Customer-Base Analysis with Discrete-Time Transaction Data

Peter S. Fader
Bruce G. S. Hardie
Paul D. Berger[1]

September 2004

[1]Peter S. Fader is the Frances and Pei-Yuan Chia Professor of Marketing at the Wharton School of the University of Pennsylvania (address: 749 Huntsman Hall, 3730 Walnut Street, Philadelphia, PA 19104-6340; phone: 215.898.1132; email: faderp@wharton.upenn.edu; web: www.petefader.com). Bruce G. S. Hardie is Associate Professor of Marketing, London Business School (email: bhardie@london.edu; web: www.brucehardie.com). Paul D. Berger is Professor, Marketing Department, School of Management, Boston University (email: pdberger@bu.edu). The second author acknowledges the support of the London Business School Centre for Marketing and the hospitality of the Department of Marketing at the University of Auckland Business School.

Abstract

Customer-Base Analysis with Discrete-Time Transaction Data

Many businesses track repeat transactions on a discrete-time basis. These include: (1) companies with transactions that occur at regular intervals (such as subscription renewals), (2) firms that frequently associate transactions with specific events (e.g., a direct marketer that records whether or not customers respond to a particular catalog), and (3) organizations that simply use discrete reporting periods even though the transactions can occur at any time. Furthermore, many of these businesses operate in a noncontractual setting, so they have a difficult time differentiating between those customers who have ended their relationship with the firm versus those who are in the midst of a long hiatus between transactions. Our goal is to develop a model to predict future purchase patterns for a customer base that can be described by these structural characteristics. Our "beta-geometric/beta-binomial" (BG/BB) model allows for heterogeneity in each of the underlying behavioral processes (customers' purchase propensities while active, and time until each customer becomes permanently inactive), and yields relatively simple closed-form expressions for future expectations conditional on past observed behavior. We apply the model to a previously published dataset consisting of cruise-line transactions for a cohort of 6094 customers over a period of five years, and demonstrate the valuable insights that arise from our forward-looking modelling framework.

Keywords: BG/BB, beta-geometric, beta-binomial, customer-base analysis, customer lifetime value, CLV, RFM, Pareto/NBD

1 Introduction

Customer-base analysis seeks to use information on the history of customer purchase patterns to identify which individuals are most likely to be active (or inactive) customers, to predict future purchasing patterns by those customers listed in the firm's transaction database, and to generate estimates of the expected lifetime value of as-yet-to-be-acquired customers. What makes this so difficult in noncontractual settings is that most customers do not notify the firm when they become inactive (i.e., shift patronage to a different supplier or stop purchasing in the category altogether). The only potential evidence of this having happened is an unusually long hiatus since the last recorded purchase.

The Pareto/NBD model was developed by Schmittlein, Morrison, and Colombo (1987) to provide answers to these standard customer-base analysis questions. It requires only two pieces of information about each customer's past purchasing history: his "recency" (when his last transaction occurred) and "frequency" (how many transactions he made in a specified time period). Using these two key summary statistics, Schmittlein et al. derive expressions for a number of managerially relevant quantities, such as the probability that an individual is still active given his observed purchasing behavior, and the expected number of transactions in the next period (e.g., year) conditional on the customer's observed purchasing behavior.

The NBD component of the model assumes that, while the customer is active, transactions can occur at any point in time. This is satisfied in the empirical validations of the model presented in Schmittlein and Peterson (1994) and Fader, Hardie, and Lee (2004a), which use data on the purchasing of office products in a B2B setting and the purchasing of CDs at an online retailer, respectively.

In many settings, however, opportunities for transactions occur at discrete intervals. For example, I can either go to church each Sunday, or not; each week I can buy the latest copy of *People* magazine at a newsvendor, or not. In other settings, purchasing data are discretized for analysis and reporting purposes. For example, a cruise-ship company may characterize customer behavior in terms of whether or not each customer went on a cruise in 2000, 2001, 2002, etc. Similarly, a mail-order company may code customer behavior in terms of whether or not the customer placed an order after receiving each catalog. In all of these cases, it would not be appropriate to use the Pareto/NBD for a customer-base analysis exercise; rather, we require a discrete-time analog of the model.

The objective of this paper is present the beta-geometric/beta-binomial (BG/BB) model as such a discrete-time analog. We first outline the assumptions underpinning this model and derive expressions for a number of managerially relevant quantities. This is followed by an illustrative empirical analysis. We conclude with a discussion of several issues that arise from this work.

