Papers

Behram Hansotia

is the co-founder of InfoWorks, where he was president and CEO until June 2004. He has a PhD in management science from the University of Illinois at Urbana-Champaign and his work experience includes several years in academia, industry and management consulting. He serves on the Editorial Board of the Journal of Database Marketing & Customer Strategy Management and has published more than 40 papers, mostly on database marketing.

Keywords: cash flow, hazard rates, hazard models, mixed regression models, stationarity, Markov chains

Defining customer equity

Behram Hansotia
226 North Street
South Haven, MI 49090
USA
Tel: +1 (269) 216 4188
E-mail: bhansotia@comcast.net

Issues and challenges in estimating customer equity

Behram Hansotia

Received: 26 January 2006

Abstract

This paper starts by describing key applications of customer equity (CE), followed by a discussion of estimating CE using Markov chains and statistical modelling techniques. A simulation-based approach and an analytic approach are presented for the Markov chain models, and direct and components-based approaches are presented for the statistical methods. In the components-based approach, methods for estimating the key inputs of CE, namely customer cash flows, revenues, attrition rates, purchase timing and purchase probabilities, are discussed. Both subscription services and product-purchase-type businesses are addressed. The key assumption of stationarity, needed for estimating CE, is also presented. The paper ends with a very brief discussion of some of the organisational requirements for implementing a CE-based marketing system.

Introduction

The concept of customer equity (CE), which is closely related to customer lifetime value (CLV), is central to the rapidly emerging strategy of running a business, namely CRM, or customer relationship management. CE is generally defined as the total value of current and future customers. Here, customers are viewed as assets and like all assets have value and the potential of generating cash flow streams. The net present value (NPV) of these customer-generated cash flows is referred to as CE. Implementing a CRM strategy essentially comprises designing and implementing a series of processes and tools that help a company increase the value of its current and future customers. Since a company's equity is the NPV of its cash flows and since customers are responsible for generating the company's cash flows, enhancing CE results in enhancing the equity of the company. Hence a CRM strategy is all about increasing the value of the company.

Since customers vary in the amount they purchase, the types of products they purchase, the frequency of their purchases and the duration of their relationship with a company, typically there is considerable variance in the magnitude of customers' equities. Hence, to be a useful tool, CE estimates need to be made at the customer level, or at least at the customer segment level. Since customer-level estimates of cash flows and relationship duration have very high variances, it is recommended that CE is estimated at a segment level. In this paper CE is used interchangeably with CLV: based on the application it will discuss the expected CE of a prospect, the CE of new and veteran customers and the firm's total CE,

which may or may not include the value of future customers based on the application.

The key applications of customer equity are in:

- developing decision rules for customer acquisition, development and retention
- managing and tracking customer performance based on key inputs used in computing CE
- assessing the value of a company.

This paper starts by discussing how CE can be used as a practical tool in customer acquisition, development and retention, in managing the customer base and in assessing the value of a firm. The next section presents issues dealing with estimating cash flows, specifically the estimation of different costs. It also presents the key requirements for estimating CE and the two broad approaches to estimating CE through Markov chains and regression modelling. There is further discussion of the Markov chain modelling approach and how CE can be estimated using Monte Carlo simulation and through analytic techniques. The paper then turns to modelling approaches for computing CE for subscription services as well as for businesses that sell products. A direct method and a components-based approach are presented. The next section presents the key assumption in using the models' outputs to estimate CE, and practical approaches to handling the situation when the assumptions are not met. This is followed by a brief discussion of some of the key organisational requirements for implementing a CE-based strategy for running a company. The paper concludes with a brief summary and touches on the complexities of estimating CE in highly dynamic markets.

Applications of customer equity

The key application of CE is in customer life-cycle management. The concepts underlying CE can be harnessed to develop an entire customer-focused marketing system with the key goal of enhancing the value of the customer franchise.¹⁻³ Customer life-cycle management includes acquiring the right types of customers, developing both new and veteran customers, proactively retaining customers at risk, saving customers who have indicated that they wish to terminate their relationship with the company and sometimes winning back lapsed customers. All these activities require marketing effort, and CE and its inputs can provide the framework for making decisions on marketing spend.

