hazard function for before and after data. Similarly, in Fig. 8.1b, we plot predicted hazard function for superstar and non-superstar. First note that hazard function is non-monotonic. As we saw in the previous section, the hazard is very high during the first 2–3 weeks (In short, many albums exit the Billboard chart within couple of weeks). However, past 3 weeks, hazard is decreasing. Thus once the album survives the first few weeks, it has a lower probability of exiting (or high probability of surviving longer). Also, note that the hazard functions are not proportional. It is because of the unobserved heterogeneity (gamma distribution) we introduced.

As our results indicated, hazard rate has increased significantly for albums in the post period (2000–2002) compared to pre period (1997–1999). Similarly, superstars have less hazard of exiting the Billboard 100.

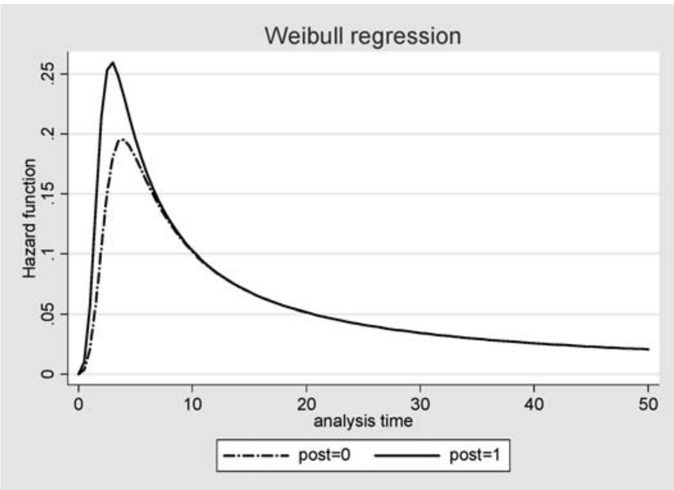


Fig. 8.1a Hazard rate in the Pre and Post periods

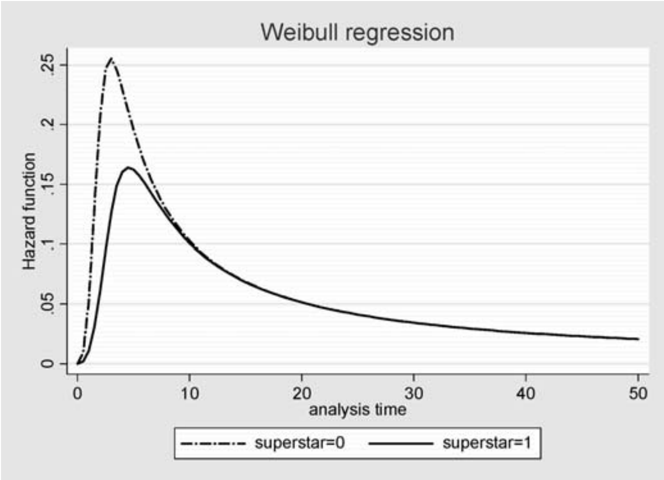


Fig. 8.1b Hazard rates for superstars and non-superstars

Similarly, in Fig. 8.2, we plot the predicted survival conditioned on exogenous factors (namely pre and post period and superstar or non-superstar). Notice that there is less than 20% chance of an album surviving beyond 10 weeks and once the albums survive 10 weeks, its survival rate more or less remains unchanged. Also, as the hazard graph indicated, survival rate is higher for superstars and survival rate in the post period has gone down significantly.

