LISR 10, 237–255 (1988)

Modeling Arrival Patterns of Library Book Orders

MICHAEL D. COOPER

School of Library and Information Studies
University of California
Berkeley, CA 94720

Book orders that libraries place to vendors may or may not be filled depending on the availability of the material, the quality of the vendor, and a host of other factors. This paper analyzes a sample of more than 4,700 orders for academic materials in terms of the probability that the material will arrive and the factors that determine arrival. Product limit estimates and accelerated failure time models were used to explain variations in arrival time.

Thirty six percent of the orders analyzed required claiming. There were significant differences in vendor performance, with a low of 15% of orders to one vendor versus 74% to another needing claiming. Delivery times for materials varied significantly by vendor as well, with some vendors supplying materials in a mean of 82 days while others required a mean of 205 days.

The accelerated failure time model analyzed the number of days from order to receipt of an item. This was done in terms of the interaction between vendor; type of material; year of publication; and whether it was ordered rush, found in *Books in Print* and/or claimed. In comparison to a reference group, there were significant variations in the performance of some vendors, as well as better delivery times for rush and reprint items. Delivery time was not influenced by year of publication of the item ordered.

1.0 INTRODUCTION

In "Random Vendor Assignment in Vendor Performance Evaluation," Barker (1986) described an experiment conducted at the University of California, Berkeley General Library Acquisitions Department in which orders for library materials were randomly assigned to vendors to test vendor performance. This paper reanalyzes the data developed by Barker using product limit survival estimates and accelerated failure time models. With these

The data used in this paper was provided by Joseph Barker, head of the Acquisitions Department of the General Library of the University of California, Berkeley. Mary Heath of the University of California Library Systems Office preprocessed it for my use. The referees made valuable suggestions for improving the paper.

Original Manuscript Submitted November 1987; Manuscript Accepted March 1988.

models, acquisitions processing is viewed as a birth–death process in which orders are placed (birth), some are fulfilled (death), and some are never completed. Orders that are incomplete at the end of the experimental period are said to be *right-censored,* that is, there is no known arrival date for them. The product limit models differ from the accelerated failure time models in that the former make no assumptions about the probability distribution of arrival time. The accelerated failure time models are a generalization of multiple regression techniques to handle censored date.

1.1 The Barker Experiment

The methodology used to gather data on vendor performance is described in Barker (1986). It is briefly reviewed here to provide a context for the modeling. A set of independent variables were used to describe each item ordered by the University of California, Berkeley General Library during the experimental period. These include the language of publication, location of the publisher, whether or not the item was listed in *Books in Print* (*BIP*), whether the item was ordered "Rush," and the type of material ordered (e.g., art book, document, reprint). Items were randomly assigned to vendors with orders for multiple copies of the same item treated as single copy orders to multiple vendors. As orders were claimed and then received, the order database was updated. Three hundred days after the last order was placed recording was ended.

2.0 DESCRIPTIVE DATA ABOUT THE ORDERING PROCESS

The performance of 16 vendors was analyzed for this experiment. Six of the vendors were located in Western Europe. Four of the six were used to supply French-language material, and the other two supplied British imprints. Two of the U.S. vendors were located in the western part of the country, and four were in the east. The amount of material ordered from each vendor is roughly bimodal: either the library ordered about 4 or about 9% of the total (close to 4,700 items) from each vendor (Table I).Only about 16% of the items ordered were not in English.

Materials ordered were classified by category. These categories indicate whether the item was listed in *BIP;* whether it was ordered for rush delivery; whether it was an art book, a reprint, or of British, Canadian, or French imprint.[1] The breakdown of materials ordered from each vendor is presented in Table II.[2] Close to half the items were found in *BIP* and more than 16% of them were French.

[1] The original data included orders to a number of other vendors and classified materials in additional ways. Vendors with low use frequencies were dropped from the analysis and some categories were eliminated or collapsed into others.

[2] The difference in the total number of orders in the tables is due to missing values for some of the classificatory variables.

TABLE I
Number of Items Ordered from Vendors by Language of Publication

Vendor	Language English	French	Other	Total Orders	% Distribution
Location—United States					
Academic Book Center, OR	397	0	1	398	8.4
Ambassador Book Service, NY	227	0	0	227	4.8
Blackwell North America, OR	450	1	1	452	9.6
The Book House, MI	475	0	1	476	10.1
Book Service International, CT	429	1	1	431	9.1
Coutts Library Service, NY	423	0	0	423	9.0
Midwest Library Service, MO	382	0	0	382	8.1
Yankee Book Peddler, NH	172	0	0	172	3.6
Roy Young Booksellers, NY	373	0	1	374	7.9
Location—Europe					
Aux Amateurs de Livres, Paris	11	182	1	194	4.1
B.H. Blackwell, Oxford	216	0	1	217	4.6
Otto Harrassowitz, Wiesbaden	10	180	4	194	4.1
Martinus Nijhoff, The Hague Neth.	1	182	2	185	3.9
B.F. Stevens & Brown, Surry England	214	0	1	215	4.5
Jean Touzot Librairie, Paris	7	177	2	186	3.9
Hubert Wilson, Margate England	200	0	1	201	4.3
TOTAL	3987	723	17	4727	
% Distribution	84.3	15.3	0.4		100

The ordering covered by the survey is for scholarly materials that are not generally published by major publishers. For example, in the sample of 4,822 orders, 106 of them had the John Wiley imprint. This was the *largest* number of titles from a single publisher in the sample! A large proportion of Berkeley's acquisitions (not covered in this analysis) are through blanket and standing orders. In these orders, the distribution is entirely different.