Customer acquisition

A key task of marketing is to acquire new customers. Based on the industry and the company's business model, marketers use a variety of channels to acquire customers. Financial services companies, for instance, may use direct marketing, including the internet and the mail, as well as print, TV, radio and billboards. The fundamental criterion in determining the marketing spend level is that at no time should the acquisition cost of a customer exceed the customer's equity. Since the

Decision rule for customer acquisition

acquisition cost may be calculated as the ratio of the cost of the marketing contact to the prospect's probability of responding, this rule is also equivalent to marketing only to those prospects whose expected CE exceeds the cost of the marketing contact. Here, expected CE is the product of the probability of a prospect becoming a customer as a result of the marketing effort and the new customer's estimated CE. To the extent that any pool of prospects can be segmented, this rule should also be applied at the segment level.

For instance, in the case of direct marketing based on a large list of prospects, if the company has some information on the prospects then this list may be segmented based on expected CE. In practice this would require models to estimate the probability of response of each prospect (segment) and the prospect's CE on becoming a new customer. Only prospects whose expected CE is larger than the marketing contact cost should then be selected for the marketing effort (see Hansotia and Wang⁴ for different scenarios relating to customer acquisition decision rules).

In the case of the internet, where the company has the opportunity to place advertisements at portals or websites, or to use search engines, CE needs to be estimated by the source of the new customer. For instance, if a website helps generate 100 new customers a month, each with an average CE of \$30, and the cost of the ad on the website is \$4,000 a month, this website does not make any economic sense. The total customer equity generated is \$3,000 while the marketing cost is \$4,000.

The analysis just described is static, and the company should continue to monitor the acquisition rate and the CE of the customers being generated at that site on an ongoing basis. If we assume that in the short term the average CE does not change, then at least the acquisition rate needs to be monitored on a monthly basis. Typically, acquisition rates from an ad will fall over time and the company may want to try a different ad, or even take the ad off for some time by using a pulsing advertisement placement policy. This implies ongoing testing of new creative formats and messages as well as the timing of ad placements. This of course is critical in analysing the efficiency of all advertising.

All acquisition efforts where prospects contact the company as a result of some advertisement they have read, seen or heard should be based on acquisition rates, the cost of the ad and the CE of the customers being generated by the ad. In theory, all advertisements that generate a positive net expected CE (expected CE after marketing cost) should be deployed.

Customer development

Soon after a new customer is acquired, the challenge of developing the customer begins. This may entail enticing the customer to place a second order by using a great relevant offer. It is recommended that companies separately focus their energies on developing new and veteran customers. If they do not, new customers may fall through the cracks in any customer development efforts. It is remarkable how many companies have nearly 50 per cent of their customer base as one-time buyers. New customers often behave differently than veteran customers, since their knowledge of the company is often limited and so is their propensity to buy from the

New versus veteran customers

company. From the company's standpoint, there is limited information on new customers, and if they are lumped with veteran customers they will not look as attractive. If a company prioritises its customer development efforts, it may then inadvertently neglect some new customers who may have significant potential for development.

In estimating CE the typical assumption is that the company's marketing practices are fixed. Since marketing practice impacts on customers' behaviours, namely purchases and relationship duration, one cannot simultaneously estimate CE in an ever-changing marketing environment or a rapidly changing marketplace. This is reminiscent of the famous Heisenberg principle in physics: one cannot simultaneously measure the position and velocity of an electron, since the very act of measurement imparts energy to the electron, affecting the results. Hence it is recommended that the company fix its marketing practices first before attempting to estimate CE. (This assumes that minor variations in marketing practice do not have an appreciable impact on customer cash flows or attrition rates.)

As far as new customer development goes, unless a company is absolutely certain that a specific new customer will not buy again (because he has passed away or left the country for good), at least some effort should be made to develop each and every one of them. This does not mean that the same effort needs to be applied to each new customer. CE models based on prospect information as well as information from the initial purchase could be used to estimate the remaining value (CE less the cash flow from the initial purchase) of each new customer. The marketing spend over this initial phase of the customer relationship can then be made proportional to each customer's predicted remaining value.

In developing veteran customers, most companies that use financial decision rules to determine which veteran customers they should contact typically use a short-term decision rule: only those customers whose *increase* in expected cash flow (as a result of the development effort) is estimated to exceed the cost of the marketing effort are targeted for development.^{5,6} Here the development effort could be a single contact with an offer, or a series of contacts over time. Rather than the increase in expected CE as the guiding metric, most companies will typically use increase in expected cash flow over a given time period, since the objective usually is to create an immediate increase in profits. (One would of course love to create a permanent increase in profits, but unfortunately that does not happen frequently.) But one still needs to be careful in testing and analysing such development efforts to ensure that they do not cannibalise sales excessively from future periods. This can be done by monitoring customer cash flows for both treated and control groups before and after the marketing campaign.

