Markus Wübben & Florian v. Wangenheim

Instant Customer Base Analysis: Managerial Heuristics Often "Get It Right"

Recently, academics have shown interest and enthusiasm in the development and implementation of stochastic customer base analysis models, such as the Pareto/NBD model and the BG/NBD model. Using the information these models provide, customer managers should be able to (1) distinguish active customers from inactive customers, (2) generate transaction forecasts for individual customers and determine future best customers, and (3) predict the purchase volume of the entire customer base. However, there is also a growing frustration among academics insofar as these models have not found their way into wide managerial application. To present arguments in favor of or against the use of these models in practice, the authors compare the quality of these models when applied to managerial decision making with the simple heuristics that firms typically use. The authors find that the simple heuristics perform at least as well as the stochastic models with regard to all managerially relevant areas, except for predictions regarding future purchases at the overall customer base level. The authors conclude that in their current state, stochastic customer base analysis models should be implemented in managerial practice with much care. Furthermore, they identify areas for improvement to make these models managerially more useful.

Keywords: customer base analysis, heuristics, relationship marketing, direct marketing, customer management

Consider a marketing executive at a catalog retailer who faces the following challenges: First, she wants to distinguish customers in the customer base who are likely to continue buying from the firm (active customers) from those who are likely to defect or from those who have already defected (inactive customers). This information should help (1) identify profitable, inactive customers who should be reactivated; (2) remove inactive, unprofitable customers from the customer base; and (3) determine active customers who should be targeted with regular marketing activities, such as new catalogs or mailings. Second, she wants to generate transaction forecasts for individual customers to identify the company's future 10% best customers, or to compute customer lifetime value (CLV). Such information should help her target those groups with perks, differential mailing frequencies, and loyalty program offerings. Third, she wants to predict the purchase volume of the entire customer base to make provisions for capacity planning, to compute the firm's customer equity, and to know when customer acquisition efforts need to be strengthened.

Markus Wübben is a doctoral student (e-mail: markus.wuebben@wi.tum.de), and Florian v. Wangenheim is Chair of Services and Technology Marketing (e-mail: florian.wangenheim@wi.tum.de), TUM Business School, Technische Universität München. The authors thank Katherine N. Lemon, Bruce Hardie, Pete Fader, Dan Goldstein, and the anonymous *JM* reviewers for their valuable comments during the preparation of this manuscript. This work was supported by the German Research Foundation (DFG).

For the executive, the central problem in successfully coping with these tasks is that the time at which a customer defects from the firm is unobservable. The customer may have been disenchanted with the purchased product or the provider and now buys at a different supplier, the customer may have moved to another city, or the customer may have even passed away. This phenomenon exists for most service providers that operate in noncontractual settings: For example, when a customer purchases from a catalog retailer, walks off an aircraft, checks out of a hotel, or leaves a retail outlet, the firm has no way of knowing whether and how often the customer will conduct business in the future (Reinartz and Kumar 2000).

In contrast, in a contractual setting, the buyer–seller relationship is governed by a contract, which often predetermines not only the length but also the usage pattern of the relationship (e.g., telephone and Internet "flat-rate" services, magazine subscriptions). In this context, hazard regression or logistic regression models (Bolton 1995; Li 1995) provide promising approaches in determining the probability that a customer will still be with the firm at a particular future time. In the noncontractual setting, the state-of-the-art approach in determining the activity and future purchase levels of a customer is the Pareto/NBD model (Schmittlein, Morrison, and Columbo 1987; Schmittlein and Peterson 1994). The Pareto/NBD model has recently been employed in several studies (Fader, Hardie, and Lee 2005a, b; Ho, Park, and Zhou 2006; Krafft 2002; Reinartz and Kumar 2000, 2003), and its implementation has been recommended on an even larger scale (Balasubra-

ISSN: 0022-2429 (print), 1547-7185 (electronic)

manian et al. 1998; Jain and Singh 2002; Kamakura et al. 2005; Rust and Chung 2006). Recently, Fader, Hardie, and Lee (2005a) introduced the BG/NBD model, which is a variant of the Pareto/NBD model but is much easier to implement and estimate. Both models are attractive because they (1) make forecasts of individuals' future purchase levels and (2) operate on past transaction behavior. More precisely, they operate solely on the frequency and recency information of a customer's past purchase behavior. The Pareto/NBD model has an additional feature; for each customer, it yields the probability that he or she is still active.

In light of increased calls for closer cooperation between marketing academics and practitioners, it must be of concern for academics that these models have not found their way into managerial practice. Instead, a survey by Verhoef and colleagues (2002) shows that simple heuristics are still commonly applied.

Given the time and money costs associated with implementing complex stochastic models in managerial practice, the marketing executive will be convinced to make use of the academic methods only when their superiority is clearly demonstrated on the aggregate level and, even more important, on the individual customer level. However, practitioners are not the only ones who would benefit from such insights. For research, it is important to know the circumstances under which the predictions of these models can be trusted to produce accurate forecasts for future implementation of these models in, for example, CLV research (e.g., Reinartz and Kumar 2000, 2003).

Few studies have compared the performance of complex versus noncomplex models for customer purchase behavior and lifetime value prediction. Donkers, Verhoef, and De Jong (2007) find that using complex methods instead of simple models for CLV prediction in a contractual setting (insurance company) does not substantially improve predictive accuracy. In a semicontractual context, Borle, Singh, and Jain (2008) find that a simple RFM (recency, frequency, and monetary value) model performs as well as the Pareto/NBD model that includes monetary value (Schmittlein and Peterson 1994) in predicting CLV. They also propose a hierarchical Bayesian model that works better than both the Pareto/NBD and the RFM models in the semicontractual setting. However, none of the studies on the stochastic models (Fader, Hardie, and Lee 2005a; Schmittlein, Morrison, and Columbo 1987; Schmittlein and Peterson 1994) have validated their predictions on the individual customer level in a noncontractual setting using multiple data sets from different industries. The current research aims to fill this gap.

In what follows, we briefly cover heuristics in managerial practice and provide an introduction to the Pareto/NBD and BG/NBD models. We then describe the data sets from three different industries on which we performed our validation. Next, we present the results of the predictive performance of the models versus the simple management heuristics. Finally, we discuss our findings and offer recommendations for using customer base analysis models in academic research and managerial practice.

Heuristics in Managerial Practice

Heuristics are often mentioned in managerial literature on customer relationship management or database marketing (e.g., Hughes 2006; Novo 2004). For example, Blattberg, Getz, and Thomas (2001) suggest the recency–sales matrix, which is a slightly modified variant of the hiatus heuristic. The most prominent example of a heuristic in practical customer management is probably the RFM framework, which was introduced by Alden's catalog company in the 1920s to value customers to decide which customer should receive a catalog (Roel 1988).