2.1 Characteristics of Materials Claimed

Orders that do not arrive within a specific period of time are claimed by sending an inquiry to the vendor about the item's status. When orders are placed, the Berkeley Acquisitions Department determines dates when claiming should take place. These dates vary by category of material and location of vendor.[3] The dates serve as guidelines for claiming action. If the department receives status information about an item before it is time to claim, the claim may not be sent.[4]

[3] Orders to U.S. and British vendors are claimed 16, 32, and 48 weeks from order date. Orders for "Rush" items from U.S. vendors are claimed 8, 16, and 24 weeks from order date, and orders from French vendors have a 24, 40, and 56-week claiming cycle.

[4] To a certain extent, this introduces a bias in the analysis since vendors who supply frequent status reports may not receive as many claims. Too many status reports from vendors might simply be ignored in claiming decisions because of the time to review them.

TABLE II
Number of Items Ordered from Vendors by Category of Material

Vendor	Category of Material									
	In BIP	Not in BIP	In BIP Rush	Art	British	Canadian	French	Reprint	Total Orders	% Distribution
Location—United States										
Academic Book Center, OR	288	59	26	0	1	0	0	14	388	8.4
Ambassador Book Service, NY	147	30	15	19	0	0	0	9	220	4.7
Blackwell North America, OR	294	54	24	0	0	35	0	14	421	9.1
The Book House, MI	280	56	21	40	0	35	0	14	446	9.6
Book Service International, CT	286	56	26	41	0	0	0	15	424	9.2
Coutts Library Service, NY	287	54	25	0	0	34	0	14	414	9.0
Midwest Library Service, MO	280	54	26	0	0	0	0	14	374	8.1
Yankee Book Peddler, NH	137	22	9	0	0	0	0	3	171	3.7
Roy Young Booksellers, NY	278	55	25	0	0	0	0	12	370	8.0
Location—Europe										
Aux Amateurs de Livres, Paris	0	0	0	0	0	0	194	0	194	4.2
B.H. Blackwell, Oxford	0	0	0	0	217	0	0	0	217	4.7
Otto Harrassowitz, Wiesbaden	0	0	0	0	0	0	194	0	194	4.2
Martinus Nijhoff, The Hague Neth.	0	0	0	0	0	0	185	0	185	4.0
B.F. Stevens & Brown, Surry England	0	0	0	0	215	0	0	0	215	4.7
Jean Touzot Librairie, Paris	0	0	0	0	0	0	186	0	186	4.0
Hubert Wilson, Margate England	0	0	0	0	201	0	0	0	201	4.4
TOTAL	2277	440	197	100	634	104	759	109	4620	
% Distribution	49.3	9.5	4.3	2.2	13.7	2.2	16.4	2.4		100

Tables III, IV, and V summarize the frequency of claiming by category of material, publication date and vendor. The tables reflect the actual claiming actions taken, but the pattern of claiming may be influenced by the claiming algorithm employed.

Overall, about 63–64% of all orders were never claimed. Of the 36–37% that had to be claimed, 68–69% were claimed once, 21–22% were claimed twice, and 10% were claimed three times.

TABLE III
Percent Distribution of Orders Claimed By Category of Material

Category of Material	Total Number of Orders	% Never Claimed	% Claimed	% Claimed Once	% Claimed Twice	% Claimed Three Times
In BIP	2284	72	28	66	21	13
Not in BIP	445	53	47	62	30	8
In BIP-Rush	197	43	57	75	12	13
Art	161	61	39	46	35	19
British	634	41	59	70	23	7
Canadian	104	41	59	80	18	2
French	760	74	26	83	14	3
Reprint	121	56	44	49	30	21
% Distribution	100	64	36	68	22	10
Total Number of Orders	4706	3002	1704	1165	371	168

Note: Columns 5, 6, and 7 give the % distribution of the value in column 4. For example, of the 36% of the art book orders that were claimed, 46% of the 39% were claimed once.

TABLE IV
Percent Distribution of Orders Claimed by Publication Date of Item

Category of Material	Total Number of Orders	% Never Claimed	% Claimed	% Claimed Once	% Claimed Twice	% Claimed Three Times
Before 1970	166	58	42	56	23	21
1970–1974	101	61	39	67	15	18
1975–1979	294	57	43	65	21	14
1980	160	63	37	63	22	15
1981	179	63	37	81	13	6
1982	347	55	45	68	22	10
1983	1108	62	38	71	21	8
1984	2092	66	34	71	19	10
1985	388	63	37	53	37	10
% Distribution	100	63	37	69	21	10
Total Number of Orders	4835	3055	1780	1217	380	183

Note: Columns 5, 6, and 7 give the % distribution of the value in column 4.