Customer retention

Customer retention starts at the time of customer acquisition if a company uses a CE-based decision rule to target prospects. Since retention rates are a key input in computing CE, such companies will typically be acquiring customers with higher-than-average retention rates. Since the customer

acquisition rule targets prospects with the highest net expected CE, the prospects who have high expected relationship durations as customers will typically always be targeted.

Decision rules

The decision rule for selecting customers for proactive retention efforts is identical to that of developing veteran customers: target only those customers whose increase in expected cash flow exceeds the cost of the retention effort. Since the objective of the retention effort first and foremost is to reduce attrition rates, the increase in expected cash flow should be due to improvements in retention rates. The decision rule for saving customers who have indicated that they wish to terminate their relationship with the company and the decision rule for targeting lapsed customers (winback) is basically the same as the rule for acquiring new customers, except that in this case it is the CE associated with the second relationship with the company.

Selection of marketing projects

Optimising marketing spend

In managing marketing based on CE, the marketing vice president needs to allocate sub-budgets for customer acquisition, retention and development. In each case, however, the marketing dollars spent, whether on acquisition, retention or development, should create increases in CE larger than the dollars expended, or else the company is destroying value. If the change in net expected CE of the customer acquisition, retention and development projects is known, then as a first step one would want to consider only those projects with (positive) increases in net expected CE. If the total cost of these projects exceeds the marketing budget, then obviously all the projects cannot be undertaken. The optimal solution to identifying the best subset of projects that maximise the total net expected CE created under the budgetary constraint can be obtained by solving an integer (0/1) programming problem, referred to as the knapsack problem.

The mathematical formulation of this basic knapsack problem is presented in Table 1. (This capital budgeting problem is referred to as the knapsack problem because it has the same structure as a situation where a decision maker must select objects of different weights and values to fill his knapsack. He has to pick the objects so as to maximise the total value of his knapsack, constrained by the total weight he can carry and the requirement that he pick no more than one object of each type.)

Table 1: Capital budgeting as a knapsack problem

Assume there are a total of N projects with positive net expected CE values.
Let x_i be the set of decision variables, such that $x_i = 1$ implies the project is selected and $x_i = 0$ implies the project was not selected, for $i = 1, \dots, N$
Let v_i be the net expected CE for project i and $v_i \geq 0$, for $i = 1, \dots, N$
Let c_i be the cost of project i , for $i = 1, \dots, N$ and let B be the total marketing budget.
If Z represents the total net expected CE generated by the set of selected projects, the capital budgeting knapsack problem is to find x_i so as to:
Maximise $Z = \sum_i (v_i x_i)$ for $i = 1, \dots, N$
Subject to: $\sum_i (c_i x_i) \leq B$ and $x_i = 0$ or 1 for $i = 1, \dots, N$

This mathematical programming approach is most useful when there are several discrete opportunities, or projects, for increasing CE through acquisition, development and retention. Lining up these projects based on total increase in CE and picking off the highest-ranked projects until the budgetary constraint is met will not necessarily yield the optimal project set.

The basic capital budgeting problem can easily be extended to include different budgetary requirements for each of the different project types — acquisition, retention and development — as well as other constraints deemed necessary by management, for a variety of reasons. Problems like the knapsack problem can be readily solved today in EXCEL with an appropriate solver add-on. Finally, once such a model is developed it becomes a useful negotiation tool with the company to explore the impact of budgetary increases on increases in CE. If necessary, the optimal solution can be linked to a pro forma income statement spreadsheet to describe the impact on the company's income over multiple years.

Assessing firm value

If a business is for sale, how should its value be imputed? Typically companies use a multiplier of sales, earnings or EBITDA (earnings before interest, taxes, depreciation and amortisation), depending on the industry. A far better way of assessing value may be through CE. Here customers of different quality may be grouped into segments and each segment's CE can be assessed, based on historic purchasing patterns, attrition rates and marketing spend. The sum of the segment CEs should then be a fair assessment of the value of the customer base. To value an ongoing concern, however, one may want to add the CE of all future customers to the CE of the current customer base.⁷

The individuals purchasing a new business often believe they can run it better, or believe they may be able to integrate it with their existing company or processes so as to increase the value of the business or the combined companies. Unless the business can be purchased for an amount significantly lower than the value imputed based on CE, it should not be purchased unless additional CE can be created.