The diffusion of simple heuristics can be explained by the opportunity costs that managers experience when investing time and effort in managerial decision making. Rieskamp and Hoffrage (2008) show that under opportunity costs, people tend to make decisions quickly. A framework for explaining contingent decision behavior—in this case, choosing between complex methods and simple heuristics—focuses on the accuracy and cognitive effort characterizing the available strategies (Payne, Bettman, and Johnson 1993). The basic hypothesis of the effort/accuracy framework is that the strategy used to make a decision represents a balancing of the goals of being as accurate as possible and conserving limited cognitive resources. Creyer, Bettman, and Payne (1990) find that when the goal is to maximize accuracy rather than minimize effort, people acquire more information, take more time, are more alternative based, and are ultimately more accurate. However, given opportunity costs, managers aim to minimize time and effort and therefore resort to well-established, simple heuristics. As Richard Abdoo, chair and chief executive officer of Wisconsin Energy Corporation, notes, "As we move to a deregulated marketplace, we don't have this slow process of hearings and review and two years to make a decision. We now have to make decisions in a timely manner. And that means that we process the best information that's available and infer from it and use our intuition to make a decision" (Hayashi 2001, p. 61).

There are reasons to believe that simple heuristics may work better than more complex strategies for various types of tasks, even though they often require less information and computation (Gigerenzer, Todd, and The ABC Research Group 1999). Support for Grigerenzer's work is manifold (Bröder 2000, 2003; Bröder and Schiffer 2003; Lee and Cummins 2004; Newell et al. 2004; Newell and Shanks 2003; Newell, Weston, and Shanks 2003; Rieskamp 2006; Rieskamp and Hoffrage 1999; Rieskamp and Otto 2006). A survey conducted by Jagdish Parikh (discussed in Buchanan and O'Connell 2006, p. 40) shows that executives "used their intuitive skills as much as they used their analytical abilities, but they credited 80% of their successes to instinct." Especially in direct marketing, experienced managers are likely to be accurate (Morwitz and Schmittlein 1998) because this environment is characterized by two properties that are essential for learning to occur: repetition and feedback (Camerer and Johnson 1991; Goldberg 1968). "Repetition" refers to the repeated occurrence of the task, and "feedback" means that the outcome of the manager's decision is easily observed and evaluated. The previously described decisions to distinguish between active and inactive customers and between high- and low-value customers both are highly repetitive and offer feedback.

A survey conducted in May 2002 by executive search firm Christian and Timbers reveals that 45% of corporate executives now rely more on instinct than on facts and figures in running their businesses (Bonabeau 2003). Nevertheless, even if facts and figures are used, managers still rely on intuition and long-standing methods. A survey by Verhoef and colleagues (2002) on 228 database marketing companies shows that cross-tabulation and RFM analysis are the most popular methods for response modeling. In the context of the current study, at least two of the three companies (airline and apparel retailer) whose customer bases are analyzed apply simple recency-of-last-purchase (hiatus) analysis to distinguish active from inactive customers. For example, the managers of the airline who have "expert" knowledge of their customer base informed us that the cutoff time was nine months. This finding is in line with an article in the *New York Times* (Wade 1988) that illustrates the use of the hiatus heuristic in frequent-flier programs. For future purchase–level determination, average past purchase behavior is often employed as a simple predictor for future behavior. The managers of focal firms also confirmed this. A series of short telephone interviews with eight people responsible for customer management within their firms revealed that the hiatus heuristics was applied in all companies to determine active and inactive customers. For determining future best customers, somewhat varying approaches were applied, but number of past purchases was always a central variable (e.g., in an RFM-type approach).

Stochastic Customer Base Analysis Models

Both the Pareto/NBD and the BG/NBD models were developed to model repeat-buying behavior in a setting in which customers buy at a steady (albeit stochastic) rate and eventually become inactive at some unobserved time. The information they operate on consists solely of customers' past purchase behavior. More precisely, for each customer, the models operate on three values ($X = x$, t, T), where $X = x$ is the number of purchases made in time frame $(0, T]$, with the last purchase occurring at time t, where $0 < t \leq T$. In addition, the models must be calibrated on the customer base to which they are applied. This calibration process yields several model parameters that describe the purchase and dropout process of the analyzed customer base.

Pareto/NBD

The Pareto/NBD model builds on the assumption that purchases follow Ehrenberg's (1988) NBD model, whereas dropout events follow a Pareto distribution of the second kind. More precisely, the Pareto/NBD model assumptions are as follows:

1. Individual Customer
   - Poisson purchases: While active, each customer makes purchases according to a Poisson process with purchase rate λ.
   - Exponential lifetime: Each customer remains active for a lifetime, which has an exponentially distributed duration with dropout rate μ.
2. Heterogeneity Across Customers
   - Individuals' purchase rates distributed gamma: The purchasing rate λ for the different customers is distributed according to a gamma distribution across the population of customers.
   - Individuals' dropout rates distributed gamma: The customers' dropout rates μ are distributed according to a gamma distribution across the population of customers.
   - Rates λ and μ are independent: The purchasing rates λ and the dropout rates μ are distributed independently of each other.

Among other things, the Pareto/NBD model yields the following information:

- $P(\text{Active}|X = x, t, T)$ is the probability that a random customer with purchase pattern ($X = x$, t, T) (whose individual purchase rate and dropout rate may be unknown) is active at some time T (see Schmittlein, Morrison, and Colombo 1987, Equations 11, 12, and 13).
- $E(X^*|X = x, t, T, T^*)$ is the expected number of transactions X^* of a random customer with purchase pattern ($X = x$, t, T) (and unknown individual purchase rate and dropout rate) in time $(T, T + T^*]$ (see Schmittlein, Morrison, and Colombo 1987, Equation 22).

BG/NBD

There is only one assumption in the BG/NBD model that differs from the assumptions of the Pareto/NBD model. Whereas the Pareto timing model assumes that dropout of a customer can occur anytime, the BG/NBD model assumes that dropout occurs only directly after purchases. This slight change greatly reduces the complexity of the model because a beta-geometric (BG) model can be used to represent the dropout phenomena instead of the exponential gamma (Pareto) model. More precisely, the BG/NBD model assumptions are as follows:

1. Individual Customer
   - Poisson purchases: While active, each customer makes purchases according to a Poisson process with purchase rate λ.
   - Geometric lifetime: Each customer remains active for a lifetime, which is distributed over the number of transactions according to a (shifted) geometric distribution with dropout probability p.
2. Heterogeneity Across Customers
   - Individuals' purchase rates distributed gamma: The purchase rate λ for the different customers is distributed according to a gamma distribution across the population of customers.
   - Individuals' dropout probabilities distributed beta: The customers' dropout probabilities p for different customers is distributed according to a beta distribution across the population of customers.
   - Rates λ and p are independent: The purchase rates λ and dropout probabilities p are distributed independently of each other.

Among other things, the BG/NBD model yields the following information:

- E(X*|X = x, t, T, T*) is the expected number of transactions X* of a random customer with purchase pattern (X = x, t, T) (and unknown individual purchase rate and dropout rate) in time (T, T + T*] (see Fader, Hardie, and Lee 2005a, Equation 10).

The BG/NBD model also includes the expression P(Active| X = x, t, T) to compute the probability that a customer is active at some time T, but the application of this expression is limited to customers whose individual purchase rate λ and dropout probability *p* is known. However, determining the dropout probability *p* for an individual is virtually impossible.