TABLE V
Percent Distribution of Orders Claimed by Vendor

Vendor	Total Number of Orders	% Never Claimed	% Claimed	% Claimed Once	% Claimed Twice	% Claimed Three Times
Location—United States						
Academic Book Center, OR	398	67	33	67	25	8
Ambassador Book Center, NY	227	53	47	59	19	22
Blackwell North America, OR	452	68	32	58	31	11
The Book House, MI	476	68	32	68	22	11
Book Service International, CT	431	48	52	69	18	13
Coutts Library Service, NY	423	47	53	72	18	10
Midwest Library Service, MO	382	78	22	53	20	27
Yankee Book Peddler, NH	172	85	15	65	31	4
Roy Young Booksellers, NY	374	82	18	78	20	2
Location—Europe						
Aux Amateurs de Livres, Paris	194	85	15	73	24	3
B.H. Blackwell, Oxford	217	62	38	72	21	7
Otto Harrassowitz, Wiesbaden	194	62	38	82	14	4
Martinus Nijhoff, The Hague Neth.	185	76	24	85	13	2
B.F. Stevens & Brown, Surry England	215	33	66	78	17	5
Jean Touzot Librairie, Paris	186	73	27	88	10	2
Hubert Wilson, Margate England	201	26	74	61	30	9
% Distribution	100	63	37	69	21	10
Total Number of Orders	4727	2995	1732	1196	362	174

Note: Columns 5, 6, and 7 give the percent distribution of the value in column 4.

Table III shows that 72% of items found in *BIP* never needed to be claimed. If the items are not listed in *BIP* the percentage dropped to 53. Remarkably, French language material needed the least claiming attention

of all. Art and reprint material needed the most follow-up in terms of the number of second claims.

In comparison to the type of material claimed, the percentages for material claimed by year of publication are remarkably stable (Table IV). At this stage in the analysis there is no direct relation between the need for claiming and the year of publication. Material with older imprint dates require less claiming than newer items, perhaps because the vendor can immediately report on availability. Material published most recently (1985) required more second claims than any other category, but there were other publication years that had much higher proportions of materials claimed (e.g., before 1970, 1975–79, 1982).

There is considerable difference in the number of items claimed by vendor (Table V). For the U.S. vendors, the proportion of material never claimed ranged from a high of 85% for Yankee Book Peddler to a low of 47% for Coutts Library Service. Rankings of performance based on frequency of claiming would place Yankee Book Peddler, Roy Young Bookseller, and Midwest Library Service at the top; Coutts Library Service, Book Service International, and Ambassador Book Service at the bottom; and Academic Book Center, Blackwell North America, and The Book House in the middle range. There are, however, some interesting anomalies in the claiming distributions. While Midwest Library Service is one of the best in terms of needing little follow-up, a lot of it is needed when required (27% of all orders needing claiming required a third claim). This pattern held for Ambassador Book Service, which was in the lowest performance level in terms of claims.

The variability in performance carried over to the European vendors as well. Only 26% of the orders placed with Hubert Wilson and 33% placed with B.F. Stevens and Brown arrived without claiming. This is in contrast to 85% of the orders with Aux Amateurs de Livres which never needed claiming.

3.0 PRODUCT-LIMIT ESTIMATES

Intrinsic in the process of ordering books is failure. The failure is defined as materials ordered and not arriving due to unavailability, or a variety of other reasons. To adequately characterize this phenomenon, survival models are used. The acquisitions department randomly assigned orders to vendors and as books arrived, arrival times were recorded. Three hundred days after the last order, data gathering terminated. By then, some material had arrived and some had not. Materials that had not arrived were *right-censored* for the experiment. All that is known about these orders is that they did not arrive by day 300. There is a bias in the analysis unless the censored observations are considered. In the acquisitions system, censored observations can

not be ignored in the analysis because their characteristics may reveal something about the reasons for survival.

3.1 Survivor Functions

In the acquisitions experiment there are a number of observations in which orders either arrive or not. Each observation has an associated time T which measures the number of days from order to receipt of the item. This variable has a probability density function (*PDF*) of $f(t)$. The cumulative distribution function is

$$F(t) = Prob(T \leq t) = \int_0^t f(x)dx. \tag{1}$$

The *survivor function* is used to describe the pattern of life of book orders. It is given by

$$S(t) = Prob(T \geq t) = \int_t^\infty f(x)dx. \tag{2}$$

It yields the probability that an order selected at random from the population of all orders and having a lifetime of T, will have a lifetime exceeding time t. At time $t=0$, $S(t)$ and $S(\infty)=0$.

The *hazard function* $h(t)$ specifies the instantaneous rate at which book orders arrive (die) at a point in time t, given the order does not arrive until time t:

$$h(t) = \lim_{\Delta t \to \infty} \frac{Prob(t \leq T < t + \Delta t | T \geq t)}{\Delta t}.$$

$$h(t) = \frac{f(t)}{S(t)}. \tag{3}$$

Thus, it is the probability density function divided by the survivor function.

For computational purposes, it is necessary to develop discrete versions of the continuous formulation of the models above. To do this, assume there are a series of discrete points in time $t_1, t_2, \ldots, t_k$ when a book order arrives. Corresponding to each arrival time t_j is a count, d_j, of the number of items that have arrived at that t_j. In addition, there are n_j orders that have *not* arrived at each arrival time t_j. The Product Limit Estimate of the survivor function $S(t)$ at time t_i is

$$\hat{S}(t_i) = \prod_{j=1}^{i} \frac{n_j - d_j}{n_j}$$

This is sometimes referred to as the Kaplan-Meier Estimate[5].