Estimating customer cash flows

Since CE is the NPV of customer-generated cash flows, estimating cash flows is a critical component of estimating CE. Note that cash flow is not the same as profits or earnings. The latter are computed primarily for tax purposes and include depreciation and amortisation (accrued costs) as deductible costs. At the firm level, cash flow is computed for a given time period and is the difference between revenues and all expenditures, including fixed and capital costs incurred over that time period. This is the money the firm has left over, and hence is often referred to as free cash flow.

Computing cash flows at the customer level raises a whole host of issues regarding cost allocations and how fixed costs should be allocated, if at all, in the computation. For instance, in the decision rule for customer acquisition, where the rule stipulates that customer acquisition

Cost component of cash flows

costs should not exceed CE, fixed costs should not be included in the calculations since they do not change in the short term. Another way of stating the acquisition decision rule is that the increase in CE (from new customers) should exceed the marketing contact cost. As long as the acquisition costs are all variable, fixed costs do not need to be allocated since they cancel out in computing the total increase in CE.

The decision whether to include fixed costs or not should be made depending on the application. For instance, to assess the current value of an ongoing business one would compute the CE of all current customers (that is, their remaining value) and add to that the CE of all customers the firm would acquire in the future. Note, however, if a firm is for sale, the buyers may want to pay only for existing assets and thus may ascribe a value only to current customers. In such instances, where one is valuing a business as the aggregate value of its customers, all fixed costs need to be allocated to customers, or the present value of total fixed costs (which in this case could change over time) should be added to the total CE, computed using only variable costs.

The key costs in computing customer cash flow depend on the business. For instance, for a retailer some of the variable costs include:

- cost of goods sold
- purchasing costs (buyers, etc)
- cost of returns
- cost of bad debt on consumer purchases
- cost of selling goods (sales clerks)
- cost of capital tied up in the inventory
- direct mail and e-mail promotions costs.

Fixed costs include:

- store capital costs (mortgage/rent/construction)
- utilities associated with the store
- head office costs, including management
- cost of liquidating excess inventory
- advertising and website maintenance costs.

Before any costs are estimated, a good discussion with the company's cost accountants is recommended to identify all relevant costs, how they should be allocated at the customer level and when they should be applied. It is best to allocate costs based on customer activity (ABC, or activity-based costing⁸), since this will enable a company to identify if certain customers are generating a disproportionate amount of certain costs, such as returns or bad debt. There may be several ways to allocate costs. The crudest way may be to allocate the total cost across all customers, but this may not yield any insights and could also be extremely misleading.

Purchasing costs could be allocated to the products purchased and then would accurately reflect the actual cost of purchasing different products. Returns, however, are a function not only of the types of products sold but

Estimating costs through models

also of customer characteristics, and return costs also depend on how the company handles returns. To estimate the expected average return cost for each purchase occasion, the company may want to consider building models with key product and customer characteristics as the predictor variables, to estimate the probability of a return and the average monetary amount of the return.

In building such models, one needs to account for the correlation among observations when customers contribute more than one observation to the modelling dataset. (If this is not done, the standard errors of the model's parameters will be biased, leading one potentially to draw incorrect conclusions about the significance of model parameters.) This can be accomplished using hierarchical regression (PROC MIXED in SAS⁹) where additional covariance parameters are estimated. Each purchase occasion can now be scored with the two models. The models' scores can then be combined into a single expected returns score by multiplying the two scores.

As described above, estimating costs at the customer level based on customer activities can quickly become a challenging task. A statistical approach is suggested to estimating the cost of returns for each transaction, so as to get maximum insight into the drivers of this behaviour, but the average cost of returns could be generated for any grouping of customers and products. The rows of this matrix could be defined customer segments and the columns product groupings. For each cell one could then compute the average cost of returns per purchase transaction. This approach, though an improvement over computing a global average, will not provide the detailed insights regarding specific variables that best explain returns behaviour, nor will it identify customer types creating the biggest returns problem.

The key challenge in allocating costs is the decision regarding the base of the calculation. Should it be the product purchased or the customers who made a purchase in the time period for which the cash flow is being computed, or a combination of the two, or all active customers whether they made a purchase or not? There is no simple answer to this question. A more detailed analysis is recommended if the cost is relatively high as a percentage of total costs and if there is significant variance in costs across customers.