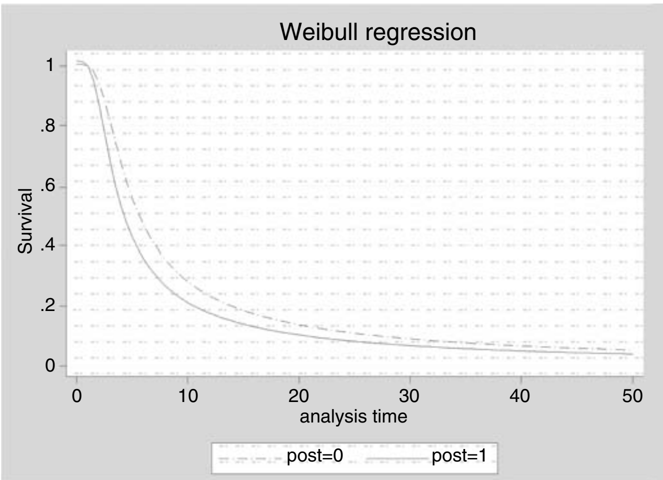


Fig. 8.2a Survival rates for albums in Pre and Post Period

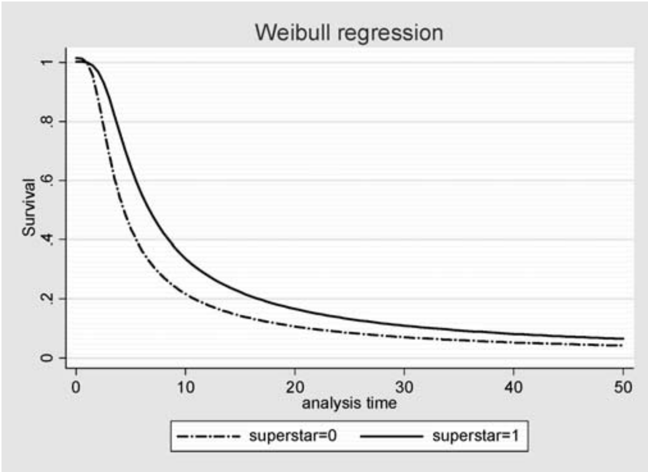


Fig. 8.2a Survival rates for superstars and non-superstars

Discussion and Conclusion

Technological advances can have significant impacts on economic markets. We have analyzed the impacts of the technological innovations of MP3 and online file-sharing on the music industry landscape. Our initial analysis focused on the development of a robust stochastic model to capture the overall dynamics of the music album life cycle on the Billboard Chart. The subsequent analysis focused on the exogenous album-related factors that impact the survivability of the Billboard Chart. In both analyses, the overall objective was to discern whether and how the recent technological innovations have fundamentally altered the music industry.

Following earlier work related to the markets for artists, we developed a stochastic process model of the life cycle of albums. Brute-force estimation yielded excellent fits for all years. The nature of the estimated models indicates a shift after the technological innovations of MP3 and online file-sharing that occurred over the 1998–1999 period. The 2000–2002 period is characterized by a much shorter life cycle. We find that music as a digital good has been significantly impacted by market changes brought about by easier information dissemination and product access to potential consumers. While overall album survival has decreased in the 2000–2002 period, the chances of survival increase dramatically after an album has survived beyond the first week during this period. This indicates a pattern that the "good" albums survive more. This also suggests that the new environment brought on by technological and other market innovations is not conducive to lower quality music albums. Easier sampling and information dissemination hurts the lower quality albums. In general, the life cycle of lower quality products will tend to diminish faster under this new environment. The analysis also suggests that albums face a shorter life cycle overall. Even with significant “churn” in the music market and related environment between the 1995–1997 and 2000–2002 periods, our simple yet robust model effectively captures the stochastic component of the chart life cycle process.

Our subsequent analysis sheds light on the exogenous factors that impact album survival and the shifts in the patterns of impact since the technological changes. We find that albums’ survival has decreased significantly in post-period. We also find that superstar effect is quite strong. Superstars’ albums tend to survive longer. Albums promoted by major labels and by females survive longer. Similarly, albums released in the month of December tend to survive longer.

In the face of file sharing networks that enable widespread sharing and downloading of music in digital forms, music companies have felt pressured to take steps to simultaneously safeguard their digital products and bolster their market performance. Their strategic decisions and actions have thus far focused on incorporating security mechanisms in the digital products themselves and on legal threats and actions against both operators of file sharing networks and individual file sharers (Bhattacharjee et al. 2006a,b). As illustrated by the Sony BMG situation earlier, embedded security measures can frustrate consumers and have significant negative impacts. Further, no security measure used by the entertainment industry so far has been foolproof. As Bhattacharjee et al. (2006a,b) detail, legal threats and actions have reduced sharing by individuals, but significant piracy opportunities remain. Further, such actions have been industry actions rather than individual firm actions.

Our focus has been on modeling life cycle on the charts and how it has been affected by significant changes in the landscape of the music market. A key finding is that the market landscape has shifted and that life cycle has shortened with lowered probabilities of surviving for each subsequent week on the chart. The significantly shorter shelf life of digital music calls for accelerated tactical and operational decision-making on resource allocations, in particular marketing and promotional efforts that target potential winners. In the latter period (2000–2002), the likelihood of surviving another week falls below one-half by the fifth week while this doesn’t occur until the eighth week in the earlier period (1995–1997). Hence music companies may well opt to move promotional efforts earlier in the cycle. Interestingly, while the landscape has shifted, the underlying drivers that govern the life cycle process appear to have remained steady. That is, even with significant “churn” in the music market and related environment between the 1995–1997 and 2000–2002 periods, our underlying model form is robust and succinctly captures the life cycle process for the entire duration. Thus the same underlying decision models, where the parameters are constantly monitored and re-estimated, would provide a music firm with a reliable benchmark to gauge and assess their suite of the music albums in the marketplace, and make better decisions in an uncertain environment.[6]

