Intern. J. of Research in Marketing 28 (2011) 281–294

Contents lists available at SciVerse ScienceDirect

Intern. J. of Research in Marketing

journal homepage: www.elsevier.com/locate/ijresmar

Will the frog change into a prince? Predicting future customer profitability

Roland T. Rust [a], V. Kumar [b,*], Rajkumar Venkatesan [c]

[a] 3451 Van Munching Hall, University of Maryland, College Park, MD, 20742, USA
[b] 35 Broad Street, J. Mack Robinson School of Business, Georgia State University, Atlanta, GA, 30303, USA
[c] Darden School of Business, P.O. Box 6550, Charlottesville, VA, 22906-6500, USA

ARTICLE INFO

Article history:
First received in April 26, 2010 and was under review for 4½ months
Available online 12 September 2011

Area Editor: Harald Van Heerde

Keywords:
Dynamic customer profitability
Customer lifetime value
Efficient marketing
Business to business marketing
Customer relationship management
Simulation model

ABSTRACT

More and more companies have customer databases that enable them to analyze customer profitability over time. These companies often seek to determine the most important customers as indicated by their current or historical profitability and focus attention on them. Focusing on profitable customers can result in more efficient use of marketing resources, but this approach neglects the fact that customers can evolve over time. Some customers begin as low-profit customers but eventually develop into high-profit customers. Others may start out as high-profit customers but become unprofitable over time. Previous efforts to predict future profitability have been relatively unsuccessful, with relatively simple, naïve models often performing just as well as or better than more sophisticated ones. Our paper presents a new approach to predicting customer profitability in future periods that performs significantly better than naïve models. We estimate the models on data from a high-tech company in a business-to-business context and validate the models' predictive ability on a holdout sample.

We show that a model based on simulation of customer futures provides large improvements over naïve extrapolation of average profits. By using the simulation model to select customers, ROI from marketing efforts is projected to increase by 58%.

1. Introduction

In a popular Grimm's fairy tale, a frog is vastly underestimated by a young princess who does not realize that he will eventually be transformed into a prince. Like the frog, some customers may be badly underestimated in terms of their profitability (and some may be overestimated). Businesses could benefit from knowing how many of their "frog" customers are really "princes" in disguise.

1.1. The management scenario

Let us consider a typical management scenario. XYZ company, a business-to-business (B2B) marketer, directs marketing actions (e.g., sales calls, promotional mailings, telephone contacts, relationship-building visits) to individual customers. The firm also keeps a database that includes a record of these actions, along with each customer's record of sales and profitability. The firm realizes that some customers may become more important over time, whereas others may become less important. The firm also knows, from previous experience and academic research, that focusing more marketing effort on customers who are likely to increase in profitability will yield better results (just like watering a plant that grows in fertilized soil will be more effective than watering a plant growing in soil that is not fertilized). Therefore, the firm would like to identify the customers whose business is likely to grow so that it can allocate more marketing efforts toward them. Likewise, the firm would like to identify those customers whose business is likely to decline so that it can focus less attention on them.

Many businesses have become aware that targeting profitable customers can make marketing spending more efficient (Mulhern, 1999; Zeithaml, Rust, & Lemon, 2001; Kumar, Venkatesan, Bohling, & Beckmann, 2008), an approach that is increasingly feasible because of the proliferation of historical customer databases. It is even better to target the customers who *will be* profitable (Reinartz & Kumar, 2003). Accurate estimates of future profitability would allow firms to make better marketing resource allocation decisions for individual customers (Bolton, Lemon, & Verhoef, 2004). Future profitability is an appropriate metric for resource allocation because it is forward looking and includes both future revenues and costs of serving the customer (Venkatesan & Kumar, 2004). It is best practice for firms to consider customers' future profitability and responsiveness to marketing when allocating resources. In this study, we focus on accurately predicting one of the inputs into the resource allocation issues, future customer profitability. A more sophisticated modeling framework that allows heterogeneous or customer-specific marketing response coefficients is required to consider both customers' level of future profits and their responsiveness to marketing in making

* Corresponding author at: J. Mack Robinson School of Business, Georgia State University, USA. Tel.: +1 404 413 7590; fax: +1 832 201 8213.
E-mail addresses: rrust@rhsmith.umd.edu (R.T. Rust), vk@gsu.edu (V. Kumar), Venkatesanr@darden.virginia.edu (R. Venkatesan).

0167-8116/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijresmar.2011.05.003

resource allocation decisions. Attempts to predict future customer profitability have been relatively unsuccessful, with very simple models often performing just as well as more sophisticated ones. The purpose of this paper is to provide a new approach to predicting future customer profitability; to estimate, test, and validate it on a large-scale customer database; and to demonstrate that it performs better than existing methods.

1.2. Managing the customer pyramid

Managers have long been aware that some customers are more profitable than others and have known that paying attention to more profitable customers can produce better results. One approach involves building a "customer pyramid" to classify customers into different profitability tiers (Rust, Zeithaml, & Lemon, 2000, pp. 187–231; Zeithaml et al., 2001). In the Rust-Zeithaml-Lemon customer pyramid, customers are classified into Platinum (the most profitable), Gold, Iron, and Lead (customers on whom the company loses money). Research in the banking industry (using a two-tier model) showed that increases in customer satisfaction in the top profitability tier were more effective for increasing future profits than were increases in the lower tier (Rust et al., 2000, pp. 199–200; Zeithaml et al., 2001).

The business world routinely discriminates among its customers on the basis of profitability. For example, FedEx instituted an approach that categorized its customers into three groups—the good, the bad, and the ugly (Brooks, 1999). The company worked hard to nurture the "good" customers and actively discouraged the "ugly" customers. This strategy should work well if profitability is stable, but what if it is not?

1.3. The evolving customer

Fig. 1 shows the profitability paths of two customers culled from the database of an actual B2B company. These customers were deliberately chosen for the purpose of comparison. We see that for the first twelve quarters (a period of three years), one customer is much more profitable than the other. The pattern reverses, however, in the next twelve quarters. If the company had based its marketing efforts on current or historical profitability, it would have wrongly identified the customer that was most profitable in the first twelve quarters as being the better target. Because the company can only affect future profits, current and historical profits are unimportant. With this consideration in mind, the company would like to classify its customers according to their future profitability as opposed to how profitable they are now or were in the past.

1.4. Customer lifetime value and customer equity

The attention to future profitability is also the basis for models of customer lifetime value (Berger & Nasr, 1998; Jain & Singh, 2002; Venkatesan & Kumar, 2004; Kumar et al., 2008) and customer equity (Blattberg & Deighton, 1996; Rust et al., 2000; Rust, Lemon, & Zeithaml, 2004). For example, research shows that directing

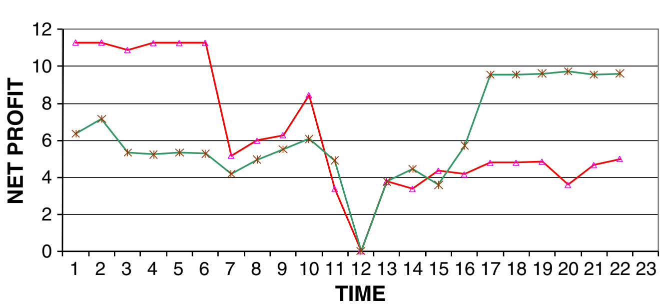

Fig. 1. The evolving customer.

marketing attention to customers with a high estimated customer lifetime value can lead to higher profitability (Venkatesan & Kumar, 2004). While Venkatesan and Kumar (2004) find that forward-looking metrics such as customer lifetime value are better at identifying profitable customers, they do not evaluate the accuracy of the profitability predictions of their model.

Models for predicting customer profitability and customer lifetime value have adopted two approaches—a brand-switching approach that uses customer surveys and historical data to project future customer profitability (Rust, Lemon, Zeithmal, 2004) and a customer database approach that projects future customer profitability only on the basis of the customer's interactions with the focal firm, with customer spending at competitors unknown (e.g., Fader, Hardie, & Lee, 2005a, 2005b; Reinartz & Kumar, 2003; Rust & Verhoef, 2005; Venkatesan, Kumar, & Bohling, 2007). In this paper, we address the latter case, in which a company possesses a historical customer database but has no information about customer purchases from competitors.

1.5. Predicting future customer profitability

Surprisingly, attempts to predict future customer profitability with information obtained from customer databases have not been entirely successful. Campbell and Frei (2004) show that it is easier to predict future profitability for some customers than for others, even for customers within the same profit tier. In general, they find that although using current profitability to predict future profitability explains considerable variance, the practice is problematic because this method misclassifies many customers. Malthouse and Blattberg (2005) build a variety of models based on regression and neural networks and test their ability to predict future customer profitability. They find that their best models misclassify most customers who are predicted to have high profitability.

Donkers, Verhoef, and de Jong (2007) build a model of customer retention and cross-buying that they use to project future profitability in a multiservice insurance industry setting. Their results show that the simplest model they test—maintaining the same profitability over time—performs the best and provides better predictions than other models they compared, including regression-based models and models that explicitly model customer retention and purchase probability.

Wuebben and Wangenheim (2006) derive similar conclusions. By testing stochastic models against simple management heuristic methods, they find that the simple models usually outperform the more complicated ones when it comes to predicting repurchase and whether a customer will be among the company's best in the future.

In general, the existing literature shows that predicting future profitability from customer databases is deceptively difficult and that simple methods tend to perform as well as more complicated ones. Building a realistic model that accurately predicts future customer profitability is, thus, an open research question that we attempt to address. In Section 2 we propose a new simulation model for predicting future customer profitability. Section 3 presents a number of alternative methods for predicting future customer profitability, including a model based on the BG/NBD model of Fader et al. (2005a), several simple models that are the best performers in the existing literature, and several variants of the simulation model. Section 4 presents an empirical test of the models, and Section 5 presents a discussion and conclusions.