2 Model Development

Our goal is to develop a stochastic model of buyer behavior in a setting where opportunities for transactions occur at discrete intervals and the time at which the customer becomes inactive is unobserved by the firm. A customer's purchase history can be represented as a binary string where 1 represents a purchase and 0 represents no purchase on any given transaction opportunity. We can summarize this in terms of "recency" (when the last transaction occurred) and "frequency" (how many transactions occurred in a specified time period). The notation used to represent this information is (x, n, m), where x is the number of purchases that occurred in n transaction opportunities, with the last purchase occurring on transaction opportunity $m \leq n$. For example, reading from left to right, the purchase string 0 1 0 1 1 0 0 0 can be summarized as $(x = 3, n = 8, m = 5)$. Given this summary of a customer's purchasing behavior, we are interested in calculating the probability that the customer is still active as well as an expectation of future purchasing. (A customer is "active" if there is a non-zero probability of his making a purchase in the interval between "now" and the end of the relevant time horizon.)

Our model is based on the following five assumptions:

i. While active, the customer buys on any given transaction opportunity with probability p.

ii. Heterogeneity in p follows a beta distribution with pdf

$$f(p \mid \alpha, \beta) = \frac{p^{\alpha-1}(1-p)^{\beta-1}}{B(\alpha, \beta)}, \quad 0 < p < 1 \,. \tag{1}$$

iii. An active customer becomes inactive at the beginning of the next transaction opportunity with probability q. This implies that the customer's (unobserved) lifetime is distributed across transaction opportunities according to a (shifted) geometric distribution.

iv. Heterogeneity in q follows a beta distribution with pdf

$$f(q \mid \gamma, \delta) = \frac{q^{\gamma-1}(1-q)^{\delta-1}}{B(\gamma, \delta)}, \quad 0 < q < 1 . \tag{2}$$

v. The transaction probability p and the dropout probability q vary independently across customers.

Assumptions (i) and (iii) are equivalent to those associated with the beta-binomial (BB) distribution. Similarly, assumptions (ii) and (iv) yield the beta-geometric (BG) distribution. We therefore call this the beta-geometric/beta-binomial (BG/BB) model of buyer behavior.

2.1 Model Likelihood Function

Consider a customer with the purchase string 0 1 0 1 1 0 0 0. The fact that the customer made a purchase on the fifth transaction opportunity means he was active for the first five transaction opportunities, with probability $(1-q)^5$. Therefore the probability of observing the first five elements of the purchase string (i.e., 0 1 0 1 1) is $p^3(1-p)^2(1-q)^5$. But how can the final string of three zeros have come about? There are four possibilities:

- the customer "died" at the beginning of period 6, with probability q,
- the customer was active (with no purchase) during the 6th period and died at the beginning of period 7, with probability $(1-p)(1-q)q$,
- the customer was active (with no purchases) during the 6th and 7th periods and died at the beginning of period 8, with probability $(1-p)^2(1-q)^2q$, or
- the customer was active (with no purchase) during periods 6–8, with probability $(1-p)^3(1-q)^3$.

Therefore, the probability of observing the purchase history $(x = 3, n = 8, m = 5)$ is

$$p^3(1-p)^2(1-q)^5 q + p^3(1-p)^3(1-q)^6 q + p^3(1-p)^4(1-q)^7 q + p^3(1-p)^5(1-q)^8 \,.$$

More generally, for a customer with purchase history (x, n, m),

$$L(p, q \mid x, n, m) = p^x(1-p)^{n-x}(1-q)^n + \sum_{i=0}^{n-m-1} p^x(1-p)^{m-x+i} q(1-q)^{m+i} \,. \tag{3}$$

To arrive at the likelihood function for a randomly-chosen customer with purchase history (x, n, m), we take the expectation of this over the mixing distributions for p and q:

$$\begin{aligned} L(\alpha, \beta, \gamma, \delta \mid x, n, m) &= \int_0^1 \int_0^1 L(p, q \mid x, n, m) f(p \mid \alpha, \beta) f(q \mid \gamma, \delta) \, dp \, dq \\ &= \frac{B(\alpha + x, \beta + n - x)}{B(\alpha, \beta)} \frac{B(\gamma, \delta + n)}{B(\gamma, \delta)} \\ &\quad + \sum_{i=0}^{n-m-1} \frac{B(\alpha + x, \beta + m - x + i)}{B(\alpha, \beta)} \frac{B(\gamma + 1, \delta + m + i)}{B(\gamma, \delta)} \,. \end{aligned} \tag{4}$$