Appendix: Steady State Characterization of the Stochastic Process

T_k, the total number of music albums that have appeared on the chart at the end of week k, can be expressed as

$$T_k = \sum_{m=1}^{k} C_{m,1} \tag{4}$$

$TD_{k,w}$, the total number of albums that were on the chart for exactly w weeks before falling off the charts at the end of week k, can be expressed recursively as

$$TD_{k,w} = TD_{k-1,w} + (1 - p_w) C_{k,w} \tag{5}$$

Note that $TD_{w,w-j} = 0 \quad \forall j \geq 1$. Therefore (5) can be expressed as

$$TD_{k,w} = (1 - p_w) \sum_{m=1}^{k-w} C_{m+w,w} + TD_{w,w} \tag{6}$$

Further, we have

$$TD_{w,w} = (1 - p_w) C_{1,1} \prod_{j=1}^{w-1} p_j \tag{7}$$

$$C_{m+w,w} = C_{m+1,1} \prod_{j=1}^{w-1} p_j \tag{8}$$

Using (7) and (8), (6) can be expressed as

$$TD_{k,w} = (1 - p_w) \prod_{j=1}^{w-1} p_j \sum_{m=1}^{k-w} C_{m+1,1} + (1 - p_w) \prod_{j=1}^{w-1} p_j C_{1,1} \tag{9}$$

Simplifying, we obtain

$$TD_{k,w} = (1 - p_w) \prod_{j=1}^{w-1} p_j \sum_{m=1}^{k-w+1} C_{m,1} \tag{10}$$

From (4) and (10), we have

$$\frac{TD_{k,w}}{T_k} = (1 - p_w)\prod_{j=1}^{w-1} p_j \left(\frac{\sum_{m=1}^{k-w+1} C_{m,1}}{\sum_{m=1}^{k} C_{m,1}} \right) \tag{11}$$

For finite values of w,

$$\lim_{k\to\infty}\left(\frac{TD_{k,w}}{T_k}\right) = (1 - p_w)\prod_{j=1}^{w-1} p_j \tag{12}$$

which is independent of k, thus yielding our steady state.

Notes

The authors gratefully acknowledge the Treibick Family Endowed Chair, the Treibick Electronic Commerce Initiative, the XEROX CITI Endowment Fund, the GE Endowed Professor Fund, The Center for Internet Data and Research Intelligence Services (CIDRIS), and the Gladstein Endowed MIS Research Lab for support that made this work possible. Rahul Telang acknowledges the generous financial support of the National Science Foundation (NSF) through the CAREER award CNS-0546009.

1. MP3 is a commonly used audio compression technology.
2. Recently, legal threats from RIAA may be changing the landscape a bit (e.g., see Bhattacharjee et al. (2006)).
3. The top 100 albums per week are available free at http://www.billboard.com/bbcom/charts/chart_display.jsp?f=The+Billboard+200&pageNumber=Top+1-10&g=Albums
4. Our empirical data show that the probability of an album re-appearing on the chart is minimal.
5. This also avoided the inclusion of periods with expected frequencies less than 5, a consideration when we analyze the appropriateness of our stochastic model.
6. Sudip Bhattacharjee is an Associate Professor and Ackerman Scholar in the Department of Operations and Information Management in the School of Business, University of Connecticut. Email: Sudip.Bhattacharjee@business.uconn.edu. Ram D. Gopal is GE Endowed Professor of Business in the Department of Operations and Information Management in the School of Business, University of Connecticut. Email: Ram.Gopal@business.uconn.edu. James R. Marsden is the Treibick Family Endowed Chair in e-Business and Board of Trustees Distinguished Professor. Email: Jim.Marsden@business.uconn.edu. Rahul Telang is an Assistant Professor of Information Systems and Management at the Heinz School, Carnegie Mellon University. Email: rtelang@andrew.cmu.edu.

References

Adler, Moshe, "Stardom and Talent," *American Economic Review*. 75 (March 1985) 208–212.

Albert, Steven, "Movie Stars and the Distribution of Financially Successful Films in the Motion Picture Industry," *Journal of Cultural Economics*. 22 (1998) 249–270.

Associated Press, "Norwegian hacker cracks iTunes code," *CNN*, November 27, 2003, http://www.cnn.com/2003/TECH/internet/11/27/itunes.code.ap/index.html

Bergstein, B., "Copy Protection Still a Work in Progress," *Associated Press*, http://news.yahoo.com/s/ap/20051119/ap_on_hi_te/music_copy_protection;_ylt=ArOQdPcARo.I0hnGv20mpcdk24cA;_ylu=X3oDMTBidHQxYjh2BHNlYwN5bnN0b3J5, Nov 18, 2005.