Data

We conducted our study on three different data sets from three different industries. The first data set comes from an apparel retailer and covers 46,793 customers and their purchases from January 2003 through August 2004. We based our analysis on a cohort of 2330 customers who began their buyer–seller relationship with the apparel retailer in the last week of January 2003. Thus, for this cohort, the available data cover the initial and repeat purchases for each customer over a period of 80 weeks. To calibrate the models, we used repeat-purchase data for the 2330 customers over the first 40 weeks of the 80-week period, leaving a 40-week holdout period to validate the models.

The second data set comes from a major global airline and covers 146,961 customers and their purchases from January 1999 through December 2002. The available data only provided aggregated quarterly transactions for each customer and did not include the exact purchase dates. Our analysis of this data set focused on a cohort of 2891 customers who conducted their initial purchase from the airline in the first quarter of 1999. For this cohort, we chose a calibration period of eight quarters (January 1999–December 2000), leaving eight quarters for the holdout period (January 2001–December 2002).

The third data set covers customers of the online CD retailer CDNOW. The data track 23,570 customers and their purchases from January 1997 through June 1998 (78 weeks), all of whom initiated their first purchase at CDNOW in the first quarter of 1997. Fader and Hardie (2001) already used this data in multiple studies. More precisely, we used the 2357 customer cohort available on Bruce Hardie's Web site (see Fader, Hardie, and Lee 2005a). The calibration and holdout periods are 39 weeks each. Detailed descriptive statistics of all three data sets appear in Table 1.

Analysis

Given the lack of empirical analysis for the superiority of the considered academic methods in determining active customers and forecasting future purchase levels, the following analyses try to shed light on this open question. First, we analyze how well the hiatus heuristic, which is used by the managers of the firms whose customer bases we analyze in this article, performs in comparison with the Pareto/NBD P(Active) facility. Second, we analyze how well the Pareto/NBD and BG/NBD models forecast future purchase behavior for both the individual customer and the customer base as a whole. Although aggregated sales forecasts are important statistics in terms of, for example, capacity planning or customer equity computation, we specifically focus on the forecast performance for the individual customer. This stems from the notion that there must be a decent individual customer purchase–level forecast for proper computation of metrics, such as CLV (Reinartz and Kumar 2000, 2003) or customer value segment classification (e.g., gold, silver, and bronze segments). Picking up on this idea, not only do we present mere performance measures for individuals' forecasts, but in a third analysis, we

TABLE 1
Descriptive Statistics

	Airline	Apparel	CDNOW
Sample size (n)	2891	2330	2357
Available time frame	16 quarters	80 weeks	78 weeks
Time split (estimation/holdout[a])	8/8	40/40	39/39
Available time units	Quarters	Weeks/months/quarters	Weeks
Zero repeaters in estimation periods	193	371	1411
Zero repeaters in holdout periods	1376	395	1673
Zero repeaters in estimation and holdout periods	163	184	1218
Number of purchases in estimation periods	31,479	10,855	2457
Number of purchases in holdout periods	23,033	11,351	1882
Average number of purchases per customer in estimation periods (SD)	10.88/customer (15.988)	4.658/customer (5.412)	1.04/customer (2.190)
Average number of purchases per customer in holdout periods (SD)	7.967/customer (16.810)	4.871/customer (5.598)	.798/customer (2.057)
Average T (SD)	4.393 quarters (3.006)	25.15 weeks (14.21)	6.845 weeks (10.731)

[a]Holdout period length was varied from 1 to max(holdout periods) in the analyses.

also show how well the models perform in identifying a company's future 10% (20%) best customers.

Parameter Estimation

Both the Pareto/NBD and the BG/NBD models need to be calibrated on the customer base to which they are applied. The Pareto/NBD model has four parameters (r, α, s, β), where (r, α) represent the shape and scale parameters of the gamma distribution that determines the distribution of the purchase rates across individuals of the customer base and (s, β) represent the scale and shape parameters of the gamma distribution that determines the distribution of the dropout rates across individuals. The BG/NBD model holds four model parameters (r, α, a, b) as well, where (r, α) (as in the NBD/Pareto model) determine the shape and scale of the purchase rate gamma distribution and (a, b) represent the shape parameters of a beta distribution that determines the distribution of the dropout probabilities across individuals of the customer base. For both models, we used a maximum likelihood approach under MATLAB to estimate the model parameters. Tables 2 and 3 report each cohort's parameters for the Pareto/NBD and BG/NBD models.

The Pareto/NBD and BG/NBD model parameters computed for the airline and CDNOW data sets are reasonable. According to the Pareto/NBD model, an average airline customer initiates 1.9877 transactions per quarter and remains active for 7.83 quarters.[1] An average CDNOW customer initiates .0523 transactions per week (one purchase every 19.12 weeks) and remains active for 19.26 weeks. According to the BG/NBD model, an average airline customer initiates 2.110 transactions per quarter and remains active for 4.34 quarters.[2] An average CDNOW customer initiates .0549 transactions per week (one purchase each 18.21 weeks) and purchases from the company for 73.95 weeks.

For the apparel data set, both the Pareto/NBD and the BG/NBD models compute notable results. Although the purchasing rate of an average customer (.1190 purchases per week [one purchase every 8.40 weeks] for the Pareto/NBD model and .1198 purchases per week or one purchase every 8.34 weeks for the BG/NBD model) is reasonable, the lifetime of an average customer is exceptionally long. According to the Pareto/NBD model, an average customer remains active for 909.09 weeks (~17.48 years). According to the BG/NBD model, an average customer remains active for 684.2 weeks (~13.15 years). In other words, the models predict an average apparel customer to be ultimately loyal. This effect will be reflected in very high P(Active) values of the apparel customers.

Determining Active and Inactive Customers

If the Pareto/NBD model is used for customer activity determination, each customer's P(Active) value is computed on the basis of the customer's purchase pattern in the observation period (i.e., estimation period). However, for the continuous P(Active) values to be useful in managerial application, a cutoff threshold $c_{P(Active)}$ (decision boundary) must be determined. Customers whose P(Active) value is greater than or equal to $c_{P(Active)}$ are classified as active, and customers whose P(Active) value is less than $c_{P(Active)}$ are classified as inactive. More precisely, for the Pareto/NBD model, given a cutoff threshold $c_{P(Active)}$ and a customer

[1]Within the Pareto/NBD model, r/α represents the number of purchases of an average customer in one time unit, and s/β represents the dropout rate of an average customer per time unit. The lifetime of an average customer is exponentially distributed with parameter s/β and has an expected value of 1/(s/β). Therefore, according to the estimated parameters, an average CDNOW customer remains active for 1/(.0519) = 19.26 weeks.

[2]Within the BG/NBD model, r/α represents the number of purchases of an average customer in one time unit, and an average customer remains active until time τ, which is exponentially distributed with parameter pλ and has an expected value of 1/(pλ), given that λ = r/α and p = a/(a + b). Therefore, according to the estimated parameters, an average CDNOW customer remains active for 1/(.0549 × .2463) = 73.95 weeks.

