The Pareto/NBD is Not a Lost-for-Good Model

Peter S. Fader
www.petefader.com

Bruce G. S. Hardie
www.brucehardie.com†

December 2014

1 Introduction

In their widely cited “Return on Marketing” paper, Rust et al. (2004) draw on Jackson’s (1985a) lost-for-good versus always-a-share classification of customer relationships and associate the Pareto/NBD model (Schmittlein et al. 1987) with the lost-for-good setting. This association is repeated by Gupta and Zeithaml (2006) and continues in the literature (e.g., Conoor 2010, Ma and Büschken 2011, Romero et al. 2013). Reflecting more carefully on Jackson’s work, we feel that it is more appropriate to refer to the Pareto/NBD as a “leaky” always-a-share model; we develop this view below.

2 Jackson’s “Models of Behavior”

The lost-for-good versus always-a-share distinction was introduced by Jackson (1985a,b), and further developed in a CLV setting by Dwyer (1989). It is now a widely used scheme for classifying customer relationships.

We feel that some of the ideas behind Jackson’s classification have been lost over time, and we suspect that many using her classification have not read the original sources. So let us go back to the most detailed source.

†© 2014 Peter S. Fader and Bruce G.S. Hardie. This document can be found at <http://brucehardie.com/notes/031/>.
To quote Jackson (1985a, p. 13):

Two simplified but suggestive pictures of accounts’ behavior can help in understanding customers’ commitments and likely customer behavior. These pictures, or models, of behavior can be considered the end points of a spectrum of possible behavior by real customers; they are essentially extreme examples that define a wide range of possible behaviors. Actual accounts in real situations will occupy less extreme positions along the spectrum.

In the first model (called lost-for-good), an account is either totally committed to the vendor or totally lost and committed to some other vendor. In the second model (called always-a-share), the account has a lasting but less intensive tie to the vendor.

To elaborate (Jackson 1985a, pp. 13–14),

The lost-for-good model assumes that a customer repeatedly makes purchases from some product category over time. At any one time, the account is committed to only one vendor. The account faces very high costs of switching vendors, and consequently it changes suppliers only very reluctantly. The account is likely, though not certain, to remain committed to its current supplier.

The lost-for-good model assumes that if a customer does decide to leave a supplier, the account is lost forever — or, alternatively, that it is at least as difficult and costly for the vendor to win back such an account as it was to win the customer in the first place. The model’s name emphasizes the pain of losing a lost-for-good customer. The flip side is considerably more cheerful; once won, such a customer is likely to be won for a long time, though not necessarily forever.

[...]

The second model also assumes that the customer purchases repeatedly from some product category. It assumes, however, that buyers can maintain less intense commitments than they do in the lost-for-good model and that they can have commitments to more than one vendor at the same time. The account can easily switch part or all of its purchases from one vendor to another, and therefore it can share its patronage, perhaps over time, among multiple vendors.

In an examination of the quantitative implications of the lost-for-good model, Jackson (1985a, 18–21) proposes a Bernoulli process for relationship termination, which implies that the duration of the customer’s relationship
with the firm can be characterized by a geometric distribution. In one examination of the always-a-share setting, Jackson (1985a, 21–24) uses a Bernoulli process to describe the probability of a customer purchasing from the firm on any given purchase occasion.

- If the firm can make only one purchase per discrete time-period, this is equivalent to assuming that the customer’s purchasing from the focal firm across \( n \) discrete time intervals is distributed binomial.

- If purchasing can occur at any point in time, we need to specify a category interpurchase time distribution. The assumption of exponentially distributed category interpurchase times coupled with the above assumption of Bernoulli supplier choice implies that the customer’s purchasing of the focal firm’s products follows a Poisson distribution.

In any empirical setting, we must account for customer heterogeneity. This would suggest that the beta-geometric (BG) distribution (Fader and Hardie 2007, 2014) is the obvious benchmark model for characterizing customer behavior in a lost-for-good setting. Similarly, the beta-binomial (BB) distribution (Chatfield and Goodhardt 1970; Easton 1980) and the NBD (Ehrenberg 1959; Morrison and Schmittlein 1988) are the obvious benchmark models for characterizing behavior in an always-a-share setting.

3 Examining the Basic Always-a-Share Models

As we develop our case for not calling the Pareto/NBD a lost-for-good model, let us start by assessing the empirical performance of the two basic stochastic models of buyer behavior for always-a-share settings, the BB and NBD distributions.