Allocating costs

A final point in estimating cash flows is to ensure that all cash flows are first standardised to the same year (base year). This eliminates the effect of inflation. For instance, if the study period is five years in duration, say 2000–2004, and if one designates 2004 as the base year, the cash flows for the years 2000–2004 need to be factored up so they are represented in the value of the currency in 2004 (dollars, pounds, etc) This can be achieved using a price deflator/inflator index for each year. As an example, if the industry-wide prices of the products sold by the company increased by 2 per cent from 2000 to 2001, by 1.5 per cent from 2001 to 2002, by 2 per cent from 2002 to 2003 and by 2.5 per cent from 2003 to 2004, then the price inflator index for the year 2000 would be $(1.02) \times (1.015) \times (1.02) \times (1.025)$ or 1.085. All cash flows in 2000 would now be increased by a factor of 1.085. Similarly, the price inflator for the

year 2001 would be 1.063. Since CE is a forward-looking calculation, when one computes a five-year CE one is implicitly assuming that the same inflation rates will prevail over the years 2005–2009. If one is not comfortable with this assumption, one could use a constant inflation rate, say the mean (2 per cent in this case), to compute the price inflator indices.

When cash flows are computed in real dollars or pounds it is important that the discount rate used in computing the CE is also net of inflation effects. The real discount rate would then reflect only the uncertainty in cash flows and the cost of debt. In practice, the firm's cost of capital (weighted average of cost of equity and the cost of debt) could be used as the discount rate.

Approaches to estimating CE

There are two requirements for the methods discussed in this paper for estimating CE.

- Availability of a customer database with information on customer and prospect characteristics and longitudinal data on customer transactions and interactions with the company.
- A relatively stable marketing environment resulting in customer behavioural processes that do not change dramatically. Processes with this property are said to be *stationary*; a fundamental requirement of all statistical modelling. This requirement is discussed in greater detail later, along with a more rigorous definition of stationarity. The Markov chain model discussed below, however, does not necessarily require this property.

One or both of these requirements will obviously not be met for new companies or when there are radically new and innovative developments in the marketplace, such as the first few years of the commercial use of the internet. None the less, the concept of CE and its applications, as discussed above, is still valid, although the estimation of CE would now have to be based on additional managerial inputs of product usage and attrition rather than hard behavioural data alone.

In such cases a hybrid approach is recommended: use whatever data are available and marry them with managerial predictions of future customer behaviours. As hard data become available over time, update the calculations with the new data and new managerial predictions. Initially start with simpler models, and as the organisation's comfort grows with analytic-based approaches, continue to refine the models. Realistically, for all organisations, modelling is a journey. With the right managerial support and analytic staff a company should be able to continue to refine and improve its analytic decision-making capabilities.

CE at the customer or segment level may be estimated through Markov chain modelling or statistical models. If a Markov chain model is used, two approaches are feasible for estimating CE: Monte Carlo simulation or a recursive calculation-based method, reminiscent of dynamic programming.

New companies/ products

Two approaches may also be used if statistical modelling is deployed: the direct method, where the CE is modelled directly using a multiple regression model, or a components approach where the key components of CE, namely attrition/purchase probabilities and revenues, are first estimated through models and then the model predictions are used to compute a CE value. The second approach will typically provide greater insights into the drivers of these key behaviours and hence, if resources are available, is often preferred.

Markov chain models for product purchases

The Markov chain approach to computing CE is popular in industry, particularly in the catalogue and retailing industries, primarily because it is relatively easy to set up and use. As will be seen, however, it provides somewhat fewer insights into the characteristics driving customer behaviour.

The approach posits that a customer may be in one of several states and progresses, probabilistically, over time across the different states (often referred to as the system's probabilistic laws of motion). The Markov chain is a discrete model, in that typically there are a small number of states and time is also discrete, measured say in months, quarters or years. For the purpose of modelling CE, the state variable may be defined in a variety of different ways. For instance, the state may be defined by the combination of number of purchases and average purchase amount (revenues) in a period or by the cash flow in a period, or by the total amount purchased in a period. An example of the latter, where the state variable s assumes four values, would be: $s = 0$ for revenues equal to zero, $s = 1$ for revenues greater than zero but less than or equal to \$100, $s = 2$ for revenues greater than \$100 but less than or equal to \$300 and $s = 3$ for revenues greater than \$300.