TABLE 2
Results of the Pareto/NBD Maximum Likelihood Estimation

	r	α	r/α	s	β	s/β	Log-Likelihood
Apparel	1.0954	9.2029	.1190	1.0885	973.7829	.0011	–31338.7
Airline	1.4304	.7196	1.9877	2.5086	19.6408	.1277	–2150.2
CDNOW	.5533	10.5776	.0523	.6061	11.6650	.0519	–9595.0

TABLE 3
Results of the BG/NBD Maximum Likelihood Estimation

	r	α	r/α	a	b	a/(a + b)	Log-Likelihood
Apparel	1.0592	8.8371	.1198	.0324	2.6243	.0122	–31336.6
Airline	1.15186	.545781	2.11048	.456637	3.73439	.108956	–2238.34
CDNOW	.2426	4.4135	.0549	.7931	2.4260	.2463	–9582.43

with purchase pattern (X = x, t, T), the customer is classified according to the following:

- $P(\text{Active})_T \geq c_{P(\text{Active})} \Rightarrow$ Customer is classified as active, and
- $P(\text{Active})_T < c_{P(\text{Active})} \Rightarrow$ Customer is classified as inactive.

To validate the classifications, we use the holdout period according to the following scheme: If a customer has made at least one purchase in the holdout period, he or she is considered "active"; if the customer has not purchased in the holdout period, he or she is considered "inactive." This scheme induces four possible classification outcomes based on whether a customer has or has not been correctly classified as active or inactive.

Likewise, for the hiatus heuristic, there needs to be a cutoff threshold c_{hiatus} below which customers are classified as active and above which customers are classified as inactive. In other words, if a customer has not purchased for more than a time span of length c_{hiatus}, he or she is considered inactive; otherwise, he or she is considered active. More precisely, let (X = x, t, T) be a customer's purchase pattern and c_{hiatus} be a cutoff threshold. Then,

- $T - t < c_{\text{hiatus}} \Rightarrow$ Customer is classified as active, and
- $T - t \geq c_{\text{hiatus}} \Rightarrow$ Customer is classified as inactive.

We obtain the same four possible classification outcomes as in the P(Active) case.

For the Pareto/NBD model, a "natural" choice for the cutoff threshold $c_{P(\text{Active})}$ could be .5, which is in line with the work of Reinartz and Kumar (2000) on the Pareto/NBD model and the classification literature (Sharma 1996). Helsen and Schmittlein (1993) also use .5 in the prediction of purchase events in survival analysis.

For the hiatus heuristic, airline and apparel firm mangers informed us that customers were considered inactive if they had not purchased from the firm for more than nine months. For the CDNOW data set, we did not have access to this information. Furthermore, we do not know whether CDNOW uses the hiatus heuristic at all. Given that online firms operate in fast-moving markets, we decided to use a hiatus length of six months, which should match the circumstances of an online retailer.

However, neither the managers' chosen hiatus nor a P(Active) threshold of .5 may necessarily be optimal thresholds in terms of overall correctly classified customers. We observe how far the managers' chosen hiatus and the P(Active) threshold of .5 deviate from their "optimal" values and the effect of this difference in terms of the classification performance. First, we show how well the hiatus heuristic with the threshold determined by managers' expert knowledge distinguishes the active from the inactive customers in comparison with a P(Active) analysis with a natural cutoff threshold of .5. Second, we show how sensitive the classification performance is to the choice of the thresholds.

Table 4 shows the results of the first analysis. The hiatus heuristic performs better (in terms of overall correctly classified customers) than the P(Active) facility in two of the three cases; the P(Active) facility performs only slightly better on the CDNOW data set. Even more notable, the P(Active) facility fails to classify any of the inactive customers in the apparel data set correctly, whereas the hiatus heuristic classifies 47.84% of the inactive customers in the cohort correctly. This suggests that the optimal cutoff threshold $c_{P(\text{Active})}$ for this cohort may deviate considerably from .5. What about the other P(Active) cutoff thresholds $c_{P(\text{Active})}$ and hiatus heuristic cutoff thresholds c_{hiatus}? Are these optimal? If not, what are the optimal values? This is the subject of our next analysis.

As we mentioned previously, we consider a cutoff threshold of $c_{P(\text{Active})}$ or c_{hiatus} optimal if it maximizes the percentage of overall correctly classified active and inactive customers of a cohort using our classification procedure.[3] Our algorithm for finding the optimal cutoff thresholds simply iterates over the domain of valid cutoff thresholds. More precisely, for the Pareto/NBD model, the algorithm is as follows:

[3]Although it is also possible to optimize for a maximum of correctly classified active or inactive customers depending on the purchase of the marketing action (i.e., reactivation or elimination), we believe that our approach to maximize for the sum of correctly classified active and inactive customers is reasonable because it combines both approaches.

TABLE 4
P(Active) Versus Hiatus Heuristic

	Airline		Apparel		CDNOW	
	Three Quarters Hiatus	$P(\text{Active})_{.5}$	Nine Months Hiatus	$P(\text{Active})_{.5}$	Six Months Hiatus	$P(\text{Active})_{.5}$
Inactive, correctly classified (%)	84.1569	84.6656	47.8478	.0000	82.6659	87.3881
Active, correctly classified (%)	69.9667	64.3564	89.8708	100.0000	63.5965	53.0703
Overall correctly classified (%)	76.7208	74.0228	82.7467	74.8972	77.1319	77.4289
Inactive but classified active (%)	15.8430	15.3343	52.1521	100.0000	17.3340	12.6120
Active but classified inactive (%)	30.0332	35.6435	10.1291	.0000	36.4034	46.9300
Overall incorrectly classified (%)	23.2791	25.9771	17.2532	25.1072	22.8680	22.5710

Notes: Numbers represent percentage hit rate of the active/inactive class. Overall hit rate percentages are weighted according to the distribution of active/inactive customers in the data set.

- For $c_{P(Active)} \in \{0, \ldots, 1\}$, choose $c_{P(Active)}$ so that the sum of correctly classified active and inactive customers is maximized.

For the hiatus heuristic, the algorithm is as follows:

- For $c_{hiatus} \in \{0, \ldots, \infty\}$, choose c_{hiatus} so that the sum of correctly classified active and inactive customers is maximized.

Table 5 presents an overview of the analysis results. Surprisingly, for all three cohorts, the hiatus heuristic performs slightly better than the more complex Pareto/NBD model. Indeed, the optimal cutoff thresholds for the hiatus heuristic (4 quarters, 40 weeks, and 23 weeks) are close to the managers' and our chosen threshold (3 quarters, 39 weeks, and 26 weeks). If the optimal cutoff thresholds were used instead of expert knowledge, it would result in a marginal gain of only .8302% (airline), .1713% (apparel), and 1.0607% (CDNOW) in terms of overall correctly classified customers.