2. The simulation model

2.1. Future profitability

We adopt the *always a share* approach for predicting future profitability (Jackson, 1985; Rust et al., 2004) because it is more appropriate for the noncontractual setting of our study (Venkatesan & Kumar, 2004; Venkatesan et al., 2007; Kumar et al., 2008). This approach assumes that there is only dormancy in a customer–firm relationship and that customers never completely terminate their relationship with a firm. This assumption allows for a customer to return to purchasing from a firm after a temporary dormancy, and when the customer returns to the relationship, he retains the memory about his prior relationship with the firm. Hence, in this approach, we measure future profitability of a customer by predicting his purchase pattern over the prediction period, and we do not predict when a customer would terminate his relationship with the firm. We measure the future profitability of customer i in terms of net profit ($NetP_i$) and net present value of profit (NPV_i):

$$NetP_i = \sum_{t=1}^{K} (\Pi_{it} - MC_{it}) \qquad NPV_i = \sum_{t=1}^{K} \frac{(\Pi_{it} - MC_{it})}{(1+d)^t} \tag{1a}$$

where $\Pi_{it}$is gross profit (in dollars) at time t, MC_{it} is variable marketing cost, and d is the discount rate. Gross profit is the margin that the firm obtains net of cost of goods sold (COGS) and is always positive. However, the net profit can be negative because it is gross profit net of the marketing cost. Net profit can be negative when customer i makes a purchase in time t but the marketing costs exceed the gross profit in that time period or when a customer does not make a purchase in time period t (i.e., gross profit is zero) and the marketing cost directed towards customer i in time period t is positive. The predicted future profit is then:

$$Net\hat{P}_i = \sum_{t=1}^{K} \left(\hat{\Phi}_{it}\hat{\Pi}_{it} - \hat{X}_{it}C\right) \qquad N\hat{P}V_i = \sum_{t=1}^{K} \frac{\left(\hat{\Phi}_{it}\hat{\Pi}_{it} - \hat{X}_{it}C\right)}{(1+d)^t} \tag{1b}$$

where,

$\hat{\Phi}_{it}$	predicted probability that customer i will purchase in time period t,
$\hat{\Pi}_{it}$	predicted gross profit provided by customer i in time period t given purchase,
$\hat{X}_{it}$	predicted number of marketing contacts directed towards customer i in time period t,
C	unit marketing cost of contacting a customer,
K	number of periods for the profitability calculation.

Given information about purchase behavior through time period $t = 0$, the profitability measures defined in Eq. (1b) are comprised of predictions of customer and firm behavior for a K-period time horizon. The predictions pertaining to purchase behavior include the propensity for customer i to purchase in each future time period t ($\hat{\Phi}_{it}$ in Eq. (1b)) and the gross profit provided by customer i given purchase in future time period t ($\hat{\Pi}_{it}$ in Eq. (1b)). While the predictions of purchase behavior capture the revenue aspect, marketing contacts such as number of sales calls or direct mailings from the firm to a customer ($\hat{X}_{it}$) capture the cost aspects and have to be predicted for accurate net profit measurement (Venkatesan et al., 2007).

2.2. The dynamics of customer profitability

Before we describe the model in detail, we provide a brief overview of the logic of the model. Fig. 2 illustrates the dynamics of customer profitability. In Fig. 2, customer characteristics include demographics such as age and income in the case of a business-to-consumer firm and firmographics such as industry category and number of employees in the case of a business-to-business firm. From the standpoint of the firm, what matters is whether the customer purchases in period t, and if so, how much gross profit results from the purchase. Both purchase and gross profit conditional on purchase are seen as being driven by marketing contacts, past purchase behavior (e.g., purchase incidence and gross profit), customer characteristics (or observed heterogeneity across customers), and control variables such as the state of the economy. Marketing contacts in time t are affected by previous marketing contacts, past purchase behavior, and customer characteristics (e.g., Kumar et al., 2008). We view these dynamics as occurring with error, necessitating simulation over multiple replications to fully evaluate the average future profitability.

We propose a single model framework that predicts customer purchase incidence (Pur_{it}), customer gross profit (Π_{it}), [1]conditional on purchase, and firm marketing contacts (X_{it}) and also models the potential correlations among these factors. Formally, we model the joint probability of gross profit, customer purchase incidence and marketing contacts as:

$$p(\Pi_{it}, Pur_{it}, X_{it}) = p(\Pi_{it}|Pur_{it} = 1, X_{it})p(Pur_{it} = 1|X_{it})p(X_{it}) \tag{2}$$

where Pur_{it} is the indicator of purchase and is equal to 1 if customer i purchases from the firm in time t and zero otherwise.

We model the marginal probability of marketing contacts because of the belief that firms allocate marketing contacts in a nonrandom fashion, i.e., following some allocation rules. We now explain the modeling process for each component in Eq. (2).

2.3. Purchase propensity

Following the choice modeling literature, we assume that at each time period t, customer i has a latent utility (U_{it}) for purchasing from the firm. This latent utility is assumed to be a function of customer characteristics or observed heterogeneity across customers (D_i), such as firm size, industry characteristics, past purchase behavior (Y_{past}), control variables (Z_t), such as overall macro-economic trends (e.g., GDP), and marketing contacts (X_{it}), such as number of contacts made by the sales personnel. The customer characteristics that were included were the sales of an establishment (a measure of the size of the establishment) and an indicator of whether the establishment belonged to the B2B industry category. For the variable "number of employees", we use the average value of the customer characteristics over the time frame of the analyses. The customer characteristics (D_i) therefore do not vary over time:

$$U_{it} = \alpha_i + I1_{it}\beta 1 + \mu'_{it} + \mu_{it} \tag{3a}$$

where $I1_{it}$ is the covariate vector that consists of Y_{it-1}, Z_t, and X_{it}, and $\beta 1$ is the corresponding coefficient vector. We capture correlations among purchase propensity, gross profits and marketing contacts through the term μ'_{it}. The random white noise term can therefore be represented as μ_{it}. This error structure ensures that unobserved factors do not bias the parameter estimates of our model. We explain the estimation of μ'_{it} later in this section. We incorporate heterogeneity across customers through the random effect intercept term (α_i).

[1] For our empirical applications, we deal with *contribution to profit*, rather than *total profit*, as our measure of profitability to avoid issues of allocating fixed costs to customers, which is often arbitrary, and also assume that the contribution to profit is positive (otherwise, the sale would not take place). Henceforth, we use the terms "profit" and "contribution to profit" interchangeably.

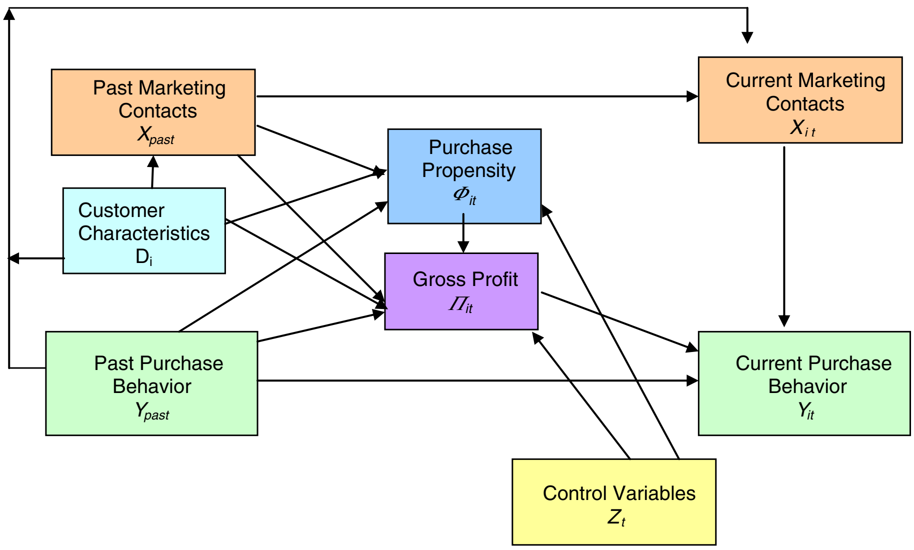

Fig. 2. Dynamics of customer profitability.

The customer-specific intercept term is modeled as a function of observed customer heterogeneity ($D1_i$):

$$a_i = D1_i\delta 1 + u_{0i} \tag{3b}$$

where u_{0i} is a customer-specific random error term, and $D1_i$ is a subset of the customer heterogeneity factors (D_i) that affect purchase propensity.

Substituting Eq. (3b) in (3a), we obtain:

$$U_{it} = D1_i\delta 1 + I1_{it}b1 + \mu'_{it} + \mu^*_{it} \tag{3c}$$

where $\mu^*_{it} = u_{0i} + \mu_{it}$. We accommodate potential correlations in purchase propensity over time for the same customer through the error term, μ^*_{it}. Let the vector $\mu^*_i = [\ \mu^*_{i1},..\ \mu^*_{i12}]'$ represent the error terms corresponding to the customer i over the twelve quarters used in our analysis. We assume that $\mu^*_{i.}$ follows a multivariate normal distribution with zero mean and variance covariance matrix R_μ with a block diagonal structure where the blocks correspond to each individual. The off-diagonal elements in each block of R_μ, $\alpha 1_{jk}$ ($j \neq k$), capture the correlation of purchase propensity across years for a single customer. We observe a purchase from the customer if the latent utility (U_{it}) exceeds a threshold, which is normalized to zero. Therefore, the probability of a customer i making a purchase in time period t is given by:

$$p(Pur_{it} = 1|X_{it}) = \Phi(D1_i\delta 1 + I1_{it}b1 + \mu'_{it}) \tag{4}$$

where Pur_{it} indicates whether customer i made a purchase in time t, and Φ is the cumulative normal distribution.