[5] See Kaplan & Meier (1958) for details. See also (Kalbfleisch & Prentice, 1982, p. 14) for formulas to calculate the mean and variance of $S(t)$.

3.2 Product Limit Estimates

Kaplan-Meier estimates were made of the survivor function; and means, standard errors, and quantiles were calculated for the number of days until arrival of an order. Table VI summarizes the descriptive statistics for three separate groupings: language of publication, category of material, and number of claims.

The mean number of days an English-language item took to arrive was 105 while French material took 188 days. Items that were in *BIP* and ordered "Rush", arrived in 82 days, but "NonRush" *BIP* items took only took 15 days longer. If an item was never claimed it could be expected to arrive in 98 days. If the library claimed an item at all, there was a substantial delay over the baseline in waiting time.

The average number of days vendors took to complete their orders is given in Table VII. Two U.S. vendors (Yankee Book Peddler and Roy Young Bookseller) have the shortest average delivery time of 78 and 82 days, respectively. A second category of vendors (Midwest Library Service, Blackwell North America, Academic Book Center, and The Book House) have delivery times in the range of 89 to 102 days. The slowest delivery came from Book Service International, Ambassador Book Service, and Coutts Library Service with mean delivery times of 117 to 123 days.

TABLE VI
Product Limit Survival Estimates for Number of Days until Order Arrives

	Quantiles				Standard	Total	%
	25%	50%	75%	Mean	Error	Orders	Censored
Language							
English	73	92	118	105	1	2552	1.5
French	156	180	229	188	3	353	6.2
Category of Material							
Art	65	86	115	100	5	108	1.9
In 'Books in Print'	71	86	107	97	1	1601	0.9
Not in 'Books in Print'	78	103	139	114	3	218	0.5
In 'Books in Print' —Rush	58	71	92	82	4	136	1.5
British	105	121	163	143	3	295	3.4
Canadian	102	126	190	147	8	63	6.3
French	159	181	229	189	3	341	5.3
Reprints	76	88	113	98	5	64	0.0
Number of Claims							
Never Claimed	71	86	105	98	1	2123	0.7
Claimed Once	120	138	175	150	2	639	1.7
Claimed Twice	232	245	278	244	5	74	16.2
Claimed Three Times	246	—	—	256	6	27	55.6

TABLE VII
Product Limit Survival Estimates for Number of Days until Order Arrives, by Vendor

Vendor	Quantiles 25%	50%	75%	Mean	Standard Error	Total Orders	% Censored
Location—United States							
Academic Book Center, OR	72	83	104	98	3	253	0.0
Ambassador Book Service, NY	90	97	126	120	5	127	4.7
Blackwell North America, OR	69	85	100	93	2	264	1.5
The Book House, MI	72	89	113	102	2	396	1.5
Book Service International, CT	93	110	124	117	3	249	2.8
Coutts Library Service, NY	93	110	134	123	3	248	1.6
Midwest Library Service, MO	64	78	94	89	2	307	0.3
Roy Young Booksellers, NY	64	75	84	82	2	248	0.4
Yankee Book Peddler, NH	61	71	82	78	3	106	0.0
Location—Europe							
Aux Amateurs de Livres, Paris	155	175	229	183	5	113	5.3
B.H. Blackwell, Oxford	86	106	127	119	5	108	3.7
Otto Harrassowitz, Wiesbaden	156	202	258	205	6	77	11.7
Martinus Nijhoff, The Hague Neth.	176	188	215	196	5	73	4.1
B.F. Stevens & Brown, Surry England	113	138	166	150	5	102	2.9
Jean Touzot Librairie, Paris	149	170	201	172	5	102	3.9
Hubert Wilson, Margate England	116	141	225	165	6	84	3.6

Among the British dealers, the best performance was provided by B.H. Blackwell of Oxford England with an average delivery time of 119 days. The best delivery time for French materials was by Jean Touzot, with the performance of Aux Amateurs close behind. Most orders placed with Otto Harrassowitz were for French materials, and a significant number of them were right-censored, indicating the orders did not arrive by the end of the 300 day limit. Harassowitz's mean arrival time was the highest in the group.

A comparison of Tables V and VII is instructive for its consistency. Table V presents an analysis of vendor performance based on claiming frequency, while Table VII analyzed performance based on delivery time. Both tables

show almost the same results. Vendors that rank high in the proportion of orders never needing claiming also have the shortest delivery time for the sample.