With respect to the optimal P(Active) thresholds, only for the CDNOW cohort, the optimal value of .44 is close to the natural cutoff value of .5, and its use improves performance by only .17% in terms of overall correctly classified customers. For both the apparel and the airline cohorts, the optimal P(Active) thresholds of .67 and .21, respectively, deviate substantially from the natural cutoff threshold of .5, and the gain from using the optimal threshold rather than the natural cutoff threshold is a considerable 7.9848% and 1.2452%, respectively, in terms of overall correctly classified customers. However, if we carefully examine the analysis results for the apparel cohort, we observe that when the optimal P(Active) of .67 is used, virtually none of the inactive customers are correctly classified (1.009%). The reason is the (estimated) exceptionally long lifetime of an average apparel customer that we already briefly covered in the parameter estimation section. This property causes P(Active) values to be close to 1 for almost all apparel customers and purchase patterns; few customers had P(Active) values in the range of .67–.9. Because we optimized for maximizing the overall correctly classified statistic and given the high percentage of apparel retailer repurchasers, it is more favorable for the optimization algorithm to classify as many active customers correctly as possible. We explore the reasons for these unrealistically high P(Active) values in the "Discussion" section. Nevertheless, the natural cutoff value of .5 may not necessarily be close to its optimal value, as the apparel and airline data sets show. This makes the interpretation of the P(Active) values counterintuitive if a P(Active) value is considered a customer's propensity to repurchase.

Predicting Future Purchase Levels

In this analysis, we focus on the Pareto/NBD and BG/NBD models' capability to predict future purchase levels cumulative for the cohort as a whole and on an individual customer basis.[4] More precisely, we benchmarked both models' predicated number of transactions against a simple management heuristic: Every customer continues to buy at his or her past mean purchase frequency.

We compare the performance of the Pareto/NBD model, the BG/NBD model, and the simple heuristic in predicting cumulated purchases on the basis of the mean absolute percentage error (MAPE) (Leeflang et al. 2000). To measure the performance on the individual customer level, we computed the (root) mean square errors ([R]MSE) for each customer over the predicted and actual transactions in the holdout period (Leeflang et al. 2000). We also computed the mean (R)MSE (median [R]MSE), which represents the mean (median) of all individual customer (R)MSE. Table 6 presents the results of the analysis.

On all three cohorts, the stochastic models outperform the simple heuristic on both the individual and the aggregate levels, and the Pareto/NBD and BG/NBD models perform almost identically. Although the stochastic models deliver decent results on the aggregated level, as the MAPE statistic shows, the results are split on the individual level. The models show poor performance in terms of the mean (R)MSE over all customers, but at least for 50% of the cohorts, the stochastic models predict future purchases precisely, as the median (R)MSE statistics show.

Identifying Future Best Customers

The previous analysis shows that the stochastic models under consideration precisely predict future purchases for

[4]See Schmittlein, Morrison, and Colombo (1987, Equation 22) and Fader, Hardie, and Lee (2005a, Equation 10).

TABLE 5
P(Active) Versus Hiatus Heuristic Using Optimal Thresholds

	Airline		Apparel		CDNOW	
	Hiatus Heuristic	P(Active)	Hiatus Heuristic	P(Active)	Hiatus Heuristic	P(Active)
Optimal cutoff threshold	4 quarters	.21	40 weeks	.67	23 weeks	.44
Inactive, correctly classified (%)	77.109	78.489	46.581	1.009	85.4752	86.3120
Active, correctly classified (%)	77.954	72.343	90.337	99.535	60.3800	56.2863
Overall correctly classified (%)	77.551	75.268	82.918	82.832	78.1926	77.5986
Inactive but classified active (%)	22.891	21.511	53.419	98.991	14.5247	13.6879
Active but classified inactive (%)	22.046	27.657	9.663	.465	39.6199	43.7133
Overall incorrectly classified (%)	22.448	24.731	17.082	17.167	21.8074	22.4014

Notes: Numbers represent percentage hit rate of the active/inactive class. Overall hit rate percentages are weighted according to the distribution of active/inactive customers in the data set.

TABLE 6
Summary Statistics for Purchase-Level Prediction

	Airline			Apparel			CDNOW		
Statistic	**BG/NBD**	**Pareto/ NBD**	**Heuristic**	**BG/NBD**	**Pareto/ NBD**	**Heuristic**	**BG/NBD**	**Pareto/ NBD**	**Heuristic**
MAPE	14.582	25.049	28.187	11.3486	12.1242	9.54866	12.15	11.49	55.69
Mean MSE	79.2123	78.2362	95.8921	4.8598	4.8857	5.02916	2.59	2.57	4.89
Median MSE	2.79177	2.3345	7.15625	1.32253	1.3309	1.38375	.061	.027	0
Mean RMSE	4.16536	4.04075	5.24889	1.6338	1.63334	1.63818	.785	.754	1.02
Median RMSE	1.67086	1.5279	2.67512	1.15001	1.15365	1.17633	.248	.166	0

50% of the individuals in a cohort. Nevertheless, this figure tells us little about the applicability of individuals' transaction forecasts in a managerial context. Often, companies implement disproportionate marketing investment strategies on the basis of a customer value rating because it is common for a small percentage of customers to account for a large percentage of revenues and profits (Mulhern 1999). For example, an airline might want to prioritize high-value customers in an overbooking occasion and deny boarding to lower-value customers. Likewise, apparel retailers may want to invite their best customers to special events (i.e., fashion shows), and an online CD store might be interested in sending sample CDs of new albums and/or artists to its best customers.

In this analysis, we assume that a company offers two levels of treatment: "best-customer" treatment and "normal-customer" treatment (Malthouse and Blattberg 2005). Optimally, a customer should receive the best-customer treatment if he or she belongs to the future best customers. Past best customers may not necessarily belong to the group of future best customers (Wangenheim and Lentz 2005). Under the assumption that a customer's future value cannot be estimated perfectly, a company can make two types of classification errors. First, a future best customer may be classified as a future normal customer and thus may be denied the treatment he or she "deserves." This misclassified and, therefore, mistreated customer may spread negative word of mouth or even switch the provider completely. Second, a future normal customer may be misclassified as a future best customer, leading to extra and unjustified spending of scarce marketing resources.

If the complex models under consideration are used to identify future best customers, they need to perform better than a simple management heuristic. This is the subject of our next analysis. More precisely, we try to identify the future 10% (20%, respectively) best customers in the customer bases (in terms of future number of transactions) on the basis of the Pareto/NBD and BG/NBD models' individual customer purchase–level prediction. This classification is benchmarked against yet another simple management judgment rule: The past 10% (20%, respectively) best customers in a customer base will also be the future 10% (20%) best customers.