Bhattacharjee, S., R.D. Gopal, K. Lertwachara, and J.R. Marsden, "Impact of Legal Threats on Individual Behavior: An Analysis of Music Industry Actions and Online Music Sharing," *Journal of Law and Economics*. 49 (1) (April 2006) 91–114.

Bhattacharjee, S., R.D. Gopal, K. Lertwachara, and J.R. Marsden, "Whatever Happened To Payola? An Empirical Analysis of Online Music Sharing," *Decision Support Systems*. 42 (1) (2006) 104–120.

Bradley, M., "Sony Aims at Pirates – and Hits Users," *Christian Science Monitor*, Nov 9, 2005, http://www.csmonitor.com/2005/1109/p14s01-stct.html

Bradlow, Eric T. and Peter S. Fader, "A Bayesian Lifetime Model for the Hot 100 Billboard Songs," *The Journal of the American Statistical Association.* 96 (2001) 368–381.

Chung, Kee H. and Raymond A.K. Cox, "A Stochastic Model of Superstardom: An Application of the Yule Distribution," *Review of Economics and Statistics*. 76 (4) (November 1994) 771–775.

Clarke, G., "DVD Jon Hacks Media Player File Encryption," *The Register*, September 2, 2005, http://www.theregister.co.uk/2005/09/02/dvd_jon_mediaplayer/

Cox, A.K. and Kee H. Chung, "Patterns of Research Output and Author Concentration in the Economics Literature," *Review of Economics and Statistics*. 73 (4) (Nov. 1991) 740–747.

Cox, D.R., "Regression Models and Life Tables," *Journal of the Royal Statistical Society Series B.* 34 (1972) 187–220.

Craver, S.A., Min Wu, Bede Liu, Adam Stubblefield, Ben Swartzlander, Dan W. Wallach, Drew Dean, and Edward W. Felten, "Reading Between the Lines: Lessons from the SDMI Challenge," *Proc. of 10th USENIX Security Symposium* (August 2001).

Felten, E.W., "A Skeptical View of DRM and Fair Use," *Communications of the ACM*. 46 (4) (April 2003) 56–61.

Felten, E.W., "Inside RISKS: DRM and Public Policy," *Communications of the ACM.* 48 (7) (July 2005) 112.

Goodley, S., "Disharmony over music pirates on the Internet," *The Telegraph*, January 9, 2003, http://www.telegraph.co.uk

Halderman, J.A., "Evaluating New Copy-Prevention Techniques for Audio CDs," *Proc. ACM Workshop on Digital Rights Management*, Washington, DC, November 2002.

Levene, Mark, Trevor Fenner, Geroge Loizou, and Richard Wheeldon, "A Stochastic Model for the Evolution of the Web," *Computer Networks.* 39 (2002) 277–287.

Patrizio, A., “Why the DVD Hack Was a Cinch,” *Wired News*, November 2, 1999, http://www.wired.com/news/technology/0,1282,32263,00.html

Reuters, “CD Crack: Magic Marker Indeed,” May 20, 2002, http://www.wired.com/news/technology/0,1282,52665,00.html

Reuters, “Sony Tests Technology to Limit CD Burning,” *CNET News.* June 1, 2005, http://news.cnet.co.uk/digitalmusic/0,39029666,39189658,00.htm

Rosen, Sherwin, “The Economics of Superstars,” *American Economic Review*. 71 (December 1981) 845–858.

Russinovich, M., “Sony, Rootkits and Digital Rights Management Gone Too Far,” http://www.sysinternals.com/blog/2005/10/sony-rootkits-and-digital-rights.html, *Mark’s Sysinternals Blog*, October 31, 2005.

Schneier, B., *Secrets and Lies: Digital Security in a Networked World*. John Wiley & Sons, Inc., New York, 2000.

Simon, Herbert A., “On a Class of Skew Distribution Functions,” *Biometrika.* 42 (1955) 425–440.

Smith, E., “Sony BMG Faces Civil Complaint Over CD Software,” http://online.wsj.com/article/SB113259581938503230.html?mod=mm_hs_entertainment, *The Wall Street Journal*. (November 22, 2005).

Yule, G. Udny, “A Mathematical Theory of Evolution, based on the Conclusions of Dr. J.C. Willis, F.R.S.,” *Philosophical Transactions of the Royal Society B*. 213 (1924) 21–87.

Ziemann, George, “RIAA’s Statistics Don’t Add Up to Piracy,” December 11, 2002, http://www.azoz.com/music/features/0008.htm