The four BG/BB model parameters $(\alpha, \beta, \gamma, \delta)$ can be estimated via the method of maximum likelihood in the following manner. Suppose we have a sample of N customers, where customer i's purchase history is denoted by (x_i, n_i, m_i). The sample log-likelihood function is given by

$$LL(\alpha, \beta, \gamma, \delta) = \sum_{i=1}^{N} \ln \left[L(\alpha, \beta, \gamma, \delta \mid x_i, n_i, m_i) \right] \tag{5}$$

This can be maximized using standard numerical optimization routines.

2.2 Derivation of $P(\textbf{active} \mid x, n, m)$

The first result of managerial interest is the probability that a customer with a given purchase history is still active, which we denote by $P(\text{active} \mid x, n, m)$.

Recall the purchase string 0 1 0 1 1 0 0 0 with associated likelihood function

$$p^3(1-p)^2(1-q)^5 q + p^3(1-p)^3(1-q)^6 q + p^3(1-p)^4(1-q)^7 q + p^3(1-p)^5(1-q)^8 \,.$$

Noting that the final term is associated with the scenario in which the customer was active (with no purchase) during periods 6–8, it follows that the probability that the customer was active in period 8 is

$$\frac{p^3(1-p)^5(1-q)^8}{p^3(1-p)^2(1-q)^5q + p^3(1-p)^3(1-q)^6q + p^3(1-p)^4(1-q)^7q + p^3(1-p)^5(1-q)^8}.$$

Since the probability of not dying at the beginning of a transaction opportunity is $1-q$, the probability that this customer is active in period 9 is

$$\frac{p^3(1-p)^5(1-q)^9}{p^3(1-p)^2(1-q)^5q + p^3(1-p)^3(1-q)^6q + p^3(1-p)^4(1-q)^7q + p^3(1-p)^5(1-q)^8}.$$

More generally, for a customer with purchase history (x, n, m), the probability he is active in period $n+1$ is given by

$$P(\text{active} \,|\, x, n, m, p, q) = \frac{p^x(1-p)^{n-x}(1-q)^{n+1}}{L(p, q \,|\, x, n, m)}. \tag{6}$$

Since p and q are unobserved, we compute $P(\text{active} \,|\, x, n, m)$ by taking the expectation of (6) over the distribution of p and q updated to take account of the information (x, n, m):

$$P(\text{active} \,|\, x, n, m, \alpha, \beta, \gamma, \delta) = \int_0^1 \int_0^1 P(\text{active} \,|\, x, n, m, p, q) f(p, q \,|\, x, n, m, \alpha, \beta, \gamma, \delta)\, dp\, dq. \tag{7}$$

By Bayes' theorem, the joint posterior distribution of p and q is given by

$$f(p, q \,|\, x, n, m, \alpha, \beta, \gamma, \delta) = \frac{L(p, q \,|\, x, n, m) f(p \,|\, \alpha, \beta) f(q \,|\, \gamma, \delta)}{L(\alpha, \beta, \gamma, \delta \,|\, x, n, m)}. \tag{8}$$

Substituting (6) and (8) (and (1), (2), (3), and (4)) in (7), and solving the double integral gives us

$$P(\text{active} \,|\, x, n, m, \alpha, \beta, \gamma, \delta) = \frac{B(\alpha + x, \beta + n - x) B(\gamma, \delta + n + 1)}{B(\alpha, \beta) B(\gamma, \delta)} \Big/ L(\alpha, \beta, \gamma, \delta \,|\, x, n, m). \tag{9}$$

2.3 Conditional Expectations of Future Purchasing

We are also interested in forecasting the expected number of future purchases for a customer with a given purchase history.