The results of the analysis appear in Tables 7 and 8. In line with the intention to identify a company's future best customers, the "correctly-classified-as-high" statistic is the one of interest. This statistic represents the fraction of the future best customers who actually have been classified as

TABLE 7
The 10% Best Future Customers

	Airline			Apparel			CDNOW		
Statistic	**BG/NBD**	**Pareto/ NBD**	**Heuristic**	**BG/NBD**	**Pareto/ NBD**	**Heuristic**	**BG/NBD**	**Pareto/ NBD**	**Heuristic**
High, correctly classified (%)	61.09	61.09	57.84	63.15	63.15	70.15	53.92	54.18	61.51
Low, correctly classified (%)	95.22	95.22	94.85	95.63	95.63	94.49	91.08	91.13	86.22
Overall correctly classified (%)	91.76	91.76	90.93	92.18	92.18	91.80	85.06	85.15	82.22
Incorrectly classified high (%)	38.90	38.90	42.15	36.84	36.84	29.84	8.91	8.86	13.77
Incorrectly classified low (%)	4.77	4.77	5.14	4.36	4.36	5.50	46.07	45.81	38.48
Overall incorrectly classified (%)	8.23	8.23	9.06	7.81	7.81	8.19	14.93	14.84	17.77

TABLE 8
The 20% Best Future Customers

	Airline			Apparel			CDNOW		
Statistic	**BG/NBD**	**Pareto/ NBD**	**Heuristic**	**BG/NBD**	**Pareto/ NBD**	**Heuristic**	**BG/NBD**	**Pareto/ NBD**	**Heuristic**
High, correctly classified (%)	64.26	63.40	63.60	67.13	67.13	73.72	61.25	61.69	71.78
Low, correctly classified (%)	89.99	90.60	89.27	91.18	91.18	89.11	84.16	84.33	72.80
Overall correctly classified (%)	84.81	85.12	84.05	86.09	86.09	85.49	77.51	77.76	72.50
Incorrectly classified high (%)	35.73	36.59	36.39	32.86	32.86	26.27	15.83	15.66	27.19
Incorrectly classified low (%)	10.01	9.39	10.72	8.81	8.81	10.88	38.74	38.30	28.21
Overall incorrectly classified (%)	15.18	14.87	15.94	14.37	14.42	15.49	22.48	22.23	27.49

future best customers by the models. The complementary figure is the "incorrectly-classified-as-low" statistic. It represents the fraction of the future best customers who have falsely been classified as future low customers. In four of the six cases, the heuristic performs better that the stochastic models in terms of correctly-classified-as-high customers. Only for the airline data set do the stochastic models outperform the heuristic. We also report the Gini coefficients in Table 9. Instead of focusing only on the 10% and 20% best customers, this measure also includes the models' performance in classifying less valuable customers. The 10% and 20% best-customers statistic and the Gini coefficient provide complementary information; a model can be good at classifying only the 10% or 20% best customers but may be less effective at recognizing the less valuable customers. However, for both the airline and the apparel data sets, the simple heuristic has smaller Gini coefficients than the stochastic models (i.e., it performs better over all customer groups). Because of data constraints, we could not compute the Gini coefficient for the CDNOW data set.

Discussion

Summary of Findings

Many researchers have outlined the usefulness and applicability of the Pareto/NBD model and, more recently, of the BG/NBD model. For example, Krafft (2002), Reinartz and Kumar (2000, 2003), and Wu and Chen (2000) employ the Pareto/NBD model in their work on customer base analysis and CLV prediction. Given this, it is all the more surprising that the current study is the first (1) to validate comprehensively these models' predictions about both the individual customer and the customer base level in noncontractual settings using multiple data sets and (2) to benchmark these models against simple management heuristics that practitioners commonly deploy.

Recall that by applying the stochastic models, we intended to assist marketing executives in (1) determining active and inactive customers, (2) generating individual customer transaction forecasts to identify the company's future best customers, and (3) determining the future purchase volume of the customer base as a whole. According to our analysis, the applicability of the focal stochastic models seems to be limited to determining the purchase volume of the customer base as a whole. In our analysis, the stochastic models showed superiority over a simple management heuristic. For determining a company's active and inactive customers and for predicting a company's future best customers, the management heuristics we applied worked as well as the stochastic models.

We performed a series of additional analyses, which, for the sake of brevity, cannot be displayed here. Specifically, we varied the holdout period length from 1 to max(holdout period length) (i.e., from 1 to 8 quarters for the airline data set and from 1 to 40 weeks for the apparel data set). At the same time, we held the estimation period constant at 8 quarters and 40 weeks, respectively. Furthermore, we varied the length of the time frame chosen for the apparel data set to have weekly, monthly, and quarterly data available.

The results revealed the same pattern as we reported previously. With regard to active/nonactive classification, the hiatus heuristic classifies an additional 1.5% of the customers correctly compared with the Pareto/NBD model for the apparel data set (averaged over 53 analyses) and an additional 1.34% for the airline data set (averaged over 8 analyses). In terms of identifying the 10% or 20% best customers of the firms, the complex methods are (very slightly)

TABLE 9
Gini Coefficients: Best Customer Classification

Data Set	Pareto/NBD	BG/NBD	Heuristic
Airline	.043330	.042874	.041942
Apparel	.071798	.071764	.053581

Notes: The CDNOW data set does not provide enough information to compute Gini coefficients.

superior in determining only the 10% best customers of the airline. For the best 20% customers and for all analyses related to the apparel data, the hiatus is at least equal and, in most cases, somewhat better than the complex models. (The detailed results of all these analyses, as well as the apparel data set itself, are available on request.)

The sensitivity analysis shows that the length of the holdout period and the aggregation level has little impact on our results. However, it could be argued that this is too little time to classify a customer as inactive. Indeed, depending on their individual purchase rate, customers who have not purchased within the holdout period may well have purchased after that period. Therefore, they are not ultimately inactive (or "dead"), but from a managerial perspective, the lengths of the holdout periods represent a reasonable marketing investment planning horizon. Therefore, it is of hardly any interest to managers whether a customer purchases after the planning horizon. However, even when we restrict our analyses to customers who are active in the holdout period and thus can be certain that our active/inactive assessment is correct, the simple heuristics at least match the performance of the more complex models.

For the apparel data set, we obtained unrealistic lifetime estimates and P(Active) values when we applied the Pareto/NBD model. (A mathematical demonstration and proof that explains why this happens for some data sets are available on request.) In essence, this will be the case when there is a relatively high number of customers in the data set who conducted their last transaction shortly before the end of the observation window and when there are high T values (i.e., an estimation time frame, such that the longer the time frame considered, the higher is the T [weeks, months, quarters]).

Managerial Implications

The finding that P(Active) classifications do not outperform the simple hiatus heuristic in determining active and inactive customers is a devastating result for what has been called the "key result of the NBD/Pareto model" (Reinartz and Kumar 2000, p. 21). As we already mentioned, in at least two cases (the apparel retailer and airline), managers are using the simple hiatus heuristic to determine customer (in)activity in their companies. Their expert assessment coincides almost perfectly with the optimal hiatus length that we determined in our analysis. This is an indication that managerial judgment may well act as a decent estimate of customer (in)activity. Consequently, researchers need to stop recommending the Pareto/NBD model to managers and fellow researchers for this purpose.

For identifying future best customers, the admittedly simple approach of assuming that past best customers are future best customers and the stochastic models deliver unconvincing results, even though we can correct Malthouse and Blattberg's (2005) 20-55 rule to a more positive figure. If we use the Pareto/NBD model or the BG/NBD model, approximately 33% (or less) of the top 20% customers are misclassified, making it a 20-33 rule.[5] Nevertheless, if disproportionate marketing investment decisions are made on the basis of the focal models (i.e., valuable customers receive better service, more perks, and so on, than less valuable customers), these are likely to be inefficient. Scarce marketing resources would be spent on less valuable customers whose behavior does not justify this best-customer treatment. However, many valuable customers who are falsely classified as less valuable customers would not receive the treatment they deserve. Being disenchanted, these customers could switch to a competitor or spread negative word of mouth (Malthouse and Blattberg 2005; Mitchell 2005).