3.1 The BB Model

Consider a major nonprofit organization located in the Midwestern United States that is funded in large part by donations from individuals. In 1995 the organization “acquired” 11,104 first-time supporters. As summarized in Table 1, we know whether or not each individual made a donation to the charity in each of the following six years.\(^1\)

We observe that an individual may make donation, do nothing for several years, and then make another donation. This is not what Jackson would call a lost-for-good setting; it is much more aligned with the always-a-share model. Based on the above discussion, the BB is the natural model for characterizing donor behavior in this dataset.

Table 1: Annual donation incidence for the 1995 cohort of first-time donors.

Fitting the BB distribution to these data, we see in Figure 1 that it provides a good fit to the actual number of 1995 first-time donors making repeat donations on 0, 1, 2, \ldots, 6 of the subsequent six years.

![Graph showing predicted and actual donation incidence](image)

**Figure 1:** Predicted (BB) versus actual frequency of (repeat) annual donation incidence.

In Figure 2 we compare the cumulative number of (repeat) donations over the six-year model calibration period and the subsequent five years with the BB model-based predictions.
Figure 2: Predicted (BB) versus actual cumulative number of donations.

We observe that the BB not only fails to track actual cumulative repeat annual donations in the six-year calibration period, but also deviates significantly from the actual donation trajectory over the subsequent five years. After 11 years, the BB model is over-forecasting by 20%. The degree of the problem is even more apparent when we look at the corresponding annual donation-incidence tracking plot (Figure 3). The BB always-a-share model predicts a constant level of repeat donation incidence, whereas it is actually decaying over time.

Figure 3: Predicted (BB) versus actual number of (repeat) donors each year.
3.2 The NBD Model

Consider the purchasing of music CDs at the online retailer CDNOW. During the first quarter of 1997, a total of 23,570 people made their first-ever purchase at the CDNOW website. (See Fader and Hardie (2001) for details of this dataset.) Focusing on a 1/10th sample of this group\(^2\), we have data on their initial (trial) and subsequent (repeat) purchases in the 39-week period ending September 30, 1997.

Figure 4 provides a graphical illustration of the patterns of purchasing observed in the CDNOW dataset, where \(\circ\) denotes the timing of the initial purchase, and \(\times\) denotes the timing of any repeat purchases made by each customer. We observe that a customer makes a purchase, does nothing for a long time, and then makes another. Reflecting on the product class, this is not appear to be consistent with Jackson’s lost-for-good model. Within Jackson’s classification scheme, it is an always-a-share setting. Based on the above discussion, the NBD is the natural model for characterizing the repeat-buying behavior in this dataset.

```
ID = 0001  \circ \times \times
ID = 0002  \circ \times
...           \circ
ID = 1178  \circ \times
ID = 1179  \circ
...           \circ
ID = 2356  \circ \times \times \times \times
ID = 2357  \circ
```

**Figure 4:** A graphical presentation of the CDNOW purchase histories.

Drawing on the analysis presented in Fader and Hardie (2004), we examine the performance of the NBD model when fitted to these repeat-purchasing data. In Figure 5 we compare the expected number of customers making \(0, 1, 2, \ldots\) repeat purchases in the period ending September 30, 1997 to the actual frequency distribution for that period. The NBD model provides a very good fit to the data when viewed in this static manner.

In Figure 6 we compare the expected total number of repeat transactions by the cohort of customers over the calibration period (January 1 to September 30) and the subsequent 39-week period, as predicted by the NBD model, with the actual repeat sales data. We observe that the NBD not only fails to track actual sales in the 39-week calibration period, but also deviates

Figure 5: Predicted (NBD) versus actual frequency of repeat transactions.

significantly from the actual sales trajectory over the subsequent 39 weeks. By the end of June 1998, the NBD model is over-forecasting by 24%.

Figure 6: Predicted (NBD) versus actual cumulative repeat transactions.

The degree of the problem is even more apparent when we look at the corresponding week-by-week repeat-transaction numbers (Figure 7). The sales figures rise through the end of week 12, as new customers continue to enter the cohort, but after that point it is a fixed group of 2,357 eligible buyers. The NBD predicts a constant level of repeat transactions, whereas in fact they are decaying over time, albeit with obvious deviations because of promotional activities and the holiday season.
4 An Alternative Classification of Relationships

A harsh reality for any marketer is that regardless of how wonderful their product or service is, or how creative their marketing activities are, the customer base of any company can be best viewed as a “leaky bucket” whose contents are continually dripping away. Customer needs and tastes change as their personal circumstances change over time, which leads them to stop purchasing from a given firm or even stop buying in the product category all together. In the end, they literally die.