Let $E(X^* \,|\, n^*, x, n, m)$ denote the expected number of purchases over the next n^* periods by a customer with purchase history (x, n, m). Assuming the customer is active at the beginning of period $n+1$, the next n^* transaction opportunities can be characterized as follows:

Transaction Opportunity	$n+1$	$n+2$	$n+3$	$n+4$	$\cdots$	$n+n^*$
P(active)	1	$1-q$	$(1-q)^2$	$(1-q)^3$		$(1-q)^{n^*-1}$
P(buy | active)	p	p	p	p		p

Noting that the sum of the first n terms of a geometric series is

$$a + ak + ak^2 + \cdots ak^{n-1} = a\frac{1-k^n}{1-k}\,,$$

it follows that, conditional on p and q, the expected number of transactions in the interval $(n, n^*]$ is

$$\frac{p}{q} - \frac{p}{q}(1-q)^{n^*}\,.$$

Multiplying this by the probability that a customer with purchase history (x, n, m) (and latent traits p and q) is still active, (6), and taking the expectation over the joint posterior distribution of p and q, (8), gives us

$$\begin{aligned} E(X^* \,|\, n^*, x, n, m, \alpha, \beta, \gamma, \delta) = {} & \frac{B(\alpha+x+1, \beta+n-x)}{B(\alpha,\beta)} \\ & \times \left[\frac{B(\gamma-1, \delta+n+1) - B(\gamma-1, \delta+n+n^*+1)}{B(\gamma,\delta)}\right] \Big/ L(\alpha,\beta,\gamma,\delta \,|\, x, n, m)\,. \end{aligned} \tag{10}$$

Perhaps of greater interest is how to use the BG/BB model in the calculation of customer lifetime value (CLV), which is "the present value of future cash flows attributed to the customer relationship" (Pfeifer, Haskins, and Conroy 2005, p. 10). Since many treatments of the calculation of CLV assume a constant cash flow per transaction per customer over time (e.g., Berger and Nasr 1998; Gupta and Lehmann 2003; Libai, Narayandas, and Humby 2002), we can factor out the value of each transaction and focus on forecasting the "flow" of future transactions (discounted to yield a present value). This number of discounted expected transactions (DET) can then be rescaled by the customer's value multiplier to yield an overall estimate of CLV. (See Fader, Hardie, and Lee (2004b) for an examination of this in the context of the Pareto/NBD model.)

Standing at the end of the nth transaction opportunity, we wish to compute the present value of the expected future transaction stream for a customer with purchase history (x, n, m), with discount rate d.[1] Let $DET(d \mid x, n, m)$ denote this quantity.

Assuming the customer is active at the beginning of period $n+1$, the basic structure of possible purchases is as follows:

Transaction Opportunity	$n+1$	$n+2$	$n+3$	$n+4$	$\cdots$	$n+t$	$\cdots$
P(active)	1	$1-q$	$(1-q)^2$	$(1-q)^3$	$\cdots$	$(1-q)^{t-1}$	$\cdots$
P(buy | active)	p	p	p	p	$\cdots$	p	$\cdots$
Discount	1	$(1+d)$	$(1+d)^2$	$(1+d)^3$	$\cdots$	$(1+d)^{t-1}$	$\cdots$

Noting that the sum of an infinite geometric series is

$$\sum_{n=0}^{\infty} ak^n = \frac{a}{1-k}\,,$$

it follows that, conditional on p and q, the present value (at the beginning of transaction opportunity $n+1$) of a customer's expected transaction stream over his remaining lifetime with discount rate d is

$$\sum_{t=1}^{\infty} p\left(\frac{1-q}{1+d}\right)^{t-1} = \frac{p(1+d)}{d+q}\,. \tag{11}$$

Multiplying this by the probability that a customer with purchase history (x, n, m) (and latent traits p and q) is still active at the beginning of transaction opportunity $n+1$, (6), gives us

$$DET(d \mid x, n, m, p, q) = \frac{p^{x+1}(1-p)^{n-x}(1-q)^{n+1}(1+d)}{(d+q)L(p, q \mid x, n, m)}\,. \tag{12}$$

Taking the expectation of this over the joint posterior distribution of p and q, (8), gives us

$$\mathsf{A} \times (1+d)\frac{B(\alpha+x+1, \beta+n-x)}{B(\alpha,\beta)}\Big/ L(\alpha, \beta, \gamma, \delta \mid x, n, m)\,,$$

[1] Suppose there are k opportunities for transactions per year. An annual discount rate of $(100 \times r)\%$ maps to a discount rate of $d = (1+r)^{1/k} - 1$.