3.3 Improved Claiming Strategies

It is possible to improve claiming strategy considerably by examining the shape of the survivor function. Figure 1 plots the survivor function for material that was found in *BIP* and for French-language materials. Recall

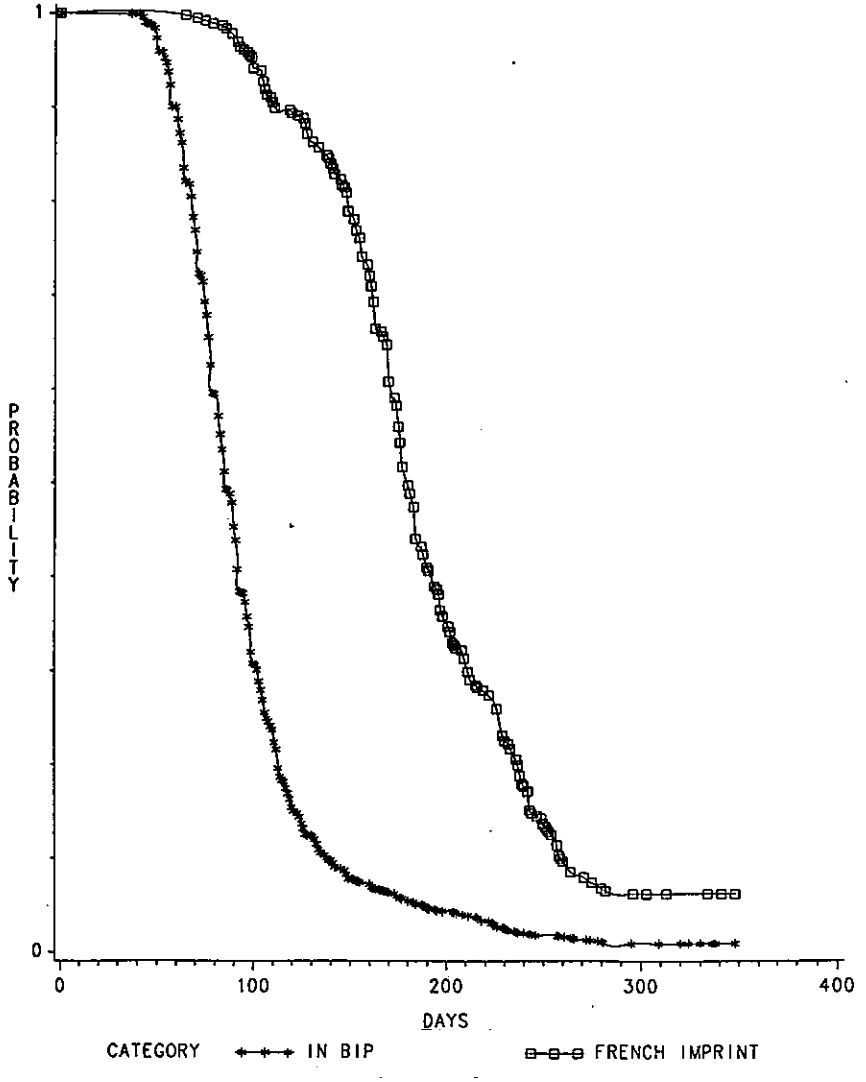

Figure 1. Survivor functions for French materials and BIP items.

that the survivor function gives the probability that the life of an order will exceed a specified number of days, t. It can be observed from Figure 1 that items in *BIP* are unlikely to arrive before 50 days has elapsed, and if they have not arrived in 150 days there is a small probability of them ever arriving. In contrast, French-language materials rarely start to arrive before 75 days has elapsed. If the item has not arrived after about 260 days, it probably will not come.

Figure 2 shows the Kaplan-Meier survivor function estimates for three representative vendors: Yankee Book Peddler, Blackwell North America, and Hubert Wilson. Yankee's survivor function shows no arrivals can be expected for 50 days after an order has been placed. After 75 days, the curve begins to reduce its rate of decline, and by 107 days, it has leveled off.

The curve for Blackwell North America has no arrivals before 50 days as well. Materials arrive for the next 50 days and then sharply decline. In contrast, Hubert Wilson's survivor function shows orders not arriving before 100 days, a leveling off in arrivals after 150 days, and continued arrivals until 275 days.

These graphs provide valuable information to assist in performing claiming operations. They tell the earliest date one could expect an item to arrive and the pattern of arrivals. Claiming strategies could be tailored to the shape of the survivor curve.[6] For example, if a claim were about to be sent to Hubert Wilson (Figure 2), it would make sense to wait 100 days before expecting anything. If nothing had been received in 150 days, that might be a good time to issue a claim. In other words, the shape of the survivor function can be an important aid to optimizing claiming strategies.

4.0 ACCELERATED FAILURE TIME MODELS

Consider the relationship between rate of failure (arrival) and a set of independent variables such as vendor, type of material, and language, using parametric models that have a probability distribution characterizing failure (arrival) time. This section shows the development of accelerated failure time models which are used to analyze the acquisitions data. The models are presented for one specific hazard distribution of arrival time, the Weibull distribution.[7]

---

[6] The decision to claim an order is obviously influenced by many other factors beside the shape of the survivor curve. The purpose of the discussion is to indicate the utility of the approach, and suggest that survivor curves be one additional type of information an acquisitions librarian uses when making a claiming decision.

[7] The presentation roughly follows that of Kalbfleisch and Prentice (1980), but with notation changes from Lawless (1982). Further details may be found in the two works. The discrete versions of the formulas are not presented nor are computational formulas.