2.4. Gross profit

We model the log of the gross profit at time t from customer i, given purchase in time t, as a linear regression:

$$\log(p_{it}|Pur_{it} = 1) = D2_i\delta 2 + I2_{it}\beta 2 + v^*_{it} \tag{5}$$

where $I2_{it}$ is the covariate vector that consists of Y_{it-1}, Z_t, and X_{it}. Similar to the purchase propensity model, customer heterogeneity is captured through $D2_i$. The covariance in profits from the same customer over time is captured through the correlation matrix R_v, which has a structure similar to R_μ in Eq. (3c). We model the gross profit for only those observations where there is a purchase (i.e., $Pur_{it} = 1$); for the rest of the observations, the gross profit is equal to zero because there is no purchase. This approach is equivalent to modeling gross profit conditional on a purchase.

2.5. Marketing contacts

Firms are likely to determine the level of marketing contacts for each customer following allocation rules that are a function of past purchase behavior and past levels of marketing contacts (Manchanda, Rossi, & Chintagunta, 2004; Donkers, Paap, Jonker, & Frances, 2006; Kumar et al., 2008). Marketing contact levels are therefore endogenous to firm expectations of future customer behavior in our sample. The observations of marketing contacts in a customer database are therefore not likely to be random. Treating such a sample as random can result in inconsistent parameter estimates.

To account for the possibility that marketing contacts are allocated in a nonrandom fashion, we model the number of salesperson contacts directed towards customer i in time t, X_{it}, as a function of its lagged values and past purchase behavior. Specifically, we estimate:

$$X_{it} = D3_i\delta 3 + I3_{it}\beta 3 + e^*_{it} \tag{6}$$

where $I3_{it}$ is a vector constructed from past marketing actions (X_{it-1}, X_{it-2}..) and past purchase behavior (Y_{it-1}, Y_{it-2}..). Once the model is estimated, we use the predicted values of marketing contacts, i.e., $\hat{X}_{it} = D3_i\hat{\delta}3 + I3_{it}\hat{\beta}3$, for X_{it} in the purchase propensity and gross profit equations. Similar to the gross profit model and the purchase propensity model, the covariance in marketing contacts from the same customer over time is captured through the correlation matrix R_ε, which has a structure similar to R_μ in Eq. (3c) and R_v in Eq. (5). This approach is similar to the framework proposed by Donkers et al. (2006) and also operationalizes the model setup proposed in Eq. (2). Accounting for the endogeneity between the level of marketing contacts and the firm's expectations of future spending allows us to estimate better coefficients of the firm's responsiveness to marketing. However, accounting for endogeneity may not be so critical for obtaining more accurate predictions of future customer profitability (Ebbes, Papies, & van Heerde, 2010).

The extant scanner panel literature has found that a customer's purchase propensity, purchase quantity, and the instrumental variables can be correlated because of unobserved factors that are not included as drivers in the model framework (Chintagunta, 1999; Villas-Boas & Winer, 1999). Not accounting for the correlation among the model components can lead to biased model estimates and poor predictive performance. We therefore model the correlations among the various error terms μ'_{it}, υ^*_{it}, and ε^*_{it} from Eqs. (4), (5), and (6), respectively. Our challenge to model this correlation is simplified because all of the error terms are assumed to be normally distributed and we can empirically observe or estimate $\hat{v}^*_{it}$ (random error in the profitability model) and $\hat{\varepsilon}^*_{it}$ (random error in the marketing contacts model). Because μ'_{it} (correlation term in the purchase propensity model) is unobserved, we assume that μ'_{it} is related to v^*_{it} and ε^*_{it} in the following fashion:

$$\begin{aligned}\mu'_{it} &= \sum_j \theta_j F_{jit} + \eta_{it} \\ &= \sum_j \theta_j \left(\gamma_j \hat{\varepsilon}^*_{it} + \alpha_j \hat{\upsilon}^*_{it}\right) + \eta_{it} \Rightarrow \mu'_{it} \\ &= \left(\sum_j \theta_j \gamma_j\right) \hat{\varepsilon}^*_{it} + \left(\sum_j \theta_j \alpha_j\right) \hat{\upsilon}*_{it} + \eta_{it}\end{aligned} \tag{7}$$

where γ_j and α_j are the factor coefficients from a principal components analysis of $\hat{\varepsilon}^*_{it}$ and v^*_{it}, respectively, j indexes the principal components, F_{jit} denotes the factor scores, and η_{it} is a random normal error term. We are unable to implement a simpler model framework such as a simultaneous equation system that also accounts for correlation in errors across the model components because purchase propensity is not directly observed.

2.6. Model estimation

Because a conventional simultaneous estimation approach is not feasible and the structure of the model is recursive, we estimate the model sequentially. Because of the recursive nature of the model, this estimation plan is equivalent to simultaneous estimation. Our estimation algorithm proceeds as follows:

Step 1: Estimate the marketing contacts component (Eq. (6)) and obtain the corresponding residuals $\hat{\varepsilon}^*_{it}$. This estimation is possible because the predictors in Eq. (6) are known.

Step 2: Estimate the gross profit component (Eq. (5)) using the predicted value of marketing contacts obtained from Step 1 as an independent variable. Obtain the corresponding residuals $\hat{\upsilon}^*_{it}$. We use the instrumental variable $\hat{X}_{it}$ in place of X_{it}. The other predictors are already known.

Step 3: Perform factor analyses of $\hat{\varepsilon}^*_{it}$ and $\hat{\upsilon}^*_{it}$ and obtain the factor coefficients (denoted as γ_j and α_j, respectively) as in Eq. (7). Set $F_{jit} = (\gamma_j \hat{\varepsilon}^*_{it} + \alpha_j \hat{\upsilon}^*_{it})$.

Step 4: Substitute F_{jit} from step 3 into Eq. (3c), i.e.,

$$U_{it} = D1_i\delta 1 + I1_{it}\beta 1 + \sum_j \theta_j\left(\gamma_j \hat{\varepsilon}^*_{it} + \alpha_j \hat{\upsilon}^*_{it}\right) + \mu^*_{it}.$$

Estimate $\beta 1$, $\delta 1$, and θ_j by maximizing the likelihood that results from Eq. (4). Note that we can obtain $\hat{\mu}'_{it}$, as shown in Eq. (7) above, from $\hat{\mu}'_{it} = \sum_j \theta_j F_{jit}$.

Our proposed model framework and the corresponding estimation algorithm allow us to make the predictions necessary for computing expected future customer profitability in a single framework while also accommodating for the correlation between the various factors used to obtain future profitability.

2.7. Predicting future profitability

When predicting future customer profitability, there is significant uncertainty in predicting purchase behavior over a multiperiod horizon. Typical sources for this uncertainty relate to poor or nonexistent information in most CRM databases on customer transactions with the competition, competitor marketing actions targeted at each customer, and customer attitudes. Sometimes the cost of increased errors can outweigh the benefits of long-term predictions. In our model framework, the uncertainty in future purchase behavior is captured by the error variables, ε^*_{it}, υ^*_{it}, and μ^*_{it}, corresponding to marketing contacts, gross profit and purchase propensity, respectively. The simulation algorithm we propose for predicting future customer profitability accounts for this uncertainty in future purchase behavior. The simulation of errors is necessary because for nonlinear CLV functions, it is more accurate to take the expectation of the predicted distribution of CLV than it is to predict CLV with the mean or expected values of the input factors, such as gross profits and marketing contacts (Venkatesan et al., 2007). The algorithm proceeds as follows:

Step 1: Use the estimation sample and the estimated model coefficients to obtain the empirical distribution (mean and variance) of the error variables in the marketing contacts and gross profit components. Let $\hat{\mu}_{\varepsilon*}$ and $\hat{R}_{\varepsilon*}$ represent the mean and correlation matrix of the error variable ($\hat{\varepsilon}*$) in the marketing instruments component (Eq. (6)). Similarly, let $\hat{\mu}_{v*}$, and $\hat{R}_{v*}$ represent the mean and variance of the error variable ($\hat{v}*$) in the gross profit component (Eq. (5)).

Step 2: For each customer i, generate a random variable from the error distribution of the marketing contacts component corresponding to future time period t and replication r, $\hat{\varepsilon}*_{itr} \sim N\left(\hat{\mu}_{\varepsilon*}, \hat{R}_{\varepsilon*}\right)$, and predict the marketing contacts for the future time period t and replication r, ($\hat{X}_{itr}$), from Eq. (6). Specifically, we obtain the predicted marketing contacts ($\hat{X}_{itr}$) from:

$$\hat{X}_{itr} = D3_i\hat{\delta}3 + I3_{it}\hat{\beta}3 + \hat{\varepsilon}*_{itr}.$$

Step 3: Generate a random variable from the error distribution of the gross profit component for customer i in time period t and replication r, i.e., $\hat{v}*_{itr} \sim N\left(\hat{\mu}_{v*}, \hat{R}_{v*}\right)$.

Step 4: Use Eq. (7) and the random variables generated in Step 2 to generate the correlation term variable for purchase propensity, $\hat{\mu}'_{itr}$. Use Eq. (3) to predict the latent utility of purchase ($\hat{U}_{itr}$) and the corresponding probability of purchase ($\hat{\Phi}_{itr}$) from Eq. (4). Specifically, we obtain $\hat{\Phi}_{itr}$ from:

$$\hat{\Phi}_{itr} = \Phi\left(D1_i\hat{\delta}1 + I1_{itr}\hat{\beta}1 + \hat{\mu}'_{itr}\right).$$

Generate a uniform random variable, $u \sim [0,1]$; if u is less than $\hat{\Phi}_{itr}$, then customer i is predicted to purchase in time period j for replication r, i.e., $\hat{P}ur_{itr} = 1$. We use the predicted marketing contacts $\hat{X}_{itr}$ within $I1_{itr}$ in the prediction of $\hat{\Phi}_{itr}$.