where

$$\mathsf{A} = \int_0^1 \frac{1}{d+q} \frac{q^{\gamma-1}(1-q)^{\delta+n}}{B(\gamma,\delta)}\, dq$$

letting $s = 1 - q$

$$\begin{aligned} &= \frac{1}{B(\gamma,\delta)} \int_0^1 \frac{1}{(1+d)-s}(1-s)^{\gamma-1} s^{\delta+n}\, ds \\ &= \frac{1}{B(\gamma,\delta)(1+d)} \int_0^1 s^{\delta+n-1}(1-s)^{\gamma-1}\big(1-\big(\tfrac{1}{1+d}\big)s\big)^{-1}\, ds \end{aligned}$$

which, recalling the integral representation of the Gaussian hypergeometric function[2]

$$= \frac{B(\gamma,\delta+n+1)}{B(\gamma,\delta)(1+d)}\, {}_2F_1\big(1,\delta+n+1;\gamma+\delta+n+1;\tfrac{1}{1+d}\big)\,,$$

Therefore,

$$\begin{aligned} DET(d \,|\, x,n,m,\alpha,\beta,\gamma,\delta) = \frac{B(\alpha+x+1,\beta+n-x)B(\gamma,\delta+n+1)}{B(\alpha,\beta)B(\gamma,\delta)} \\ \times\, {}_2F_1\big(1,\delta+n+1;\gamma+\delta+n+1;\tfrac{1}{1+d}\big)\Big/ L(\alpha,\beta,\gamma,\delta \,|\, x,n,m)\,. \end{aligned} \tag{13}$$

Noting[3] that ${}_2F_1(a,b;c;1) = \big[\Gamma(c)\Gamma(c-a-b)\big]/\big[\Gamma(c-a)\Gamma(c-b)\big]$, it follows that setting $d = 0$ in (13) yields the following expression for the (undiscounted) expected number of transactions over the customer's remaining lifetime:

$$\begin{aligned} &E(X^* \,|\, x,n,m,\alpha,\beta,\gamma,\delta) \\ &\qquad = \frac{B(\alpha+x+1,\beta+n-x)B(\gamma-1,\delta+n+1)}{B(\alpha,\beta)B(\gamma,\delta)}\Big/ L(\alpha,\beta,\gamma,\delta \,|\, x,n,m)\,, \end{aligned} \tag{14}$$

which is $\lim_{n^*\to\infty} E(X^* \,|\, n^*, x, n, m, \alpha, \beta, \gamma, \delta)$.

In some situations we would like to compute DET for a randomly-chosen customer (i.e., without any conditioning on the purchase history (x, n, m)). Taking the expectation of (11)

[2] http://functions.wolfram.com/07.23.07.0001.01

[3] http://functions.wolfram.com/07.23.03.0002.01

over the mixing distributions for p and q, (1) and (2), gives us

$$DET(d \mid \alpha, \beta, \gamma, \delta) = \left(\frac{\alpha}{\alpha+\beta}\right) {}_2F_1\big(1, \delta; \gamma+\delta; \tfrac{1}{1+d}\big)$$

This expression for $DET(d)$ can be viewed as the (discounted) expected number of transactions for an as-yet-to-be-acquired customer. Multiplying the corresponding CLV number by the probability of acquisition represents an upper bound on what should be spent on acquiring a customer.

3 Illustrative Application

We apply the BG/BB model to data from a cruise-ship company with the disguised name "Joyful Voyages, Inc." (JV). We consider the cohort of 6094 customers who made their first cruise with JV in 1993, and track their repeat cruises over the following four years (1994–1997). While there is a small number of customers who make more than one cruise a year, these data have been summarized in terms of whether or not the customer took a cruise in each year — see Figure 1. As such, the behavior of each customer is characterized by one of 16 binary strings of length four: from 1 1 1 1 (for a customer who took a cruise in 1994, 1995, 1996, and 1997) to 0 0 0 0 (for a customer who took no repeat cruises with JV). (See Berger, Weinberg, and Hanna (2003) for further details about this dataset.)

Maximum likelihood parameter estimates of the four BG/BB model parameters $(\alpha, \beta, \gamma, \delta)$ are obtained by maximizing the log-likelihood function given in (5).[4] The maximum value of the log-likelihood function is −7130.7, which is associated with the following parameter estimates: $\hat{\alpha} = 0.657, \hat{\beta} = 5.193, \hat{\gamma} = 173.761, \hat{\delta} = 1882.928$.[5] (All parameter estimates are significant at $p < 0.001$.) The good fit of the model to the data is illustrated in Figure 2, which compares the expected numbers of customers taking cruises in $0, 1, \ldots, 4$ of the following four years to the actual frequency distribution.