Transition matrix and stationarity

Central to a Markov chain is the transition matrix of transition probabilities. For the above example, the transition matrix will typically be a 4×4 (four rows and four columns) matrix, with each cell (i, j) representing the conditional probability of the system being in state $s = j$ at time t , given it was in state $s = i$ at time $t - 1$. The Markov chain is said to be stationary if these conditional probabilities are the same for all values of t . This requirement is, however, not essential for estimating customer CE using a Markov chain. If a Markov chain is not stationary, it will have different transition matrices for each transition. If one considers the first purchase to occur at time zero and the total amount purchased in the first year to determine the value of s , the state variable, then the system can never start in state $s = 0$. Hence the first transition matrix will be a 3×4 matrix, since the system can always enter state $s = 0$ in the second year. Also, if a customer does not buy anything in his second year, he may buy nothing in year three or he may again buy something in his third year with the company. This would mean that the second transition matrix will be 4×4 . That is, he can start in state zero at the end of his second year with the company and remain in that state at the end of year three, or buy something, putting him in one of the other states, $s = 1, 2$ or 3 , at the end of year three.

It may very well be that there are different transition matrices for the initial years with a company, and after this time has elapsed the system may settle down into stationarity. For example, one may have the initial transition matrix $T_{3 \times 4}(1)$, a second different transition matrix $T_{4 \times 4}(2)$ and thereafter the same transition matrix, ie $T_{4 \times 4}(3) = T_{4 \times 4}(4) = \dots = T_{4 \times 4}$. It is recommended that the equality of the transition matrices be tested rigorously to ensure that differences are not due to sampling variances.

Estimating CE through simulation

Monte Carlo simulation can easily be used to model the progression of the system over time. To do that for each state one first needs an estimate of the customer's cash flow when she is in that state. One also needs to estimate the transition matrix and determine if it is stationary or becomes stationary after a few periods. This is best accomplished by tracking a large cohort of customers over time, identifying the states they pass through and then computing the average cash flow associated with each state. The percentage of customers that move from state i to state j provides an estimate of the transition probability t_{ij} . As discussed above, one also needs to test whether the transition matrices are stationary. If they become stationary after a few transitions, one will need separate estimates of the initial transition matrices and the stationary transition matrix.

Applying Monte Carlo simulation

Once the cash flows and transition probabilities have been estimated, the Monte Carlo simulation calculations can proceed. For the example discussed above, where there are four states based on annual revenues, one assumes a customer starts in state $s = 1$ and by generating a random number between zero and one (from a uniform distribution) one identifies the end state the customer will occupy after the first transition. For instance, if the first row of the initial transition matrix, $T(1)$, is (0.3 0.4 0.2 0.1), corresponding to, for year two, a 30 per cent chance of zero revenues and a 40 per cent chance of revenues under \$100 etc, then if the random number that is generated is say 0.65, it implies that the customer will land in state $s = 1$, since the random number is between 0.3 and 0.7 (0.3 + 0.4). But if the random number equals say 0.83 one assumes the customer will land in state $s = 2$, since 0.83 is larger than 0.7 but less than 0.9.

This process is repeated over several transitions, using the appropriate transition matrices to place the customer in the corresponding state at the end of each transition and always noting the state the customer lands in. Then, using the estimated cash flows associated with each state the customer passes through, one calculates the NPV of this customer. One can decide to compute an NPV over a finite number of years, to be conservative, or one can continue this process until the NPV does not increase substantively.

This process is again repeated several times, always starting in state 1 and computing the NPV of each customer. The average of these NPV estimates is the estimated expected NPV of a customer starting in state $s = 1$. This process is then repeated for starting states $s = 2$ and 3. One

can also compute confidence intervals around these estimates to recognise the sampling error of a finite number of simulations.

Analytic approach

A long-term CE for each starting state i can be calculated when the transition matrix, T , is stationary. If one denotes C_i as the annual cash flow associated with state i , d the periodic discount factor, $(1/(1 + \text{discount rate}))$, t_{ij} as the transition probability of moving from state i to state j , M as the number of states and V_i as the expected long-run CE for a system starting in state i and continuing indefinitely, then V_i may be obtained by solving the following equations:

$$V_i = C_i + d * \sum_j (t_{ij} * V_j), \text{ for all states } i \text{ and } j = 0, 1, \dots, M$$

The second term above represents the long-run expected CE discounted by one period, and when it is added to the cash flow in the first period associated with state i one gets the long-run expected CE at the start of the process for state i . For simplicity, the above assumes the cash flows occur at the start of each period, but the equations can easily be modified to accommodate cash flows occurring at the mid-point of each period. There are $M + 1$ equations above and $M + 1$ unknowns, and hence the V_i may be easily obtained by solving these equations simultaneously.