Step 5: If, $\hat{P}ur_{itr} = 1$, in Step 4, predict gross profit at time t for customer i and replication r, $\hat{\Pi}_{itr}$, from Eq. (5); otherwise, set $\hat{\Pi}_{itr}$ to zero. When predicting gross profit from Eq. (5), use the predicted marketing contacts, $\hat{X}_{itr}$, from Step 2 and the random variable from Step 3. We obtain the predicted gross profits $\hat{\Pi}_{itr}$ from:

$$\hat{\Pi}_{itr} = \exp\left(D2_i\hat{\delta}2 + I2_{it}\beta 2 + \hat{v}*_{it}\right).$$

Step 6: Repeat Steps 2–5 until the end of the prediction horizon, i.e., for twelve quarters or $t = 12$. Using the predicted values of marketing contacts ($\hat{X}_{itr}$), purchase propensity ($\hat{\Phi}_{itr}$), and gross profit ($\hat{\Pi}_{itr}$), compute the predicted net profits ($Net\hat{P}_i$) and net present value of profitability ($N\hat{P}V_i$) for customer i corresponding to replication r. The predicted values of $\hat{X}_{itr}$, $\hat{\Phi}_{itr}$, and $\hat{\Pi}_{itr}$ at time t are used as lagged variables in the prediction for time period $t+1$.

Step 7: Repeat Step 6 R times, where R is the total number of replications. (We set R equal to 1000.) The expected values of TP and NPV for customer i are then the averages of the values obtained in the R replications.

A simulation approach to prediction works better than a mean one-step-ahead approach because there is state dependence and because failure to consider the entire distribution of possible outcomes leaves out many of the states. To see why the simulation approach is better, consider a very simple Markov model in which there are states A (profit = 2), B (profit = 1), and C (profit = 0). Let us suppose that the customer begins in state B and that the probabilities for the next period, conditional on B, are $p(\text{A;B}) = .2$, $p(\text{B;B}) = .6$, and $p(\text{C;B}) = .2$. Now suppose that the probabilities conditional on A are $p(\text{A;A}) = .5$, $p(\text{B;A}) = .3$ and $p(\text{C;A}) = .2$. Also, the probabilities conditional on C are $p(\text{A;C}) = .1$, $p(\text{B;C}) = .1$ and $p(\text{C;C}) = .8$. Now let us consider the expected total profit after two periods for someone who begins in state B. The possible outcomes (and associated total probabilities) are AA (.10), AB (.06), AC (.04), BA (.12), BB (.36), BC (.12), CA (.02), CB (.02), and CC (.16). Ignoring the distribution and estimating the mean profitability one step ahead will cause the model to predict a profitability of 1.0 in period 1 and again in period 2, for a total profit of 2.0. The actual expected total profit can be calculated to be 1.2. By contrast, a simulation that captures the correct transition probabilities will, on average, recover the correct amount. Simply projecting the means leaves out the state dependence effects from high or low values, which can be asymmetric and result in a bias.

Next, we present alternatives to the proposed simulation model. These alternatives include several simple existing models that have been top performers in the literature, a new model based on the BG/NBD model, and several variants of the simulation model.

3. Alternative models

3.1. Existing models

3.1.1. Current profitability

This very naïve model simply takes the profit from the most recent period for each customer and projects it indefinitely into the future. Donkers et al. (2007) find this model to be better than some more complicated models in mean absolute deviation of predicted customer lifetime value. Campbell and Frei (2004) also find support for this model.

3.1.2. Average profitability

This model projects future profit as being equal to the historical average for that customer. This model is very similar to the current profitability model but may be more stable if customer profitability fluctuates.

3.1.3. Trend in profitability

This model estimates a simple linear regression on profitability for each customer, with time as the predictor. Donkers et al. (2007) and Malthouse and Blattberg (2005) test variants of this model.

3.2. New models

3.2.1. The BG/NBD model

Models by Fader et al. (2005a and 2005b) have provided excellent predictions of whether the customer is still "alive"—that is, whether the customer continues to do business with the firm. Given the success of these models, we attempted to create a new model based on theirs, even though predicting future profitability was not the intended purpose of their models. This model, which is based on a customer dropout and repeat purchase model proposed by Fader et al. (2005a), assumes that customer purchases follow the BG/NBD framework and that profitability conditional on purchase remains constant and equal to the customer's historical average. This BG/NBD model is known to produce results similar to the Pareto/NBD model (Fader et al., 2005b) but is much easier to implement (Fader et al., 2005a).

Next, to evaluate the various components of our simulation model, we propose below a set of competing nested models that are simplified variants of the simulation model.

3.2.2. Proposed model without simulation and error correlation

This model includes separate gross profits and purchase propensity models, and we do not simulate the errors when predicting CLV.

3.2.3. Proposed model without simulation

Here, we allow for correlation between the gross profits and purchase propensity models but do not simulate errors when predicting CLV.

3.2.4. Simulation model without replication

In this model, each customer's future profitability is projected using the mean estimates from each equation. (This approach is equivalent to simulating one replication in which the error terms are all equal to zero.)

3.2.5. Simulation model without customer heterogeneity

One of the basic notions of CRM is that customers are heterogeneous, and accounting for this heterogeneity in customer valuation and in marketing contacts directed towards customers can increase firm profitability (Venkatesan & Kumar, 2004). We therefore compare the performance of the proposed model framework without accounting for customer heterogeneity. In this framework, we do not estimate the customer-specific intercept parameter that is specified in Eqs. (3c), (4), (5), and (6). Instead, we estimate a single intercept parameter that is common to all customers in these equations. For example, μ^*_{it} in Eq. (3c) is equal to μ_{it} in this model and is distributed normally, with a mean of zero and variance σ^2_μ.

3.2.6. Simulation model without the marketing contacts instrument

In this model, we do not use the marketing contacts component. We use the observed level of marketing contacts in a quarter directly in the purchase propensity and gross profit components. Specifically, we do not estimate Eq. (6), and we use the errors from the gross profit component directly to operationalize μ_{it} instead of using Eq. (7). This benchmark model therefore assumes that marketing contacts are allocated at random to each customer. The performance of this model allows us to assess the extent of bias that results from assuming that marketing contacts are random when in fact the firm allocates marketing contacts according to an allocation rule, i.e., in a nonrandom fashion.

4. Empirical testing

4.1. Artificial data

To ensure that the model is identified, we estimate our model on artificial data and evaluate the extent to which our model and estimation strategy recover the true parameters of the data. The results of this simulation are provided in Table 1. Three different artificial datasets were generated by varying the error variance in the gross profits (σ^2_v) and marketing contacts (σ^2_ε) equations for 150 customers over twelve quarters. We see that the correlation between the estimated and true parameters that are used to create the artificial data increases when we: (a) allow for correlation between the purchase propensity and gross profit models; (b) account for the endogeneity of marketing contacts through the instrumental variables; and (c) include customer heterogeneity. The accuracy of predictions declines for all of the models as the error variance increases. The correlation between the true and estimated parameters is reasonable in the proposed model, even for the artificial data with the highest error variances. We show in Appendix A that the CLV predictions from the

Table 1
Correlations between true and estimated parameters in artificial data.[a]

	Separate models for purchase propensity, gross profit and no IV	Include correlation between purchase propensity and gross profit (selection model)	Include marketing contacts IV	Include customer heterogeneity
$\sigma^2_v = 1e5, \sigma^2_\varepsilon = 10$				
Purchase propensity model	0.85	0.89	0.96	0.98
Gross profit model	0.88	0.92	0.97	0.98
Marketing contacts model	NA	NA	0.94	0.97
$\sigma^2_v = 1e10, \sigma^2_\varepsilon = 20$				
Purchase propensity model	0.8	0.85	0.91	0.92
Gross profit model	0.78	0.86	0.89	0.93
Marketing contacts model	NA	NA	0.88	0.91
$\sigma^2_v = 1e15, \sigma^2_\varepsilon = 30$				
Purchase propensity model	0.72	0.76	0.82	0.88
Gross profit model	0.69	0.78	0.83	0.87
Marketing contacts model	NA	NA	0.79	0.89

NA = Not Applicable.
[a] Models are nested.

proposed model are better correlated with the observed CLVs than are simpler versions of the proposed models and naïve models using current, average and trend in customer profitability.[2]

4.2. Data

For our empirical analysis, we use customer data from a firm that sells a number of high-technology products and services to other business customers. The firm's products typically require maintenance and frequent upgrades; these provide the variance necessary for satisfactory model estimation. We randomly sampled a set of 191 customers from a cohort of customers who started purchasing in the first quarter of 1997.[3] We use the first twelve quarters of data (first quarter 1997 to fourth quarter 2000) as the calibration sample to estimate the models. We use the next twelve quarters (first quarter 2001 to fourth quarter 2003) as the holdout sample to evaluate the accuracy of our profitability predictions because contributions to profit beyond three years are lower in value and because the product offerings can change substantially over time (Gupta & Lehmann, 2005; Kumar et al., 2008). Firms in this industry prefer to predict CLV up to three years.

4.3. Drivers for the simulation model

Following past research on customer equity and CLV (Reinartz & Kumar, 2003; Rust et al., 2004; Venkatesan & Kumar, 2004), we identify several drivers of customer profitability and purchase propensity. The drivers are classified as customer characteristics, marketing contacts and control variables. The customer characteristics variables are further classified as exchange characteristics and customer heterogeneity. The exchange characteristics define and describe the nature of the customer–firm exchange, whereas customer characteristics capture customer heterogeneity. Different exchange characteristics that we include as predictors of purchase propensity (Eqs. (3a)–(3c) and (4)) and gross profit (Eq. (5)) include past customer spending level, past purchase incidence, cross-buying, frequency of past purchase activity, and the marketing contacts by the firm. Cross-buying for customer i in time period t is measured as the number of different product categories from which a customer purchased until time t. Past purchase activity measures the level of customer activity in the most recent two quarters. Past purchase activity in time period t is set to 2 if a customer purchased in both time periods $t-2$ and $t-1$, to 1 if the customer purchased in either $t-1$ or $t-2$ and to 0 otherwise.