The fundamental question is whether we are in a business setting where the loss (or “death”) of an individual customer is actually observed by the firm (e.g., the customer terminates their contract or fails to renew their fixed-term contract) or one where it is unobserved (i.e., “they just silently attrite” (Mason 2003, p. 55)). The key challenge in this latter setting is how to differentiate those customers who have ended their relationship with the firm (without informing it) from those who are simply in the midst of a long hiatus between transactions. (While we can never know for sure which of these two states a customer is in, we can use statistical models to make probabilistic statements.)

It is now standard to use the term “contractual” to characterize a relationship when the death of a customer is observed by the firm, and the term “noncontractual” to characterize a relationship where the death of a customer is unobserved by the firm.\footnote{This classification on the basis of the observability of customer death was proposed by Schmittlein et al. (1987). The term contractual has long been used as a label for settings where the death of customer is observed. To the best of our knowledge, the use of the term} The two datasets considered above

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure7.png}
\caption{Predicted (NBD) versus actual weekly repeat transactions.}
\end{figure}
clearly come from noncontractual settings.

At first glance, the contractual versus noncontractual distinction seems the same as Jackson’s lost-for-good versus always-a-share classification, with a number of researchers treating them as equivalent labels (e.g., Bauer et al. 2003; Calciu 2009; Kumar and Reinartz 2012). However, this is not the case. We identify two differences: the first, which is slightly pedantic, concerns the role of competition, while the second concerns the recognition of the leaky bucket phenomenon.

- Jackson’s view of a lost-for-good relationship is one of serial monogamy: “an account is either totally committed to the vendor or totally lost and committed to some other vendor” (Jackson 1985a, p. 13). No such assumption is made in the contractual classification. In some contractual settings, it is only possible to use one service provider (e.g., standard residential utilities). However, in other contractual settings it is not uncommon for a customer to have relationships with several different providers at any point in time (e.g., credit cards). Similarly, whereas the notion of competitive offerings lies at the heart of the definition of an always-a-share setting, no assumptions about the competitive context are made when defining noncontractual settings. When not purchasing from the focal firm, the customer could be purchasing from a competitor or simply not be making a category purchase. In short, what exactly is happening with respect to competitors is immaterial to the definition of a noncontractual setting.

- Jackson’s classification only recognizes the reality of customer attrition in the lost-for-good case; it is assumed to be non-existent in always-a-share settings. In contrast, it is central to the contractual versus noncontractual classification. Admittedly, Jackson (1985a, p. 13) sees always-a-share as a limiting case — “[a]ctual accounts in real situations will occupy less extreme positions along the spectrum.” Our problem with the always-a-share label is that it does not acknowledge the empirical reality of customer attrition and had led researchers to develop models that ignore this phenomenon (e.g., Kumar et al. 2008; Rust et al. 2011; Venkatesan and Kumar 2004; Venkatesan et al. 2007).

noncontractual as a label for settings where the death of a customer is not observed was introduced to the marketing literature by Werner Reinartz in his Ph.D research (Reinartz 1999).

4We also note that the “subscription” versus “repertoire” classification proposed by Sharp et al. (2002) is very similar to Jackson’s lost-for-good versus always-a-share classification.

5Recognizing the pedantic nature of this point, we are willing to treat lost-for-good and contractual as effectively being equivalent. This is reinforced by the fact that the basic probability model of customer behavior in contractual settings, the BG distribution, is a nature extension of the probabilistic process proposed by Jackson herself. However, we are not willing to do this for always-a-share and noncontractual.
We prefer the contractual versus noncontractual classification to the lost-for-good versus always-a-share classification as it is agnostic when it comes to the issue of competition and, more importantly, brings to the fore the issue of the observability of customer attrition, which has obvious implications for the types of marketing metrics and statistical models the analyst should use when analyzing a given customer database.