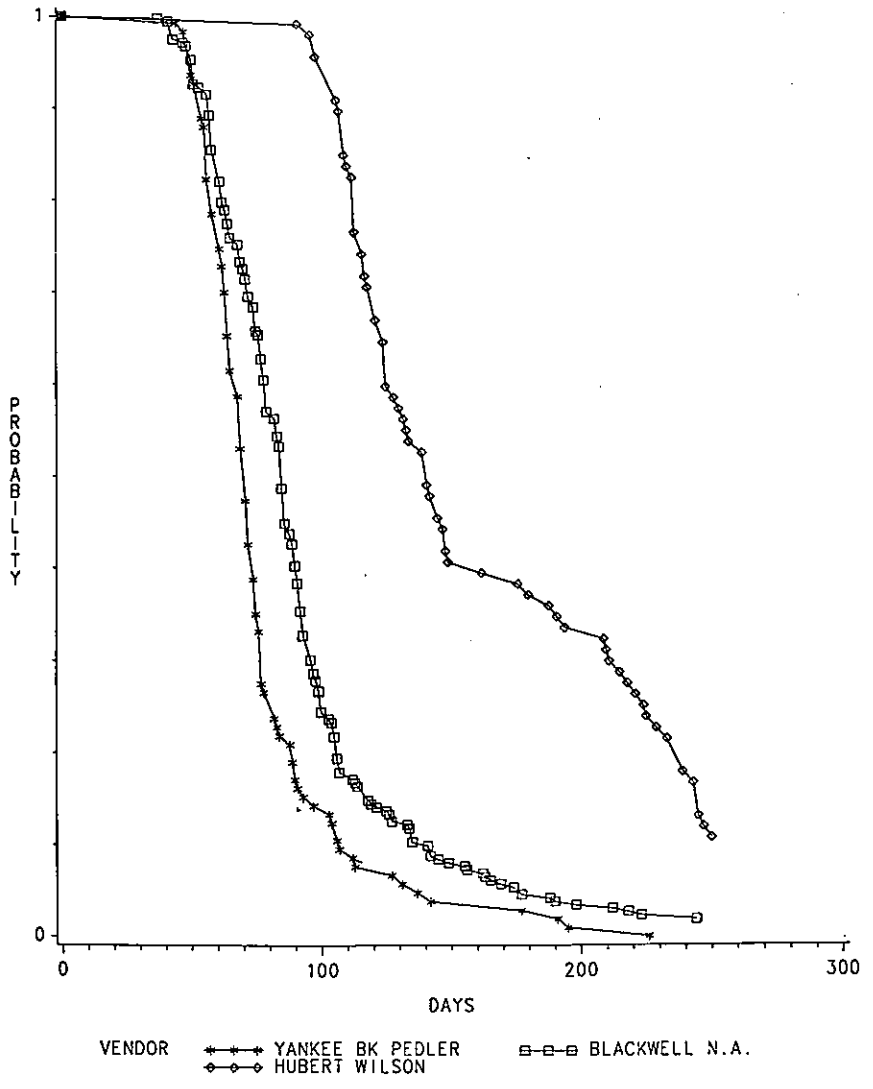

Figure 2. Survivor functions for selected vendors.

The distribution of failure times is reflected in the shape of the hazard function. Using (2), the survivor function can be rewritten in terms of the hazard function (Lawless, 1982, p. 9) to yield

$$S(t) = \exp\left(-\int_0^t h(x)dx\right). \tag{4}$$

From (3) we have

$$f(t) = h(t)\, S(t). \tag{5}$$

Thus, the PDF from (5) and (4), written in terms of the hazard function, is

$$f(t) = h(t) \exp\left(-\int_0^t h(x)dx\right) \tag{6}$$

Equations (4) and (6) provide a basic structure into which appropriate hazard functions can be substituted.

4.1 Exponential and Hazard Models

For an exponential distribution of failure times, the hazard function is constant:

$$h(t) = \lambda \tag{7}$$

where $t \geq 0$ and $\lambda > 0$. Using (7), the PDF from (6) becomes

$$f(t) = \lambda e^{-\lambda t} \tag{8}$$

and the survivor function with an underlying exponential distribution of failure times becomes [from (4) and (7)]

$$S(t) = \exp\left(-\int_0^t \lambda dx\right) = e^{-\lambda t}. \tag{9}$$

The hazard function for the Weibull distribution is given by Kalbfleisch & Prentice, 1980, (p. 23) as

$$h(t) = \lambda p (\lambda t)^{p-1} \tag{10}$$

where $p > 0$ and $\lambda > 0$. From (6) and (10) the PDF is

$$f(t) = \lambda p (\lambda t)^{p-1} \exp\left[-(\lambda t)^p\right], \tag{11}$$

and from (4) and (10) the survivor function is

$$S(t) = \exp\left[-(\lambda t)^p\right]. \tag{12}$$

The Weibull hazard function is characterized by p which is termed the shape parameter. If $p > 1$ the function is monotone increasing. If $p < 1$ it is monotone decreasing, and if $p = 1$ it is constant. The λ value is a scale parameter which has the effect of changing the scale of the horizontal or time axis. [Lawless, pp. 16–17].

4.2 Weibull Model with Covariates

Now consider the relation between arrival time and a set of independent variables such as language of publication, vendor, and year of publication.

These characteristics can be defined as $\mathbf{z}$ where $\mathbf{z}=(z_1,z_2,\ldots,z_n)$, and the associated regression parameters are β where $\beta'=(\beta_1,\beta_2,\ldots,\beta_n)$.