The firm contacts customers through several channels, including in-person visits from salespeople, direct mail and telesales. In this study, we define marketing actions as the total number of firm contacts with a customer (across all channels) in a particular quarter. Because R-squared did not differ between models that allowed separate effects for the marketing contacts across the different channels and a model that combined the contacts across all of the channels, we combined all of the channels into a single measure in our study.[4] Further, our main interest in this study is to provide a framework for accurately predicting future customer profitability while sufficiently accommodating the endogeneity of marketing actions. Drivers of marketing actions are based on both theory and discussions with firm representatives regarding how the firm allocates marketing touches to a customer. The drivers of marketing actions include past customer spending, past levels of marketing contacts, and customer relationship (or exchange) characteristics such as cross buying. When assigning predictors, we ensured that there was at least one unique predictor for each dependent variable, i.e., marketing contacts, purchase propensity and profitability, to ensure model identification (Greene, 1993). Next, we present results from model estimation. Although we assessed the influences of several drivers, we only show those that had significant effects and did not result in an unacceptable level of multicollinearity. We provide the descriptive statistics of the drivers of customer profits used in our model in Table 2.

4.4. Model estimation for the simulation model

The first twelve quarters (first quarter of 1997 to fourth quarter of 2000) of data are used to estimate the parameters of the simulation model. The estimated parameters are provided in Table 3. The table shows the results of estimating the equations for purchase propensity and gross profit as well as the instrumental variables' regression on marketing contacts. The Adjusted R^2 values are equal to 0.73, 0.70, and 0.69 for the purchase propensity, gross profit and marketing contacts components, respectively, indicating good in-sample fit. [5] All of the signs are in the expected directions. The error variables from the marketing contacts component ($\hat{\varepsilon}^*_{it}$) and the gross profit component (v^*_{it}) loaded onto a single factor, with a factor score of 0.69 for both

[2] The model estimation and prediction codes and the artificial data are available to interested readers from the authors.
[3] We used a cohort of customers for our analysis to avoid potential left-censoring issues.
[4] We estimated a linear effect because a nonlinear, quadratic or log effect of marketing did not add explanatory power in our data.
[5] We evaluated the R^2 values by comparing the residual variance obtained from the model with the residual variance obtained from using the mean in the calibration sample across all customers as the null model.

Table 2
Operationalization and descriptive statistics for drivers of CLV.

Variable	Operationalization	Mean	Standard deviation
Cross-buying	Number of different product categories from which a customer has purchased.	10.10	7.80
Lagged indicator of purchase	Equal to 1 if customer made a purchase in the previous quarter, 0 otherwise.	0.70	0.40
Gross profit	Gross profit provided by the customer in the current quarter.	57,047	156,654
Lagged gross profit	Gross profit provided by the customer in the previous quarter.	54,801	155,931
Two-period lagged gross profit	Gross profit provided by the customer two quarters prior to the current quarter.	53,954	154,172
Lagged growth in gross profit	Difference between lagged gross profit and two-period lagged gross profit	848	95,792
Lagged average gross profit	Average of the gross profit provided by the customer from the first quarter until the previous quarter	52,582	131,764
Log of lagged total marketing contacts	Log of the number of contacts through all channels (salesperson, direct mail and telesales) to a customer in the previous quarter	−1.90	4.40
Lagged average total marketing contacts	Average number of marketing contacts through all channels to a customer per quarter from the first quarter until the previous quarter	1.70	2.60
Two-period lagged total marketing contacts	Number of marketing contacts through all channels to a customer two quarters prior to the current quarter.	3.36	6.27

Table 3
Parameter estimates for the simulation model.

Purchase propensity		Gross profit		Marketing contacts	
$R^2 = .75$		$R^2 = 0.72$		$R^2 = 0.73$	
Adj. $R^2 = 0.73$		Adj. $R^2 = 0.70$		Adj. $R^2 = 0.69$	
N = 2292		N = 1605		N = 2292	
Variable	Parameter estimate	Variable	Parameter estimate	Variable	Parameter estimate
Intercept	−1.20***	Intercept	6.50	Intercept	0.99***
Cross-buying	0.20***	Lagged gross profit	0.39***	Log of lagged total marketing contacts	0.22***
Lagged indicator of purchase	1.55***	Predicted total marketing contacts	0.03**	Lagged gross profit	2.20e−06***
Lagged activity	0.18**	Cross-buying	0.16***	Lagged average total marketing contacts	1.91***
Lagged growth in gross profit	0.02**	Two-period lagged gross profit	0.23***	Cross-buying	0.20***
Number of employees	0.05***	Lagged average gross profit	0.26***	Two period lagged total marketing contacts	0.11***
B2B	0.18***	Lagged growth in gross profit	0.02**	Number of employees	0.06***
GDP	0.21***	Number of employees	1.30***	GDP	0.04*
θ_1	1.00e−04***	GDP	0.15***		
$\rho_{R\mu}$		$\rho_{R\nu}$	3.60e−07***	$\rho_{R\varepsilon}$	−0.57**
σ_μ^2		σ_ν^2	1.07e10***	σ_ε^2	24.55***

*Significant at $\alpha = 0.10$, ** Significant at $\alpha = 0.05$, *** Significant at $\alpha = .01$. Significance of the coefficients is based on two-sided tests. Adj. R^2 = Adjusted R^2.

of the variables. Histogram plots indicated that both the marketing contacts component error variable ($\hat{\varepsilon}^*_{it}$) and the gross profit component error variable ($\hat{\nu}^*_{it}$) were normally distributed. [6]We discuss the drivers of each component below. We report bootstrapped standard errors that are based on 100 replications of our model estimation.

4.4.1. Purchase propensity

The results indicate that customers who: (a) have purchased across different product categories, i.e., higher cross-buy (0.20, $p<0.01$); (b) bought in the previous quarter, i.e., lagged indicator of purchase is equal to 1 (1.55, $p<0.01$); (c) have exhibited higher levels of past purchase activity (0.18, $p<0.05$); and finally, (d) exhibited growth in their gross profits over the previous two quarters (0.02, $p<0.05$) are more likely to purchase in the current quarter. We observe that the drivers included in the marketing contacts model are not able to capture the relationship between marketing contacts and purchase propensity (i.e., the predicted number of marketing contacts does not have a significant influence). However, the residual error correlation between marketing contacts and gross profit has a positive and significant ($\hat{\theta}_1 = 1.0e{-}04$, $p<0.01$) relationship with purchase propensity. This result implies that our model framework is able to capture the relationship between marketing contacts and purchase propensity that is not accounted for by the instrumental variable model.

Customer characteristics have a significant effect on purchase propensity supporting the need to account for customer heterogeneity in the purchase propensity model. We find that customers that have more employees (0.05, $p<0.01$) and customers who belong to the B2B industry category (0.18, $p<0.01$) are more likely to make a purchase. Macro-economic trends (captured by GDP in this study) did not significantly influence customer purchase propensity. The correlation parameter in the error correlation matrix for purchase propensity (R_μ) is significant and positive ($\hat{\rho}_\mu = 0.05$, $p<0.01$). This result implies that customers' purchase propensities are positively correlated over time and provides support for the random effects model structure of the purchase propensity model.

4.4.2. Gross profit

Similar to purchase propensity, customers with a higher level of cross-buying (0.16, $p<0.01$) and a higher level of past purchase activity (0.38, $p<0.01$) are expected to yield a higher gross profit in the current quarter. Past customer spending levels, captured by (a) the level of gross profit in the previous two quarters (0.23, $p<0.01$) and (b) the average level of gross profit in all previous quarters except the most recent two quarters (0.26, $p<0.01$), positively influence a customer's spending in the current quarter. We also find a significant and positive influence of marketing contacts on a customer's current spending level (0.03, $p<0.01$). Similar to purchase propensity,

[6] The Anderson–Darling test (Stephens, 1974) failed to reject the null hypothesis that the data are normally distributed at $\alpha = 0.05$.

Table 4
Changes in profitability over time.

		Future profitability (periods 13–24)		
		Low[a]	Medium[b]	High[c]
Past profitability (Periods 1–12)	Low[a]	23	17	7
	Medium[b]	32	47	18
	High[c]	6	16	25

[a] Lowest quartile.
[b] Middle two quartiles.
[c] Highest quartile.

observed customer heterogeneity factors have a significant influence on gross profit. We find that customers that have more employees (1.30, $p<0.01$) are likely to yield a higher gross profit. We observe that customers are more likely to provide higher revenue when the macro-economic trends are good, i.e., higher GDP (0.15, $p<0.01$). Finally, the correlation parameter in the error correlation matrix for the profitability model (R_v) is positive and significant ($\hat{\rho}_v = 0.02$, $p<0.01$). This result implies that gross profit is positively correlated over time and provides support for modeling the random effects model structure of the gross profit component.

4.4.3. Marketing contacts

The level of marketing contacts for a customer depends on that customer's recent purchase behavior, which is captured through lagged gross profit (2.20e−6, $p<0.01$). While a customer's past purchase behavior, in general, determines whether a customer is contacted, the specific level of marketing contacts directed towards a customer in a particular month is influenced to a large extent by the level of contacts for the customer in the two prior quarters (0.22, 0.11, $p<0.01$) and by the average level of contacts allocated to the customer in the past, except the most recent two quarters (1.91, $p<0.01$). Finally, the positive influence of cross-buying (0.20, $p<0.01$) reflects the firm's strategy of focusing marketing activities on customers who have been actively purchasing products across several categories. The correlation parameter in the error correlation matrix for the marketing contacts model (R_ε) is negative and significant ($\hat{\rho}_\varepsilon = -0.024$, $p<0.01$). The drivers of marketing contacts included in the model imply that the firm maintains a high level of contacts for customers who were contacted frequently in the past. The correlation parameter, on the other hand, shows that the firm slightly decreases the level of contacts for high-contact customers (and vice versa) over time with factors not captured in the model, such as salesperson inputs.

The firm seems to focus its marketing contacts on customers that have more employees (0.06, $p<0.01$) but does not seem to focus on any one industry category in particular. The results do not provide strong evidence that the marketing contacts of the firm are influenced by macro-economic trends. We evaluate the accuracy of predictions from the simulation model and competing models in the next section.