It appears that the managerial applicability of the Pareto/NBD and BG/NBD models is limited to customer equity computation. For this purpose, to both managers and academics, we recommend using the BG/NBD model because of its relatively easier implementation, faster computation, and superior performance compared with the Pareto/NBD model and the simple heuristic. Given the increasing interest in valuing firms on the basis of customer equity (Gupta, Lehman, and Stuart 2004), the stochastic models are good candidates for valuing customer bases in noncontractual settings. Thus, it would be worthwhile to benchmark these models against the approach that Gupta, Lehman, and Stuart (2004) use.

Limitations and Further Research

We determined the top 10% (20%) customers solely on the basis of the number of transactions because the monetary value of transactions was not available in the data. A more managerial top 10% (20%) customer analysis would employ the monetary value of customers. However, we believe that if monetary value were incorporated into our computations, overall model comparison results would not shift substantially; in the end, the Pareto/NBD and BG/NBD models were designed to predict future transactions, not monetary value. Incorporating monetary value would only introduce an additional source of distortion.

For research, much more work still needs to be done. An appealing characteristic of the stochastic models described initially is that they work only on recency and frequency purchase information. The prediction error may partially be explained by the lack of attitudinal information, such as customer satisfaction, repurchase intention, or commitment. However, this does not explain why the stochastic models are not better than models that work on even less and simpler information. Therefore, the question arises whether the information currently used should be substituted by other or augmented by additional information. For example, Wangenheim and Lentz (2005) show that trend in revenues (i.e., the slope of revenue regressed on time) is an important predictor of a customer's life-cycle pattern and improves the accuracy of CLV predictions.

Further research should also address a more general question in predicting future customer purchase patterns: How much purchase information is needed to make predictions about future buyer behavior? For example, in the CDNOW data set (Fader and Hardie 2001), more than 50% of the customers in the data set had not made any purchases since the initial trial (and thus have [0, 0, T] purchase pat-

[5]Our analysis confirms Malthouse and Blattberg's (2005) 80-15 rule.

terns). Can any model, as sophisticated as it may be, make reliable forecasts for a customer who has conducted only one transaction with a supplier? Malthouse and Blattberg (2005) examine the effect of the length of the prediction period on the accuracy of the predictions, and Schmittlein and Peterson (1994) examine how many periods and customers should be included during estimation of the Pareto/NBD model parameters, but to the best of our knowledge, no work has addressed the question of how many transactions a customer needs to have conducted before reliable forecasts can be made.

Another notable aspect of our study is that the heuristics the firms used worked astonishingly well. Thus, it would be a worthwhile course for further research to examine how such heuristics emerge in the context of customer management and customer relationship management and how such knowledge can be integrated into relationship management solutions.

Conclusion

This article examines the performance of what have frequently been called state-of-the-art models in customer activity determination and purchase-level prediction in noncontractual settings. To validate these models, we not only used metrics and methods recommended in the statistical literature but also simulated the implementation of those models in managerial practice. However, we find no clear evidence for the superiority of these models for managerially relevant decisions in customer management compared with simple methods that our industry partners used.

As academics, when we lament about practitioners' resistance to using advanced research methods developed in academic research, we too easily forget that model validation in a strict statistical sense is not equivalent to model validation in the spirit of managerial relevance. Although the standard statistical tests are necessary for gaining acceptance in the academic community, they represent a necessary, but not a sufficient, condition for gaining acceptance in the managerial world. Academics must not merely present models' good statistical fit but also suggest how they improve managerial decision making to convince practitioners to adopt them. In the end, this will lead to better cooperation between academics and practitioners. Other examples in the literature on customer purchase predictions show that under certain circumstances, relatively simple heuristics are outperformed by specially developed complex models (Borle, Singh, and Jain 2008). With regard to the Pareto/NBD and BG/NBD models, it is clear that these models carry potential for managerial use in customer management, but it is yet to be shown for which distinct managerial decision they show superior performance.

REFERENCES

Balasubramanian, Sridha, Sunil Gupta, Wagner Kamakura, and Michel Wedel (1998), "Modeling Large Data Sets in Marketing," *Statistica Neerlandica*, 52 (3), 303–324.

Blattberg, Robert C., Gary Getz, and Jacquelyn S. Thomas (2001), *Customer Equity: Building and Managing Relationships as Valuable Assets*. Cambridge, MA: Harvard Business School Press.

Bolton, Ruth N. (1998), "A Dynamic Model of the Duration of the Customer's Relationship with a Continuous Service Provider: The Role of Satisfaction," *Marketing Science*, 17 (1), 45–65.

Bonabeau, Eric (2003), "Don't Trust Your Gut," *Harvard Business Review*, 81 (5), 116–23.

Borle, Sharad, Siddharth Singh, and Dipak Jain (2008), "Customer Lifetime Value Measurement," *Management Science*, 54 (1), 100–112.

Bröder, Arndt (2000), "Assessing the Empirical Validity of the 'Take-the-Best' Heuristic as a Model of Human Probabilistic Inference," *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 26 (5), 1332–46.

——— (2003), "Decision Making with the 'Adaptive Toolbox': Influence of Environmental Structure, Intelligence, and Working Memory Load," *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 29 (4), 611–25.

——— and Stefanie Schiffer (2003), "Take the Best Versus Simultaneous Feature Matching: Probabilistic Inferences from Memory and Effects of Representation Format," *Journal of Experimental Psychology: General*, 132 (2), 277–93.

Buchanan, Leigh and Andrew O'Connell (2006), "A Brief History of Decision Making," *Harvard Business Review*, 84 (1), 32–41.

Camerer, Colin F. and Eric J. Johnson (1991), "The Process-Performance Paradox in Expert Judgment: How Can Experts Know So Much and Predict So Badly?" in *Toward a General Theory of Expertise: Prospects and Limits*, K. Anders Ericsson and Jacqui Smith, eds. Cambridge, UK: Cambridge University Press, 195–217.

Creyer, Elizabeth H., James R. Bettman, and John Wayne Payne (1990), "The Impact of Accuracy and Effort Feedback and Goals on Adaptive Decision Behavior," *Journal of Behavioral Decision Making*, 3 (1), 1–16.

Donkers, Bas, Peter C. Verhoef, and Martijn de Jong (2007), "Modeling CLV: A Test of Competing Models in the Insurance Industry," *Quantitative Marketing and Economics*, 5 (2), 163–90.

Ehrenberg, Andrew S.C. (1988), *Repeat-Buying, Theory and Applications*, 2d ed. London: Griffin.

Fader, Peter S. and Bruce G.S. Hardie (2001), "Forecasting Repeat Sales at CDNOW: A Case Study," *Interfaces*, 31 (Part 2 of 2), 94–107.

———, ———, and Ka Lok Lee (2005a), "'Counting Your Customers the Easy Way: An Alternative to the Pareto/NBD Model," *Marketing Science*, 24 (2), 275–85.