[4]Note that the entire analysis can easily be performed in a spreadsheet environment; a copy of the associated Excel spreadsheet is available from the authors.

[5]The magnitude of $\hat{\gamma}$ and $\hat{\delta}$ suggests that the beta distribution for q is effectively a spike at $E(q) = 0.084$. Fitting a model that does not allow for heterogeneity in q (i.e., a G/BB model) yields the same value for the log-likelihood function. This suggests that, for this dataset, there is no heterogeneity in the dropout parameter q.

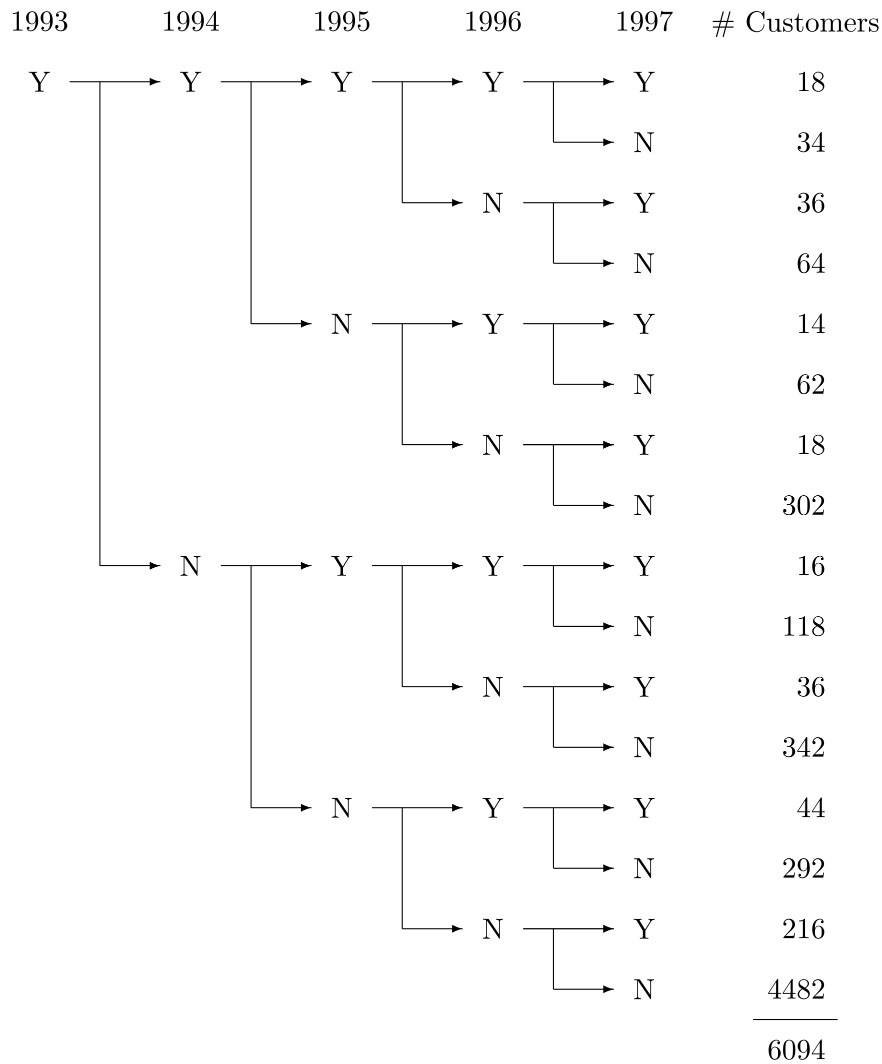

Figure 1: Summary of purchasing by the 1993 cohort of new customers

For each of the 16 possible purchase histories, we use (9) to compute the probability that a customer with that purchase history will still be active at the beginning of 1998. This quantity is reported in Table 1 as a function of recency (the year in which the last cruise was taken) and frequency (the number of years in which the customer took a repeat cruise).