Several other drivers, such as variance in marketing contacts, number of upgrades, zip code of the customer's headquarters, number of customer service calls, and number of sales through the Internet were also included in our model but were found to have insignificant effects beyond the variables already in the model.

4.5. Accuracy of predictions

Customer profitability (i.e., net profit) is not always stable. Table 4 shows the relationship between past profitability (the estimation sample, periods 1–12) and future profitability (the validation sample, periods 13–24). Customers were classified into three groups (lowest quartile, middle half, highest quartile) by profitability. We can see from this table that customers evolve over time. For example, we see that 7 (15%) customers switched from the lowest profitability quartile to the highest quartile. We also see that 6 (13%) customers switched from the highest profitability quartile to the lowest quartile. Consistent with previous research (Malthouse & Blattberg, 2005), there is considerable movement among profitability tiers. It would appear as though accurately predicting future customer profitability would require predicting movement from low profitability to high and vice versa. An effective model of future customer profitability would be capable of predicting at least some of this movement.

The model estimates obtained from the calibration sample provided input to the simulation algorithm that was used to predict future customer profitability. Future profitability was estimated by predicting the marketing contacts, purchase propensity and gross profit values in the holdout time frame (1st quarter of 2001 to 4th quarter of 2003). We also compared the predictive accuracy of the simulation model against the predictions from the various alternative models.

Table 5 provides the mean absolute deviation (MAD) between the predicted and observed values of profit for the holdout period for the various competing and nested models. The average total net profit over periods 13 through 24, or three years (as shown in Eq. (1a)), in the holdout sample is $867,524. Table 5 also provides the mean absolute deviation of the net present value of the future profits over three years with an annual discount rate of 15%. We see from Table 5 that the simulation model is the most accurate both for predicting future net profitability and for predicting net present value of profitability. The mean difference between net profits over three years predicted from the simulation model and the corresponding observed net profits over three years, i.e., the MAD of the simulation model, is $409,000. The MAD of the NPV of future net profits from

Table 5
Model comparison — predicting future customer profitability.

Model	AIC	Future profit[a]		NPV of future profit[b]	
		MAD[c]	MAPE[d] (%)	MAD	MAPE (%)
Current profitability	NA	5.54*	64	4.34*	60
Average profitability	NA	5.23*	60	4.14*	57
Trend in profitability	3607	4.61*	53	3.75*	52
BG/NBD	3521	4.44*	51	3.64*	50
Proposed model without simulation and correlation among components [a]	4002	5.22*	60	4.11*	57
Proposed model without simulation [a]	3994	5.20*	60	4.09*	57
Simulation model without replication [a]	3935	5.19*	60	4.05*	56
Simulation model without customer heterogeneity [a]	3487	4.20	48	3.40	47
Simulation model without the marketing contacts instrument[a]	3754	5.08*	59	3.95*	55
Simulation model[a]	3258	4.09	47	3.28	45

*Significantly higher than the MAD of the Simulation Model at $\alpha = 0.016$. The significance of the model comparisons is based on one-tailed tests. NA = Not Applicable.
[a] Nested models.
[b] Reported values are in $100,000 s.
[c] MAD = Mean Absolute Deviation.
[d] MAPE = Mean Absolute Percentage Error.

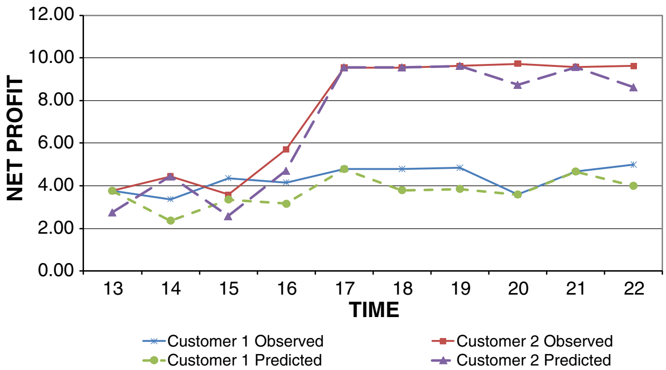

Fig. 3. Predicted and observed net profits for illustrative customers.

the simulation model is $328,000 for NPV of future profit. The Mean Absolute Percentage Error (MAPE) metric shows that the prediction error from the proposed model represents about 47% and 45% of the average future total net profits and the NPV of the future net profits, respectively. For the sake of simplicity, we provide the rest of the MAD values in units of $100,000. The second-best model for predicting the future net profit was the simulation model without customer heterogeneity (4.20), followed by the BG/NBD model with an MAD of 4.44, the trend in the profitability model (4.61), the simulation model without the marketing contacts instrument (5.08), the simulation model without replication (5.19), the proposed model without simulation (5.20), the proposed model without simulation and correlation among the CLV components (5.22), average profitability (5.23), and current profitability (5.54), respectively. The second-best model for predicting the NPV of future net profit was also the simulation model without heterogeneity (3.40), followed by the BG/NBD model with an MAD of 3.64, the trend in profitability model (3.75), the simulation model without the marketing contacts component (3.95), the simulation model without replication (4.05), the proposed model without simulation (4.09), the proposed model without simulation and correlation among the CLV components (4.11), average profitability (4.14), and current profitability (4.34), respectively. Fig. 3 shows the predicted profitabilities for the two customers shown in Fig. 1 for the twelve quarters in the holdout time period.

Predictions from the simulation model without replication performed worse than did those from the simulation model with replication. This result indicates that the multiple replications of customer futures are critical for obtaining accurate predictions of future customer profitability. The simulation model simulates future purchasing scenarios one customer at a time. By averaging over many possible futures for each customer, the proposed simulation algorithm allows us to obtain a more accurate prediction of future profitability. The simulation model without customer heterogeneity performs better than all other comparison models but worse than the simulation model that includes customer heterogeneity. This result indicates the importance of accounting for customer heterogeneity in customer profitability models.

We treat our forecasting exercise as a within-subjects experimental design where the different models are considered to be treatments administered to the same customer. For each model, the absolute deviation between the forecasted and observed lifetime values and the absolute deviation between the forecasted and observed NPV of future profit from each customer are considered the response variables. A one-tailed F-test indicated that the MADs from the various comparison models and the proposed model were significantly different from each other at $\alpha = 0.10$. We then conducted seven pairwise comparisons of the MADs, i.e., comparing the MADs from each of the comparison models with the MAD from the proposed model.

Table 6
Simulation model prediction accuracy — changes in customer profitability.

		Actual change in profitability		
		Decline	Stable	Growth
Predicted change in profitability	Decline[a]	50 (89%)	5 (9%)	1 (2%)
	Stable[b]	12 (13%)	68 (73%)	13 (14%)
	Growth[c]	2 (4%)	10 (20%)	38 (76%)

[a] Decline defined as >20% decline in total profitability.
[b] Stable defined as ≤20% change in total profitability.
[c] Growth defined as >20% growth in total profitability.

A dependent one-sided t-test with Bonferroni adjustment[7] was used to conduct the pairwise comparisons (Keppel, 1982). Under the Bonferroni adjustment, we use $\alpha = 0.016$ to evaluate whether the prediction from a comparison model is significantly lower than the prediction from the proposed model. For both customer lifetime value and the NPV of future profit, the pairwise comparisons indicate that the simulation model provides significantly better predictions than all of the comparison models except for the simulation model without customer heterogeneity.

Table 6 provides some additional insight into the prediction accuracy of the simulation model. In this table, we compare the predicted change in profit to the actual change in profit. For simplicity of exposition, we form three groups of customers: 1) a "decline" group, defined as profitability declining more than 20%; 2) a "stable" group, defined as profitability staying within 20% of the previous value; and 3) a "growth" group, defined as profitability increasing by more than 20%.[8] We can see from this table that among the 56 customers that were predicted to decline in profitability, 50 (89%) were accurately predicted. Among customers that were predicted to be stable in profitability, 68 (73%) were accurately predicted, and 38 (76%) were accurately predicted to grow in profitability. Thus, most predictions were accurate for each predicted category. The off-diagonal elements show customers that were mispredicted according to this three-group scheme. For example, 13 (14%) customers were predicted to be stable in profitability, but the observed profits showed growth. In Appendix A we show that the percent of correct predictions ranged from 44% to 71% for the naïve models, such as current, average and trend in customer profitability. The proposed model therefore performs better than the alternative naïve models for predicting changes in customer profitability.

5. Discussion

5.1. Customer evolution

Following the ideas underlying the customer pyramid, many businesses target customers according to their profitability. More profitable customers typically receive more attention than do less profitable customers. The logical extension of this approach is to target customers according to their *future* profitability, or their customer lifetime value. The problem with this approach is that most attempts to predict future profitability have been notably unsuccessful. In fact, very simple models, such as assuming that current profitability levels will continue, have been shown to be just as good as more complicated models.

Thus, obtaining a better understanding of how customers evolve and obtaining accurate predictions of future profitability are challenges of pressing importance. We need to have a better

[7] The Bonferroni adjustment was used because multiple comparisons are performed on the same data.

[8] We also evaluated the model's accuracy by defining an increase in profitability of more than 10% as "growth", a decrease in profitability of more than 10% as "decline" and staying within 10% profitability as "stable". The substantive conclusions of the results do not change with this alternative definition.