5 Extending the Basic Always-a-Share Models

While both the BB and NBD models do an excellent job of capturing the distribution of repeat transactions in the model calibration period (Figures 1 and 5), they fail to track the evolution of sales over time (Figures 2 and 6). Their “straight line” predictions of period-by-period buying (Figures 3 and 7) follow from the always-a-share assumption of stationarity. (The BB model assumes that each individual buys from the focal supplier with a given probability that is constant over time for any given individual but varies across individuals. Likewise, the NBD model assumes that each individual buys from the focal supplier at a given rate that is constant over time for any given individual but varies across individuals.)

The decline in period-by-period transactions, as observed in Figures 3 and 7 and presented in a stylized manner in Figure 8, indicates that the buying process in a noncontractual setting is not stationary.

![Figure 8: A stylized representation of repeat-buying by a cohort of buyers in a noncontractual setting.](image)

Schmittlein et al. (1987) capture this nonstationarity by assuming that a customer’s relationship with the firm has two phases: he is “alive” for some period of time, then “dies”. The unobserved (and unobservable) time at which each customer dies is treated as-if random from the perspective of the analyst. While each customer is alive, their buying behavior is assumed to be stationary.
The logic of this latent attrition solution is illustrated in Figure 9. Each customer is buying according to their own stationary process (i.e., at their own constant rate). Ignoring the effect of random purchasing around their means, individual customers purchase the product at steady but different underlying rates, giving us the straight dashed lines associated with customers A–E. The first decline in total sales (the solid line) is associated with the (unobserved) death of customer A, followed by that of D then C. We expect customers B and E to eventually die as customers; it is simply the case that they have not done so by the end of the observation period.

![Figure 9: A stylized representation of a latent-attrition model.](image)

For settings where transactions can occur at any point in time, Schmittlein et al. propose that the time to (unobserved) death be characterized by the Pareto distribution of the second kind, and that buying while alive be characterized by the NBD. This gives us the Pareto/NBD model. Figure 10, taken from Fader et al. (2005), examines the performance of the Pareto/NBD model when fitted to the CDNOW data. We observe that the Pareto/NBD model accurately tracks the actual (cumulative) repeat sales trajectory in both the 39-week calibration period and the 39-week forecast period, under-forecasting by less than 2% at the end of week 78.

For settings where transactions are best recorded in discrete time, Fader et al. (2010) propose that the time to (unobserved) death be characterized by the BG distribution, and that buying while alive be characterized by the BB distribution. This gives us the BG/BB model. Figure 11, taken from Fader et al. (2010), examines the performance of the BG/BB model when fitted to the donation data considered above. We note that the BG/BB model accurately tracks the actual cumulative number of donations in both the six-year calibration period and the five-year forecast period, under-forecasting at 2006 by a mere −0.65%. 

Figure 10: Predicted (Pareto/NBD) versus actual cumulative repeat transactions.

Figure 11: Predicted (BG/BB) versus actual cumulative number of donations.
Recall that, following the logic of Jackson, the BB and NBD distributions are the natural baseline models for always-a-share settings. The key insight from the analyses presented in Section 3 is that such “pure” always-a-share models (i.e., ones that do not accommodate latent attrition) will tend to do a poor job of tracking repeat purchasing in a longitudinal validation period. In contrast, the BG/BB and Pareto/NBD models do an excellent job of capturing the “leakage” of customers observed in practice. As such, we feel that it is appropriate to call them “leaky” always-a-share models.

6 Reflections on Rust et al. (2004)

We started this note by referring to the work of Rust et al. (2004). Let us now consider what could be viewed as the “offending text” (pp. 112–113):

Customer retention historically has been treated according to two assumptions (Jackson 1985). First, the “lost for good” assumption uses the customer’s retention probability (often the retention rate in the customer’s segment) as the probability that a firm’s customer in one period is still the firm’s customer in the following period. Because the retention probability is typically less than one, the probability that the customer is retained declines over time. The implicit assumption is that customers are “alive” until they “die,” after which they are lost for good. Models for estimating the number of active customers have been proposed for relationship marketing (Schmittlein, Morrison, and Columbo 1987), customer retention (Bolton 1998), and CLV (Reinartz 1999).

The second assumption is the “always a share” assumption, in which customers may not give any firm all of their business. Attempts have been made to model this by a “migration model” (Berger and Nasr 1998; Dwyer 1997). The migration model assigns a retention probability as previously, but if the customer has missed a period, a lower probability is assigned to indicate the possibility that the customer may return. Likewise, if the customer has been gone for two periods, an even lower probability is assigned. This is an incomplete model of switching because it includes purchases from only one firm.