We note that all those customers who took a cruise in 1997 have the same probability of being alive in 1998. Of course, the probability of their actually taking a cruise in 1998 will vary, with those having taken a cruise in each of the last four years having here a far higher probability than those whose first repeat cruise occurred in 1997. Looking at the 1996 column, we see that the probability of being alive in 1998 is a decreasing function of frequency. The reason for this

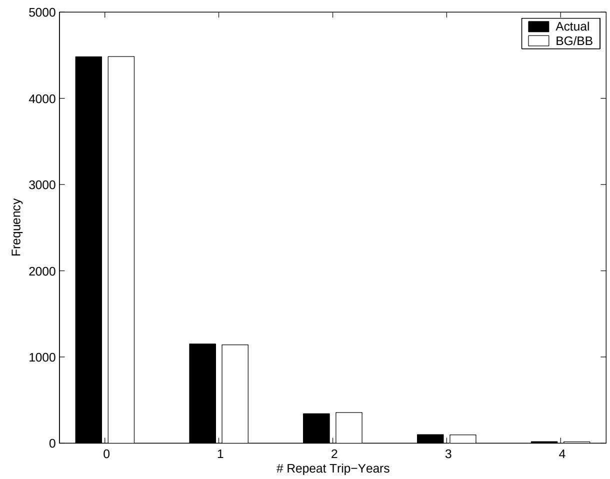

Figure 2: Predicted versus actual # years in which repeat cruises occur

	Year of Last Cruise				
# Cruise-years	1997	1996	1995	1994	1993
4	0.916				
3	0.916	0.791			
2	0.916	0.809	0.679		
1	0.916	0.822	0.721	0.612	
0					0.595

Table 1: P(active) as a function of recency and frequency

is that those customers who took a repeat cruise in only one year have a lower value of p than those who took cruises in all three years, and therefore the fact that no cruise was taken in 1997 can be attributed more to their low probability of taking a cruise on any given year than to the possibility of their having become inactive. Looking across the table, we also observe that the probability of being active in 1998 is a decreasing function of recency — the longer the hiatus since the last cruise, the more likely it is that the customer is no longer active.

In Table 2, we report the number of discounted expected transactions (DET), as computed using (13) with a discount rate of 15%, as a function of recency and frequency.[6] Let us first consider those customers who took a cruise in 1997: while they all have the same probability of being alive in 1998, their DETs are an increasing function of frequency. This reflects the fact that those customers who have taken a cruise in each of the four years have a higher latent probability

[6] As we are modelling whether or not each customer made a cruise each year, and not how many cruises were taken each year, these numbers actually represent a lower-bound for DET.

p of taking a cruise in any year (while active) than those customers whose first repeat cruise occurred in 1997 (and therefore have a lower value of p). Looking at those customers whose last cruise occurred in 1996, we note that DET is an increasing function of frequency, even though the probability of being alive in 1998 is a decreasing function of frequency. Looking across the table, we also observe that DET is a decreasing function of recency; this follows naturally from the observation that the probability of being active in 1998 decreases with a longer hiatus since the last cruise.

	Year of Last Cruise				
# Cruise-years	1997	1996	1995	1994	1993
4	2.13				
3	1.67	1.44			
2	1.21	1.07	0.90		
1	0.76	0.68	0.60	0.51	
0					0.19

Table 2: DET as a function of recency and frequency ($d = 0.15$)

We observe exactly the same pattern when we consider the corresponding undiscounted numbers, as computed using (14) — see Table 3. In fact, the undiscounted numbers are a constant 2.43 times larger than the corresponding DETs.

	Year of Last Cruise				
# Cruise-years	1997	1996	1995	1994	1993
4	5.16				
3	4.05	3.50			
2	2.95	2.60	2.18		
1	1.84	1.65	1.45	1.23	
0					0.47

Table 3: Expected number of (remaining) transactions as a function of recency and frequency

The discounted numbers reported in Table 2 are central to any financial calculations of customer value. However, the undiscounted numbers reported in Table 3 may be more meaningful to many analysts seeking to understand the relationship between a customer's purchase history and his future purchasing behavior.

4 Discussion

We have developed a discrete-time analog of the Pareto/NBD model that can be used to answer standard customer-base analysis questions in settings where opportunities for transactions occur at discrete intervals. Using data for a cohort of customers from a cruise-ship company, we have demonstrated how the model can be used to compute a number of managerially relevant quantities, such as the probability that the customer is still active as well as an expectation of future purchasing. In examining these quantities, we have observed some interesting effects of past behavior (as summarized by recency and frequency) on predictions about future behavior.