The hazard function at time t for an observation with independent variables $\mathbf{z}$ is

$$h(t;\mathbf{z})=\lambda(\mathbf{z}). \tag{13}$$

If the relation between the hazard and the covariates is linear, the functional form could be expressed as just the product $\mathbf{z}\beta$. To begin, assume an exponential distribution with $c=e^{(\mathbf{z}\beta)}$, and thus (13) becomes

$$h(t;\mathbf{z})=\lambda e^{\mathbf{z}\beta}. \tag{14}$$

Using (8), the conditional PDF is

$$f(t;\mathbf{z})=\lambda e^{\mathbf{z}\beta}\exp(-\lambda t\, e^{\mathbf{z}\beta}). \tag{15}$$

Covariates $\mathbf{z}$ can be added to the Weibull hazard model of (10) to obtain

$$h(t;\mathbf{z})=p(\lambda t)^{p-1}\, e^{\mathbf{z}\beta}. \tag{16}$$

The PDF with covariates is then derived from (11):

$$h(t;\mathbf{z})=\lambda p(\lambda t)^{p-1}\, e^{\mathbf{z}\beta}\exp[-(\lambda t)^p\, e^{\mathbf{z}\beta}]. \tag{17}$$

This equation can be written in a log-linear form for ease of fitting the equation to the data. (Kalbfleisch & Prentice, 1980, p. 32).

4.3 The Proportional Hazards Model

Kalbfleisch and Prentice (1982, p. 32) show that the effect of equation (17) is that the independent variables act multiplicatively on the hazard function. They present a general multiplicative model proposed by Cox[8] where

$$h(t;\mathbf{z})=h_0(t)e^{\mathbf{z}\beta} \tag{18}$$

Here h_0 is the base line hazard function. The PDF for a Weibull distribution is given in (11). The conditional PDF employing the proportional hazard model of (18) is

$$f(t;\mathbf{z})=h_0(t)e^{\mathbf{z}\beta}\exp[-e^{\mathbf{z}\beta}\int_0^t h_0(u)du] \tag{19}$$

and the conditional survivor function is

$$S(t;\mathbf{z})=[S_0(t)]^{\exp(\mathbf{z}\beta)}$$

where

$$S_0(t)=\exp[-\int_0^t h_0(u)du]. \tag{20}$$

[8] See Cox and Oakes (1984, p. 40) for a discussion of the model.

Alternatively,

$$S(t;\mathbf{z}) = \exp[-\int_0^t h_0(u)du]^{\exp(\mathbf{z}\beta)} \tag{21}$$

4.4 Accelerated Failure Time Model

The final step is developing the accelerated failure time model. A deficiency with the proportional hazard model is that the failure time depends only on the independent variables and not on time to failure. The accelerated failure time model corrects that problem by assuming that the covariates modify the time to failure by lengthening or shortening it.

The conditional hazard function for the accelerated failure time model is given in (Kalbfleisch & Prentice, 1980, p. 34) as

$$h(t;\mathbf{z}) = h_0(te^{-\mathbf{z}\beta})\, e^{-\mathbf{z}\beta} \tag{22}$$

and the survivor function is

$$S(t;\mathbf{z}) = \exp[-\int_0^t \exp(-\mathbf{z}\beta)\, h_0(u)du]. \tag{23}$$

As in the proportional model, the accelerated failure time model can be rewritten as a linear model. The equation can be fitted to the data using maximum likelihood estimates and Newton–Raphson techniques. They are discussed in SAS User's Guide. Statistics, 1985, pp. 507–528. Examples of the method are given in Kalbfleisch & Prentice (1980, pp. 54–67).

4.5 Analysis of Acquisitions Data

The accelerated failure time model was used to analyze the acquisitions data. The independent variables used in the model were the vendors (Table I), categories of materials (Table III), year of publication, and number of times the order was claimed. The dependent variable was the number of days between order and receipt of the item. After verifying the fit, a Weibull distribution was used as the underlying model of the data.[9] A total of 2785 orders were used in the analysis, of which 53 were censored. A total of 1813 observations had to be excluded due to missing data.

The results of the accelerated failure time model are analyzed in terms of a reference or base group of observations. The reference group for this experiment is materials ordered from "The Book House," that was in *BIP* and was published in English. The estimated coefficients and significance measures of the independent variables are given in Table VIII. The last column in the Table gives the proportionate increase or decrease in the arrival rate in relation to the baseline group. For example, the value for the vendor

[9] A plot of $log[-log\hat{S}(t)]$ versus $log(t)$ was made for most of the independent variables. It was found that the resulting curves were approximately a straight line indicating a good fit to the Weibull distribution.