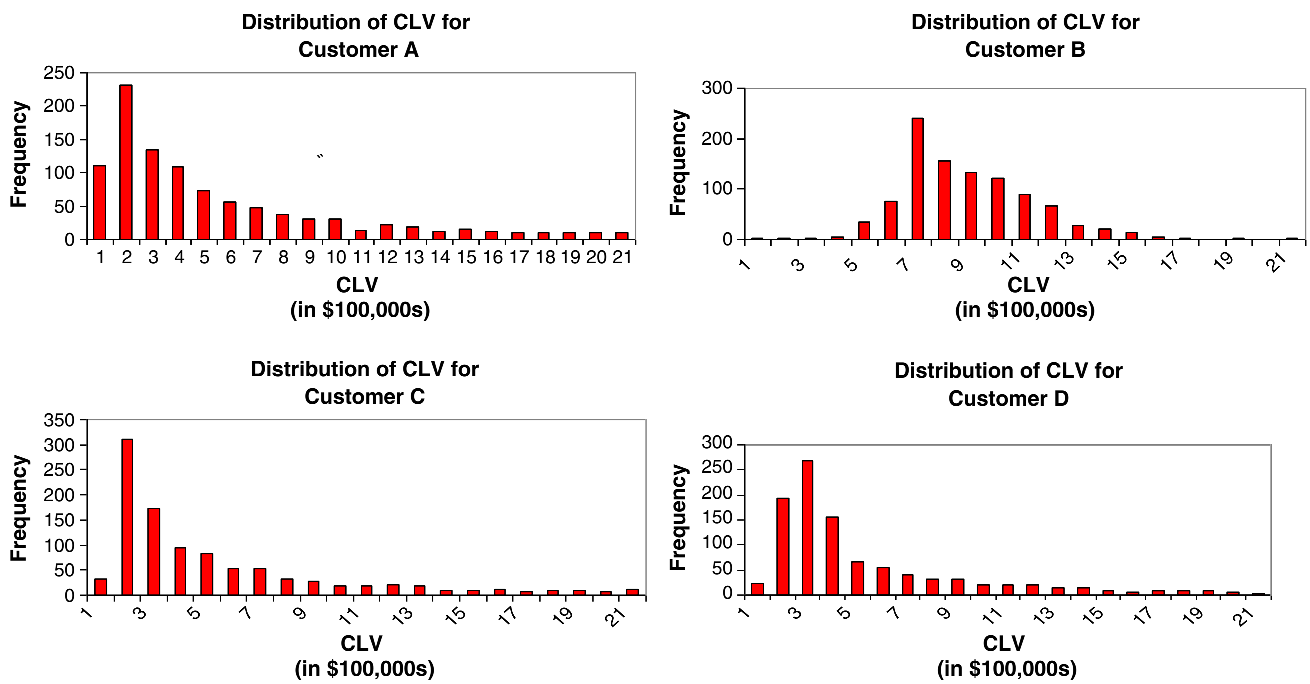

Fig. 4. Predictive distribution of future profitability (1000 simulated replications).

understanding of when customers will increase in profitability over time, when they will stay more or less the same, and when they will decrease in profitability over time. The purpose of this paper is to provide models that increase our understanding of how customers evolve over time.

5.2. Future customer profitability

Our research provides a new approach to predicting future customer profitability that outperforms existing methods. A model based on the simulation of future customer profits provides large improvements in prediction over the simpler models that have previously been shown to be the best in the literature.

The proposed simulation model performs the best by modeling purchase incidence and profit conditional on purchase. Given the endogeneity of marketing contacts, we also estimate future marketing contacts by modeling past contacts using an instrumental variables approach. We also model the joint error structure between the different equations in the model, assuming that the errors are correlated.

A key element of our model is a simulation approach. For each customer, we use the Monte Carlo simulation to project many possible futures, making random draws on the error terms. This approach results in many possible profitability trajectories (and total future profits) for each customer. We then average across the replications to obtain the expected future profitability for each customer for the time horizon desired. The results show that the new model produces large prediction improvements over the simpler competing models. In particular, the future profitability values predicted by the simulation model are more accurate than are predictions that result from projection of current profitability, projection of average profitability, use of a linear trend in profitability, or the BG/NBD approach. The simulation model also outperforms variants of the simulation model that do not utilize multiple replications, do not accommodate customer heterogeneity in the parameters or do not account for the endogeneity of marketing contacts. The results imply that an appropriate model specification and prediction algorithm would enable firms to obtain better returns from the investments they make in collecting detailed customer-level transaction information.

The simulation algorithm allows us to obtain a distribution of future profitability for each customer, as shown in Fig. 4 for four randomly selected customers. The distribution of CLV for these four customers is not normal and has a long right tail. The CLV distribution highlights the uncertainty of future profitability for a given customer. The highest density of simulated CLV values (about 47%) for customer A ranges from $100,000 to $300,000. The highest density of simulated CLV values (about 50%) for customer B ranges from $700,000 to $900,000. The distribution of CLV allows managers to determine a corresponding distribution of marketing resource allocation levels for each customer rather than a point estimate. For example, a manager may want to pay extra attention to a “high-variance” customer who has a decent chance of being very profitable in the future, even if the customer's expected future profitability was only average, to avoid missing a potential opportunity.

5.3. Managerial implications

Our research shows that using more sophisticated approaches can predict future customer profitability more accurately than simple managerial heuristics can. The proposed simulation model also provides managers with a range of profitabilities of each customer that can be used to develop several alternative customer-level resource allocation scenarios.

We find that accurately predicting changes in customer profitability has important implications for the success of customer resource allocation decisions. We illustrate these implications through a simulation exercise. Let us assume that in periods 13 through 24, the firm analyzed in this study has marketing resources (i.e., sales calls) that can be allocated to only about one quarter (25%) of its customer base. Given this resource constraint, the firm would prefer to allocate its marketing resources to customers that are expected to provide the highest profit in the future time periods, thereby increasing the total future net profit and return on investment (ROI[9]). Following the customer pyramid logic presented in this study, this would imply that the firm needs to identify the 50 or so customers who would fall in

[9] For our purposes, we define ROI as net profit divided by direct marketing costs.

Table 7
Customer Resource Allocation Implications. (Reported values are in $1000 s).

	Average observed net profit — periods 13–24	Average observed marketing cost — periods 13–24	Observed ROI — periods 13–24	Percent increase in ROI over average profitability model
Average profitability	1416	105	13.49	–
BG/NBD model	1545	87	17.76	31.70%
Simulation model	1794	84	21.36	58.30%

the highest quartile of net profit in periods 13 through 24 or future time periods (Table 4). As shown in Table 4, customers in the highest quartile in the future time periods may have belonged to any of the profit quartiles in periods 1 through 12 or in the past. Hence, we can expect models that accommodate changes in customer profitability to be better able to identify segments that will show high net profits in the future. Table 7 provides a measure of the predictive accuracy of the models for the entire customer sample, and the simulation exercise illustrates the implications of accurately predicting changes in customer profitability for the customers that give the company the highest profit. Accurately identifying future high-profit customers is a very important aspect of customer relationship management (Venkatesan & Kumar, 2004).

We focus our investigation on the best naïve model from our study, the average profitability model, which has been shown in the literature to perform as well as more sophisticated models, and the two main sophisticated models in our study, the BG/NBD model and the simulation model. Using each of the three competing models, we identify the 50 customers that are predicted to have the highest future net profits. For these 50 customers, we report the actual average observed values for the gross profits and the actual average marketing costs and ROI[10] for periods 13 through 24 in Table 7. Because we assume identical response coefficients for all customers in our model, we do not calculate the optimal marketing costs that would maximize future customer profits and ROI. In this analysis, we focus only on accurately selecting or identifying customers who will be profitable in the future. Once a customer is selected, we assume that the firm allocates marketing resources as per their status quo strategy. The ROI from a customer is therefore calculated from the observed marketing costs and profits in the holdout time periods, i.e., periods 13 through 24.

Targeting customers using the two more sophisticated models results in substantial improvements in ROI when compared with the average-profitability approach. In particular, the BG/NBD model produces a 31.7% improvement in ROI over the current profitability model, and the simulation model produces a 58.3% improvement in ROI. [11]The financial implications of using one of these models rather than focusing on the currently most profitable customers are large. The 50 most profitable customers received an average of $105,000 in marketing costs and reported an average of $1,416,000 in net profit, with a corresponding ROI of 13.49%. The customers predicted to be most profitable by the BG/NBD model received only $87,000 in marketing costs and yet produced $1,545,000 in net profit, indicating that more profit could be produced with less marketing cost by using the BG/NBD approach. The simulation model produced even stronger results, using only $84,000 in marketing costs to produce $1,794,000 in net profit. Thus, the simulation model produced $378,000 more net profit per customer, using $21,000 less in marketing costs per customer to do it.

The models presented in this paper are applicable when managers have available a database of customer purchase history and direct marketing actions. In such a case, we conclude that prediction of future profitability is not the futile endeavor that has recently been suggested by several recently published papers. More sophisticated approaches can improve accuracy significantly over simpler approaches. The drivers of purchase incidence and profit provide managers with early warning of changes in customer profitability and present them with methods of influence to improve customer profitability.

Although our research focuses on predicting total profitability over a fixed time horizon, the model could also be used to predict the net present value of future customer profitability over a given time frame. By increasing the length of the time frame, the model can also be used to estimate the customer's lifetime value, given that there is almost no contribution to customer lifetime value after about 15–20 years under the discount rates that managers typically use (Gupta & Lehmann, 2005).

5.4. Limitations

Several limitations of the current research should be noted. First, our conclusions are based on the analysis of a sample from one large-scale, industrial customer database from a single industry. Replication across multiple databases, companies, and industries would be necessary to fully establish the model's general applicability.

Second, an endogeneity issue inevitably remains. That is, if a company predicts future profitability for a customer, it is also implicitly assuming that its own marketing actions will be determined as in the past (as the instrumental variables prediction of marketing actions implies). However, if the company then uses the future customer profitability information to change its marketing actions, customer profitability will change further. The result is that increased attention to future customer profitability probably means that customers that are projected to be profitable may become even more profitable than projected. Likewise, customers that are projected to be unprofitable may become even more unprofitable than projected.

Third, our models are only applicable to companies that have a historical customer database that includes past purchase history, profitability (or contribution to profitability), and a history of direct marketing actions. There are many companies, especially those that make consumer packaged goods, that do not have such a database.