There are three natural extensions to this basic model. First, as is the case with the Pareto/NBD model, the BG/BB model will need to be augmented by a model of purchase amounts when we are interested in the overall monetary value of each customer. A natural candidate would be the gamma-gamma mixture (Colombo and Jiang 1999) that Fader et al. (2004b) use in conjunction with the Pareto/NBD model. In situations where the data have been discretized for analysis and reporting purposes (as is the case for the above cruise-ship example), there is the possibility that more than one transaction could occur in each discrete time-interval. We could arrive at the monetary-value multiplier of DET (as required for any CLV calculation) by first modelling the number of transactions (conditional on the fact that at least one transaction occurred) and then multiply this by the average value per transaction. A logical model would be the shifted beta-geometric distribution (as used by Morrison and Perry (1970) and Fader and Hardie (2001) to model purchase quantity, conditional on purchase incidence).

Second, we may want to allow for a non-zero-order purchasing process at the individual level. A good starting point would be the "Brand Loyal Model" (Massy, Montgomery, and Morrison 1970). This would effectively be an extension of Morrison et al.'s (1982) Markov chain model of retail customer behavior at Merrill Lynch; an extension in which the "exit parameter" is allowed to be heterogeneous and estimated directly from the data (as opposed to being derived from other data sources). Finally, the model can be extended to include demographic variables and measures of marketing activity. Even when the last two extensions are undertaken, the BG/BB model in its basic form will still serve as an appropriate benchmark model that can be implemented at a very low cost (i.e., in a spreadsheet), and should be viewed as the right starting point for any customer-base analysis exercise in a noncontractual setting where opportunities for transactions occur at discrete intervals.

References

Berger, Paul D. and Nada I. Nasr (1998), “Customer Lifetime Value: Marketing Models and Applications,” *Journal of Interactive Marketing*, **12** (Winter), 17–30.

Berger, Paul D., Bruce Weinberg, and Richard C. Hanna (2003), “Customer Lifetime Value Determination and Strategic Implications for a Cruise-Ship Company,” *Journal of Database Marketing & Customer Strategy Management*, **11** (September), 40–52.

Colombo, Richard and Weina Jiang (1999), “A Stochastic RFM Model,” *Journal of Interactive Marketing*, **13** (Summer), 2–12.

Fader, Peter S. and Bruce G. S. Hardie (2001), “Forecasting Repeat Sales at CDNOW: A Case Study,” *Interfaces*, **31** (May–June), Part 2 of 2, S94–S107.

Fader, Peter S., Bruce G. S. Hardie, and Ka Lok Lee (2004a), “"Counting Your Customers" the Easy Way: An Alternative to the Pareto/NBD Model,” *Marketing Science*, forthcoming.

Fader, Peter S., Bruce G. S. Hardie, and Ka Lok Lee (2004b), “RFM and CLV: Using Iso-value Curves for Customer Base Analysis,” working paper.

Gupta, Sunil and Donald R. Lehmann (2003), “Customers as Assets,” *Journal of Interactive Marketing*, **17** (Winter), 9–24.

Libal, Barak, Das Narayandas, and Clive Humby (2002), “Towards an Individual Customer Profitability Model: A Segment-Based Approach,” *Journal of Service Research*, **5** (August), 69–76.

Massy, William F., David B. Montgomery, and Donald G. Morrison (1970), *Stochastic Models of Buying Behavior*, Cambridge, MA: The MIT Press.

Morrison, Donald G., Richard D. H. Chen, Sandra L. Karpis, and Kathryn E. A. Britney (1982), “Modelling Retail Customer Behavior at Merrill Lynch,” *Marketing Science*, **1** (Spring), 123–141.

Morrison, Donald G., and Arnon Perry (1970), “Some Data Based Models for Analyzing Sales Fluctuations,” *Decision Sciences*, **1** (3 & 4), 258–274.

Pfeifer, Phillip E., Mark E. Haskins, and Robert M. Conroy (2005), “Customer Lifetime Value, Customer Profitability, and the Treatment of Acquisition Spending,” *Journal of Managerial Issues*, forthcoming.

Schmittlein, David C., Donald G. Morrison, and Richard Colombo (1987), “Counting Your Customers: Who They Are and What Will They Do Next?” *Management Science*, **33** (January), 1–24.

Schmittlein, David C. and Robert A. Peterson (1994), “Customer Base Analysis: An Industrial Purchase Process Application,” *Marketing Science*, **13** (Winter), 41–67.