TABLE VIII
Accelerated Failure Time Analysis of Acquisitions Data $n = 2732$

Variable	Estimated Coefficient	Standard Error	χ^2 Probability	Proportionate Change Over Base-line
Intercept	4.273	0.304	0.0001	
Location—United States				
Academic Book Center, OR	0.043	0.025	0.0894	0.958
Ambassador Book Service, NY	0.101	0.033	0.0020	0.904
Blackwell North America, OR	−0.026	0.025	0.2954	1.026
Book Service International, CT	0.103	0.026	0.0001	0.902
Coutts Library Service, NY	0.089	0.025	0.0004	0.915
Midwest Library Service, MO	−0.079	0.024	0.0009	1.082
Yankee Book Peddler, NH	−0.065	0.034	0.0546	1.067
Roy Young Booksellers, NY	−0.108	0.025	0.0001	1.114
Location—Europe				
Aux Amateurs de Livres, Paris	0.522	0.134	0.0001	0.593
B.H. Blackwell, Oxford	0.582	0.307	0.0577	0.559
Otto Harrassowitz, Wiesbaden	0.688	0.134	0.0001	0.503
Martinus Nijhoff, The Hague Neth.	0.584	0.136	0.0001	0.558
B.F. Stevens & Brown, Surry England	0.512	0.307	0.0951	0.599
Jean Touzot Librairie, Paris	0.393	0.133	0.0031	0.675
Hubert Wilson, Margate England	0.662	0.307	0.0311	0.516
Category of Material				
Art Books	0.047	0.040	0.2350	0.954
British Material	−0.256	0.305	0.4005	1.292
Canadian Material	0.437	0.042	0.0001	0.646
French Language Material	0.408	0.131	0.0018	0.665
In 'BIP'—Rush	−0.288	0.028	0.0001	1.334
Not in 'BIP'	0.076	0.022	0.0007	0.927
Reprint Material	−0.061	0.042	0.1397	1.063
Year of Publication	0.0001	0.0002	0.4741	1.000
Number of times order claimed	0.480	0.011	0.0001	0.619
Scale Parameter	0.304	0.004		

Notes:
1. The base-line observation for the accelerated failure time analysis was an order from 'The Book House' for an item listed in *BIP* and published in English.
2. See text for explanation of last column in table.

Roy Young Bookseller is 1.114.[10] This means that in comparison to English-language materials that were in BIP and ordered from The Book House, materials ordered from Roy Young Bookseller arrived about 11% faster. A negative coefficient implies an improvement over the baseline group.

The nonparametric analysis showed there was a small difference in the number of days between order and arrival for "Rush" versus "Non-Rush" materials. When all other variables are factored into the accelerated failure time model, the results indicate that material ordered "Rush" arrives one third faster than that in the baseline group.

The lack of significance of certain variables means they do not contribute to the explanatory power of the equation. One of the more interesting cases is "year of publication." The number of days it takes for an order to arrive is not influenced by its year of publication. The fact that an item has a British imprint also has no effect.[11] These results conform to the analysis presented earlier. Table 4, for example, showed relative invariance between publication date and claiming frequency.

The statistical significance of the scale parameter confirms a very intuitive and obvious fact that there is a strong relation between passage of time and arrival of orders.

5.0 SUMMARY AND CONCLUSIONS

This paper has presented several models to evaluate the book ordering process, and has provided quantitative information on vendor performance. A total of more than 4,700 orders placed to 16 vendors were analyzed. These orders were classified in terms of the urgency of the request, type of material, publisher, and date of publication. Overall, about 36% of all orders were claimed. There was little variation in claiming frequency by imprint date, but significant variation by vendor. At the extremes, 15% of the orders to one vendor and 74% to another needed claiming. Using product limit survival estimates, large variations were found between vendors in the mean number of days until orders arrived. The range was between 82 and 205 days.

Graphs presented in the paper showed the possibility of improving claiming strategy by analyzing the shape of the survivor function. By adjusting the time when claims are sent, better results should be obtained.

Finally, a multivariate analysis of the data was undertaken using an accelerated failure time model. This model was used to explain variations in

[10] The value is calculated as $e^{-\beta}$. In this case, $\beta = -0.108$ (from Table VIII), so $e^{-(-0.108)} = 1.114$. The reason for exponentiation of the coefficient is that the regression is performed in terms of logarithms.

[11] In these examples, the criterion used to decide significance is a χ^2 probability value less than 0.10

arrival time of orders given a set of independent variables about the order. The analysis showed the proportionate change in performance of various vendors relative to a pre-established base. It demonstrated that certain variables such as imprint date have little effect on the time until an order arrives.

One important point must be made about these results. It was made by Barker in his original paper and is repeated here:

> At U.C. Berkeley, we have learned to value vendors for their attention to detail, tenacity, precision, efficiency, friendliness, and scope of services as well as their speed and overall effectiveness in obtaining the books. Quantified vendor evaluation is not a substitute for ongoing two-way communication of needs, expectations, frustrations, and appropriate expressions of appreciation. [Barker (1986), p. 280].

REFERENCES

Barker, Joseph W. (1986). Random vendor assignment in vendor performance evaluation, *Library Acquisitions: Practice and Theory 10,* 265–280.

Cox, David R., and D. Oakes. (1984). *"Analysis of Survival Data."* Chapman and Hall, London. (Monograph on Statistics and Applied Probability).

Kalbfleisch, John D., and Ross L. Prentice. (1980). *"The Statistical Analysis of Failure Time Data."* John Wiley and Sons, New York.

Kaplan, E.L., and P. Meier. (1958). Nonparametric estimation from incomplete observations, *Journal of the American Statistical Association 53,* 457–481.

Lawless, Jerald F. (1982). *"Statistical Models and Methods for Lifetime Data."* John Wiley, New York.

SAS Institute, Inc. (1985). *"SAS User's Guide: Statistics, Version" (5th Edition).* SAS Institute Inc., Cary, NC.