5.5. Directions for future research

Although we have developed models that outperform simple managerial heuristic models in predicting future customer profitability, we are well aware that further improvement is possible. This improvement can take several forms. The individual components of the simulation model—purchase incidence, profit conditional on purchase, marketing actions, and the simulation of future customer profitability paths—all are potential opportunities for improvement. Another possible way to improve the model is to build a unified model that solves the endogeneity of marketing actions issue by somehow building the altered marketing actions into the customer profitability projection

[10] Note that we are not able to manipulate the marketing costs for the purpose of this test; instead, we report the results of the actual historical marketing allocations for periods 13–24.

[11] We also compared against the trend in profitability model, which recorded a net profit of $1,502,000 from marketing costs of $95,000, giving an ROI of 15.81%. The BG/NBD model increased ROI by 12.3%, and the simulation model increased ROI by 35.1%, over the trend-in-profitability model. The simulation model yielded $281,000 more in net profit from $11,000 less in marketing costs.

itself. Finally, there is the opportunity to validate the model across a variety of companies, countries, and industries.

5.6. Conclusions

Predicting future customer profitability is not futile; it requires a sophisticated modeling approach. We provide a new approach that significantly outperforms naïve models. A model based on simulating customer futures produced large improvements over simpler models in predicting future profitability. Our proposed model, which includes simultaneous equations and a Monte Carlo simulation approach, outperformed competing models in a large-scale empirical test. We conclude that predicting future customer profitability is possible and can be used to drive customer-specific marketing actions. We may be able to predict systematically whether some of our “frogs” will become “princes” after all.

Acknowledgments

The authors thank the Editor, the AE and the anonymous reviewers for their valuable suggestions in revising this manuscript. The authors also thank the audiences at various conferences, including the Marketing Science Conference, the Frontiers in Service Conference, EMAC and QUIS, and at various universities. The authors also thank the anonymous firm that granted us access to the data. The authors thank Renu for copy editing the manuscript.

Appendix A. Additional model performance results

Table A.1
Correlations between predicted and observed CLVs in artificial data.

	Current profitability	Average profitability	Trend in profitability	Separate models for purchase propensity, gross profit and no IV	Include correlation between purchase propensity and gross profit (selection model)	Include marketing contacts IV	Include customer heterogeneity
	$\sigma^2_v = 1e5$, $\sigma^2_\varepsilon = 10$						
Purchase propensity model	NA			0.81	0.84	0.85	0.87
Gross profit model	0.77	0.76	0.79	0.82	0.85	0.88	0.91
Marketing contacts model	NA			NA	NA	0.91	0.94
	$\sigma^2_v = 1e10$, $\sigma^2_\varepsilon = 20$						
Purchase propensity model	NA			0.69	0.74	0.82	0.85
Gross profit model	0.68	0.69	0.70	0.72	0.78	0.83	0.86
Marketing contacts model	NA			NA	NA	0.79	0.85
	$\sigma^2_v = 1e15$, $\sigma^2_\varepsilon = 30$						
Purchase propensity model				0.62	0.72	0.77	0.84
Gross profit model	0.57	0.58	0.61	0.64	0.75	0.81	0.83
Marketing contacts model				NA	NA	0.75	0.82

Na= Not Applicable.

Table A.2
Current profitability model prediction accuracy — changes in customer profitability.

		Actual change in profitability		
		Decline	Stable	Growth
Predicted change in profitability	Decline[a]	35 (63%)	15 (27%)	6 (11%)
	Stable[b]	27 (29%)	42 (45%)	24 (26%)
	Growth[c]	14 (28%)	14 (28%)	22 (44%)

[a] Decline defined as >20% decline in total profitability.
[b] Stable defined as ≤20% change in total profitability.
[c] Growth defined as >20% growth in total profitability.

Table A.3
Average profitability model prediction accuracy — changes in customer profitability.

		Actual change in profitability		
		Decline	Stable	Growth
Predicted change in profitability	Decline[a]	37 (66%)	14 (25%)	5 (9%)
	Stable[b]	27 (29%)	44 (47%)	22 (24%)
	Growth[c]	13 (26%)	12 (24%)	25 (50%)

[a] Decline defined as >20% decline in total profitability.
[b] Stable defined as ≤20% change in total profitability.
[c] Growth defined as >20% growth in total profitability.

Table A.4
Trend in Profitability Model Prediction Accuracy — Changes in Customer Profitability.

		Actual change in profitability		
		Decline	Stable	Growth
Predicted change in profitability	Decline[a]	40 (71%)	12 (21%)	4 (7%)
	Stable[b]	24 (26%)	49 (53%)	20 (22%)
	Growth[c]	10 (20%)	12 (24%)	28 (56%)

[a] Decline defined as >20% decline in total profitability.
[b] Stable defined as ≤20% change in total profitability.
[c] Growth defined as >20% growth in total profitability.

References

Berger, Paul D., & Nasr, Nada I. (1998). Customer lifetime value: Marketing models and applications. *Journal of Interactive Marketing*, *12*(Winter), 17–30.

Blattberg, Robert C., & Deighton, John (1996). Manage marketing by the customer equity test. *Harvard Business Review*, *74*, 136–144 (July–August).

Bolton, Ruth N., Lemon, Katherine N., & Verhoef, Peter C. (2004). The theoretical underpinnings of customer asset management: A framework and propositions for future research. *Journal of the Academy of Marketing Science*, *32*(2), 271–292.

Brooks, Rick (1999). Alienating customers isn't always a bad idea, many firms discover. *Wall Street Journal*, *A1*, A12 (January 7).

Campbell, Dennis, & Frei, Frances (2004). The persistence of customer profitability: Empirical evidence and implications from a financial service firm. *Journal of Service Research*, *7*(2), 107–123.

Chintagunta, Pradeep (1999). Variety seeking, purchase timing, and the lightning bolt brand choice model. *Management Science*, *45*(4), 486–498.

Donkers, Bas, Verhoef, Peter C., & de Jong, Martijn (2007). Modeling CLV: A test of competing models in the insurance industry. *Quantitative Marketing and Economics*, *5*(2), 163–190.

Donkers, Bas, Paap, Richard, Jonker, Jedid-Jah, & Frances, Philip Hans (2006). Deriving target selection rules from endogenously selected samples. *Journal of Applied Econometrics*, *21*(5), 549–562.
Ebbes, Peter, Papies, Dominik, & van Heerde, Harald J. (2010). The sense and non-sense of holdout sample validation in the presence of endogeneity. May 1, Available at SSRN: http://ssrn.com/abstract=1635250
Fader, Peter S., Hardie, Bruce G. S., & Lee, Ka Lok (2005). Counting your customers the easy way: an alternative to the Pareto/NBD model. *Marketing Science*, *24*(2), 275–284.
Fader, Peter S., Hardie, Bruce G. S., & Lee, Ka Lok (2005). RFM and CLV: using iso-value curves for customer base analysis. *Journal of Marketing Research*, *42*(November), 415–430.
Greene, William H. (1993). *Econometric analysis* (3rd Edition). Upper Saddle River, NJ: Prentice Hall, Inc.
Gupta, Sunil, & Lehmann, Donald (2005). *Managing customers as investments: The strategic value of customers in the long run.* Upper Saddle River, NJ: Wharton School Publishing.
Jackson, Barbara B. (1985). *Winning and keeping industrial customers.* Lexington, MA: D.C. Heath and Company.
Jain, Dipak, & Singh, Siddhartha S. (2002). Customer lifetime value research in marketing: A review and future directions. *Journal of Interactive Marketing*, *16*(2), 34–46.
Keppel, Geoffrey (1982). *Design and analysis: A researcher's handbook.* Englewood Cliffs, NJ: Prentice-Hall.
Kumar, V., Venkatesan, Rajkumar, Bohling, Timothy, & Beckmann, Dennis (2008). The power of CLV: Managing customer lifetime value at IBM. *Marketing Science*, *27*(4), 585–599.
Malthouse, Edward C., & Blattberg, Robert C. (2005). Can we predict customer lifetime value? *Journal of Interactive Marketing*, *19*(1), 2–16.
Manchanda, Puneet, Rossi, Peter E., & Chintagunta, Pradeep K. (2004). Response modeling with nonrandom marketing-mix variables. *Journal of Marketing Research*, *41*(November), 467–478.
Mulhern, Francis J. (1999). Customer profitability analysis: Measurement, concentration, and research directions. *Journal of Interactive Marketing*, *13*(Winter), 25–40.
Reinartz, Werner J., & Kumar, V. (2003). The impact of customer relationship characteristics on profitable lifetime duration. *Journal of Marketing*, *67*(January), 77–99.
Rust, Roland T., Lemon, Katherine N., & Zeithaml, Valarie A. (2004). Return on marketing: Using customer equity to focus marketing strategy. *Journal of Marketing*, *68*(1), 109–127.
Rust, Roland T., & Verhoef, Peter C. (2005). Optimizing the marketing interventions mix in intermediate-term CRM. *Marketing Science*, *24*(3), 477–489.
Rust, Roland T., Zeithaml, Valarie A., & Lemon, Katherine N. (2000). *Driving customer equity: How customer lifetime value is reshaping corporate strategy.* New York: The Free Press.
Stephens, M. A. (1974). EDF statistics for goodness of fit and some comparisons. *Journal of the American Statistical Association*, *69*(September), 730–737.
Venkatesan, Rajkumar, & Kumar, V. (2004). A customer lifetime value based framework for customer selection and resource allocation strategy. *Journal of Marketing*, *68*(October), 106–125.
Venkatesan, Rajkumar, Kumar, V., & Bohling, Timothy (2007). Optimal CRM using Bayesian decision theory: An application to customer selection. *Journal of Marketing Research*, *44*(4), 579–594.
Villas-Boas, Miguel J., & Winer, Russell S. (1999). Endogeneity in brand choice models. *Management Science*, *45*(10), 1324–1338.
Wuebben, Markus, & Wangenheim, Florian (May, 2006). *"Counting Your Customers the Managerial Way," presentation at the European Marketing Academy Meetings, Athens.*
Zeithaml, Valarie A., Rust, Roland T., & Lemon, Katherine N. (2001). The customer pyramid: Creating and serving profitable customers. *California Management Review*, *43*(4), 118–142.