If the transition matrix becomes stationary only after a few periods, the solution of the above equations provides the remaining value or CE for each state from this point onwards. Recursively one can then calculate the expected remaining value for each state for each of the earlier periods. The probabilities t_{ij} will now have to correspond to the values from each of the different transition matrices associated with each period.

Recursive calculations

If one wishes to compute the CE over only the initial periods, say five years, and the appropriate transition matrices have been estimated, a recursive approach can be used to estimate the five-year CE for each starting state. If one denotes V_i^n as the CE for state i with n periods to go and t_{ij}^n as the corresponding transition probability of moving from i to j , then V_i^1 represents the CE for state i at the end of period 4 and V_i^2 the CE for state i at the end of period 3. V_i^2 may now be calculated by the following:

$$V_i^2 = C_i + d * \sum_j (t_{ij}^1 * C_j), \text{ for } i \text{ and } j = 0, 1, \dots, M \text{ and where}$$

$$\sum_j (t_{ij}^1 * C_j) \text{ represents } V_j^1$$

In general, one has:

$$V_i^n = C_i + d * \sum_j (t_{ij}^{n-1} * V_j^{n-1}), \text{ for } i \text{ and } j = 0, 1, \dots, M.$$

The above equations allow one to move recursively backwards, where one uses the CE from the next period to calculate the CE at the earlier period.

The Markov chain modelling approach is relatively straightforward to apply, but it does require the analyst to define the potential states of the system; since the number of states is finite, based on how they are

defined, the CE estimates can have considerable variation. Also, non-statistical methods do not provide the insights regarding the drivers of key behaviours that the regression-type models provide.

The above approach may be extended to determine optimal treatment policies that maximise CE. Here, management has a finite set of customer treatments for each state i and the objective is to determine the optimal treatment whenever the customer is in state i . The probabilistic laws of motion are now a function of both the state of the system and the specific treatment policy. Such optimal decision policies may be calculated using the approaches of Markov decision processes. That, however, will be the subject of another paper.

Statistical approaches to estimating CE

As mentioned earlier, two approaches may be used to estimate CE using regression models: the direct approach and the component approach. The paper first discusses the direct approach for both subscription services and businesses that sell products. It next presents the component approach and separately takes up the two types of businesses. For subscription services it discusses the modelling of attrition rates and revenues and cash flows, and for businesses that involve product purchases there is a discussion of purchase rates and the estimation of revenues and cash flows. For each business type the paper then presents formulae for combining these component estimates into estimates of CE.

Direct approach

In this approach, a cohort of customers is tracked over several periods, called the study period, and the cash flow and resulting CE for each customer are calculated over this study period. The computed CE is then used as the dependent variable in a multiple regression model that predicts it as a function of prospect/customer characteristics. Only prospect characteristics may be used if one wishes to use the estimated CE in a decision rule for acquiring new customers. But if one wishes to estimate the remaining value of existing customers who have been with the company for one year, say, the CE (remaining value) for the customers is calculated starting in year two over the next several periods. In such instances both prospect characteristics as well as variables computed based on the first year's transactions may be used as predictor variables in the model.

This approach may be used for both subscription services businesses and businesses that sell products. In the former case, however, the CE is calculated only over the periods the customer is active. For instance, if a customer terminates in period three, then his CE is computed over the first three periods. In the latter case the CE is computed over the entire study period. Based on assumptions, the cash flow in periods where no purchase occurs may be zero or some negative value if certain fixed costs continue to be allocated over all periods. If the study period is say N periods long, the model would estimate an N period CE. Note that this will underestimate the true CE over the entire life of the customer relationship.

Multiple regression

Censored regression

For subscription services, however, if the study period is long enough so that say 20–25 per cent of the cohort are still active at the end of the study period, a lifetime-based CE may be estimated using censored regression. The customers who are active at the end of the study period are deemed censored observations, in the sense that their lifetime CE is not observed. What one does know, however, is that their CE will always be greater than, or at most equal to, the ones used in the regression model. Censored regression makes this explicit in estimating the model parameters through a method called the *maximum likelihood* approach. The method works reasonably well as long as the percentage of censored observations in the dataset is not too large. The percentage of censored observations recommended above is not based on research but is an educated guess, so it should be used with caution. In SAS a censored regression model may be estimated through PROC LIFEREG.