———, ———, and ——— (2005b), "RFM and CLV: Using Iso-Value Curves for Customer Base Analysis," *Journal of Marketing Research*, 42 (November), 415–30.

Gigerenzer, Gerd, Peter M. Todd, and The ABC Research Group (1999), *Simple Heuristics That Make Us Smart*. New York: Oxford University Press.

Goldberg, Lewis R. (1968), "Simple Models or Simple Processes? Some Research on Clinical Judgments," *American Psychologist*, 23 (7), 483–96.

Gupta, Sunil, Donald R. Lehmann, and Jennifer A. Stuart (2004), "Valuing Customers," *Journal of Marketing Research*, 41 (February), 7–18.

Hayashi, Alden M. (2001), "When to Trust Your Gut," *Harvard Business Review*, 79 (2), 59–65.

Helsen, Kristiaan and David C. Schmittlein (1993), "Analyzing Duration Times in Marketing: Evidence for the Effectiveness of Hazard Rate Models," *Marketing Science*, 12 (4), 395–414.

Ho, Teck-Hua, Young-Hoon Park, and Yong-Pin Zhou (2006), "Incorporating Satisfaction into Customer Value Analysis:

Optimal Investment in Lifetime Value," *Marketing Science*, 25 (3), 260–77.

Hughes, Arthur M. (2006), *Strategic Database Marketing*, 3d ed. New York: McGraw-Hill.

Jain, Dipak and Siddhartha S. Singh (2002), "Customer Lifetime Value Research in Marketing: A Review and Future Directions," *Journal of Interactive Marketing*, 16 (2), 34–47.

Kamakura, Wagner, Carl. F. Mela, Asim Ansari, Anand Bodapati, Pete Fader, Raghuram Iyengar, Prasad Naik, Scott Neslin, Baohong Sun, Peter C. Verhoef, Michel Wedel, and Ron Wilcox (2005), "Choice Models and Customer Relationship Management," *Marketing Letters*, 16 (3–4), 279–91.

Krafft, Manfred (2002), *Kundenbindung und Kundenwert*. Heidelberg: Physica-Verlag.

Lee, Michael D. and Tarrant D.R. Cummins (2004), "Evidence Accumulation in Decision Making: Unifying the 'Take the Best' and the 'Rational' Models," *Psychonomic Bulletin & Review*, 11 (2), 343–52.

Leeflang, Peter, Dick R. Wittink, Michel Wedel, and Philippe A. Naert (2000), *Building Models for Marketing Decisions*. Boston: Kluwer Academic.

Li, Shaomin (1995), "Survival Analysis," *Marketing Research*, 7 (Fall), 17–23.

Malthouse, Edward and Robert Blattberg (2005), "Can We Predict Customer Lifetime Value?" *Journal of Interactive Marketing*, 19 (1), 2–16.

Mitchell, Alan (2005), "Who Knows the Relative Worth of Customers?" *Precision Marketing*, 17 (49), 12.

Morwitz, Vicki G. and David C. Schmittlein (1998), "Testing New Direct Marketing Offerings: The Interplay of Management Judgment and Statistical Models," *Management Science*, 44 (5), 610–28.

Mulhern, Frank (1999), "Customer Profitability Analysis: Measurement, Concentration, and Research Directions," *Journal of Interactive Marketing*, 13 (1), 25–40.

Newell, Ben R., Tim Rakow, Nicola J. Weston, and David R. Shanks (2004), "Search Strategies in Decision Making: The Success of 'Success,'" *Journal of Behavioral Decision Making*, 17 (2), 117–37.

——— and David R. Shanks (2003), "Take the Best or Look at the Rest? Factors Influencing 'One-Reason' Decision Making," *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 29 (1), 53–65.

———, Nicola J. Weston, and David R. Shanks (2003), "Empirical Tests of a Fast-and-Frugal Heuristic: Not Everyone 'Takes-the-Best,'" *Organizational Behavior & Human Decision Processes*, 91 (1), 82–96.

Novo, Jim (2004), *Drilling Down: Turning Customer Data into Profits with a Spreadsheet*, 2d ed. St. Petersburg, FL: The Drilling Down Bookstore.

Payne, John Wayne, James R. Bettman, and Eric J. Johnson (1993), *The Adaptive Decision Maker*. Cambridge, UK: Cambridge University Press.

Reinartz, Werner J. and V. Kumar (2000), "On the Profitability of Long-Life Customers in a Noncontractual Setting: An Empirical Investigation and Implications for Marketing," *Journal of Marketing*, 64 (October), 17–35.

——— and ——— (2003), "The Impact of Customer Relationship Characteristics on Profitable Lifetime Duration," *Journal of Marketing*, 67 (January), 77–99.

Rieskamp, Jörg (2006), "Perspectives of Probabilistic Inferences: Reinforcement Learning and an Adaptive Network Compared," *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 32 (6), 1355–70.

——— and Ulrich Hoffrage (1999), "When Do People Use Simple Heuristics and How Can We Tell?" in *Simple Heuristics That Make Us Smart*, Gerd Gigerenzer, Peter M. Todd, and The ABC Research Group, eds. New York: Oxford University Press, 141–68.

——— and ——— (2008), "Inference Under Time Pressure: How Opportunity Costs Affect Strategy Selection," *Acta Psychologica*, forthcoming.

——— and Phillip E. Otto (2006), "SSL: A Theory of How People Learn to Select Strategies," *Journal of Experimental Psychology: General*, 135 (May), 207–236.

Roel, Raymond (1988), "Direct Marketing's 50 Big Ideas," *Direct Marketing*, 50 (May), 45–52.

Rust, Roland and Tuck S. Chung (2006), "Marketing Models of Service and Relationships," *Marketing Science*, 25 (6), 560–80.

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

——— and Robert A. Peterson (1994), "Customer Base Analysis: An Industrial Purchase Process Application," *Marketing Science*, 13 (1), 41–67.

Sharma, Subhash (1996), *Applied Multivariate Techniques*. New York: John Wiley & Sons.

Verhoef, Peter C., Penny N. Spring, Janny C. Hoekstra, and Peter S.H. Leeflang (2002), "The Commercial Use of Segmentation and Predictive Modeling Techniques for Database Marketing in the Netherlands," *Decision Support Systems*, 34 (4), 471–81.

Wade, Betsy (1988), "Mileage Points Can Fly Away," *The New York Times*, (November 20), (accessed January 14, 2008), [available at http://query.nytimes.com/gst/fullpage.html?res=940DE1DC1F3DF933A15752C1A96E948260].

Wangenheim, Florian and Patrick Lentz (2005), "Customer Portfolio Analysis: Applying Financial Risk and Volatility Measures to Customer Segmentation and Risk-Adjusted Lifetime Value Determination," (October), (accessed August 17, 2007), [available at http://ssrn.com/abstract=782064].

Wu, Couchen and Hsiu-Li Chen (2000), "Counting Your Customers: Compounding Customer's In-Store Decisions, Interpurchase Time and Repurchasing Behavior," *European Journal of Operational Research*, 127 (1), 109–119.