In one scenario (consistent with the lost-for-good assumption) when the customer is gone, he or she is gone. This approach systematically understates CLV to the extent that it is possible for customers to return. In another scenario (consistent with the migration model), the customer may leave and return. In this scenario, customers may be either serially monogamous or
polygamous (Dowling and Uncles 1997), and their degrees of loyalty may vary or even change. We can model the second (more realistic) scenario using a Markov switching-matrix approach.

The first thing to note is that while they do not explicitly state that the Pareto/NBD is a lost-for-good model, the implication is there, and it is clear how other researchers who have not read Jackson’s work incorrectly conclude that the Pareto/NBD model is a lost-for-good model, citing Rust et al.

The model developed by Rust et al. is based on a Markov switching matrix, which is a generalization of the always-a-share Bernoulli process proposed by Jackson, one in which the probability of purchasing from the focal firm at any (discrete) point in time is dependent on the supplier used in the previous period. (The diagonal elements of this matrix are repeat-buying probabilities.) As it does not account for the leaky bucket phenomenon, we know that (unlike, say, the BG/BB or Pareto/NBD) such a model will overestimate purchasing over time and therefore systematically overestimate CLV.

Later on in their empirical analysis (pp. 120–121), Rust et al. note that “previously, we proposed that some models of CLV that do not account for customers returning systematically underestimate CLV ....” To examine this, they set the off-diagonal elements of their brand-switching model to zero, which is equivalent to Jackson’s assumption of geometrically-distributed customer-relationship durations in lost-for-good settings. Not surprisingly, they find that “the lost-for-good model provides a systematic underestimation” of customer value.

Note that such a “you always buy from the same vendor until you die” story of customer behavior is clearly inconsistent with Pareto/NBD model (and its discrete-time analog, the BG/BB). The Pareto/NBD does not assume that “an account is either totally committed to the vendor or totally lost and committed to some other vendor” (Jackson 1985a, p. 13). Rather, it allows customers to come and go and come back again, as we would expect in an always-a-share setting. What it also allows for is customer attrition. Now, any model of buyer behavior that allows for latent attrition is assuming that once a customer has died, they remain dead. But it should be clear that such latent attrition is not what Jackson meant by lost-for-good.

Jackson explicitly states that the lost-for-good versus always-a-share classifications are represent extreme positions along a continuum. It should be clear from the above discussion that the Pareto/NBD is much closer to the always-a-share extreme than it is to the lost-for-good extreme. If we are to work within Jackson’s classification scheme, the Pareto/NBD has to

---

6There is no resurrection of the dead. Any dead customers “reacquired” at a later date are treated as new customers.
be viewed as a “leaky” always-a-share model. Any researcher claiming that
the Pareto/NBD is a lost-for-good model and will therefore systematically
understate CLV is simply wrong.

References

Bauer, Hans H., Maik Hammerschmidt and Matthias Braehler (2003), “The
Customer Lifetime Value Concept and Its Contribution to Corporate
Valuation,” Yearbook of Marketing and Consumer Research, Volume 1,

Calciu, Mihai (2009), “Deterministic and Stochastic Customer Lifetime Value
Models. Evaluating the Impact of Ignored Heterogeneity in Non-contractual
Contexts,” Journal of Targeting, Measurement and Analysis for Market-
ing, 17 (4), 257–271.


Conoor, Sandeep S. (2010), Customer-Base Analysis in Noncontractual Set-
tings, unpublished Ph.D dissertation, Northwestern University.

Dwyer, F. Robert (1989), “Customer Lifetime Valuation to Support Market-

OMEGA, 8 (1), 63–69.

Statistics, 8 (March), 26–41.

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. and Bruce G. S. Hardie (2004), “Illustrating the Performance
of the NBD as a Benchmark Model for Customer-Base Analysis.”
<http://brucehardie.com/notes/005/> 

Fader, Peter S. and Bruce G. S. Hardie (2007), “How To Project Customer
Retention,” Journal of Interactive Marketing, 21 (Winter), 76–90.

Fader, Peter S. and Bruce G. S. Hardie (2014), “A Spreadsheet-Literate Non-
Statistician’s Guide to the Beta-Geometric Model.”
<http://brucehardie.com/notes/032>