Component approach

For the component approach the paper separately discusses the estimation of CE for subscription services and businesses that sell products. For the former it discusses the estimation of attrition rates, revenues and cash flows and for the latter the purchase probability over time, revenues and cash flows.

Customer life

The two key components in calculating CE are customer cash flows and the duration of the cash flows, or customer life. Take customer life first. Two situations present themselves, depending on the nature of the business. For subscription businesses the attrition event is typically observed or can be readily inferred when there is no payment, even after repeated efforts to get the customer to renew. For other businesses involving product purchases, like retailing, where the customer decides when, where and what to buy, the attrition event is never observed and the absence of any purchase could imply attrition. There is greater uncertainty in purchase timing in the latter type of business and in some categories like consumer durables, where purchase intervals can be rather lengthy, the difficulty of estimating customer life increases substantially. The paper first considers subscription services businesses.

Subscription services

Computing CE using attrition rates

Insurance, banking, telecommunications and motor clubs are all examples of subscription services. In computing the CE of a new customer, one needs an estimate of the average life of a new customer. Alternatively, one needs the conditional probability of attrition of a new customer in any future period, t , given the customer survived the previous $t-1$ periods. A simple example can help illustrate the calculations (Table 2).

The example shown in Table 2 is a useful back-of-the-envelope formula for estimating CE if payments are monthly and the monthly attrition rate declines gradually (due to loyalty). If one is cautious in setting the constant attrition rate somewhat lower than what is observed in

Table 2: Calculation of CE: A simplified example

<p>Assumptions:</p> <ul style="list-style-type: none"> — monthly cash flow 'F' is fixed and occurs at the start of each month — monthly percentage attrition rate 'a' is fixed — monthly percentage discount rate 'd' is fixed. <p>CE, calculated as the NPV of cash flows, is then given by:</p> $CE = F + F(1 - a/100)/(1 + d/100) + F[(1 - a/100)/(1 + d/100)]^2 + \dots$ <p>In the above expression, in the second term, $(1 - a/100)$ represents the fraction of the customer surviving the first month and $1/(1 + d/100)$ is the one-month discount factor for the monthly cash flow, F. Likewise, $(1 - a/100)^2$ represents the fraction of the customer surviving two months and $1/(1 + d/100)^2$ is the two-month discount factor. In the above expression, the number of terms is infinite but since $(1 - a/100)/(1 + d/100)$ is less than one the series converges.</p> <p>Defining 'r' as $(1 - a/100)/(1 + d/100)$,</p> $CE = F[1 + r + r^2 + r^3 + \dots]$ <p>The terms in brackets represent a geometric series that sums to $1/(1 - r)$. Hence,</p> $CE = F/(1 - r), \text{ or}$ $CE = F/[1 - (1 - a/100)/(1 + d/100)], \text{ or}$ $CE = F[(d + a)/(100 + d)]$

the earlier periods, it may give a reasonable first-order estimate of CE. Note that if payments are annual, the attrition rates will follow a cyclical pattern and will typically spike every 12 months. This is often referred to as the renewal pattern of attrition. In such a case, CE should be estimated based on the observed monthly attrition rates.

The percentage attrition rate for any period, for any cohort of customers, may be estimated by tracking that cohort over time and noting the number of customers that attrite in each period. The attrition rate for any period is then estimated as the number that left the company in that period divided by the number of customers at risk of attrition at the start of the period. The retention rate in any period is the complement of the attrition rate and is computed as one minus the attrition rate, where the latter is stated as a fraction.

Modelling attrition rates

For subscription services, the attrition rate for a customer, in any period t, may be estimated by the conditional probability of the customer leaving the company in period t, given the customer survived the previous t-1 periods. This probability is often referred to as the discrete hazard probability or hazard rate. The table of hazard probabilities (t versus hazard probabilities) is referred to as the discrete hazard function. The hazard function can take many shapes. A declining hazard function is typically seen in many subscription services, implying increasing loyalty over time. In other cases the hazard function may be constant (ie a fixed attrition rate), and in yet other situations, like 'winback', where a company successfully reacquires a lost customer, the hazard function can be increasing (the company that just lost the customer may try to win him back or save him immediately) and then decreasing again after a while (if the save is unsuccessful).

Over the last 35 years or so statisticians have extensively studied the

Calculation of attrition rate

Survival analysis

phenomenon of the timing of events. This field is alternatively referred to as hazard analysis, or survival analysis or event history analysis, and has been and continues to be used extensively by sociologists (for instance to study recidivism rates), actuaries (mortality rates), medical researchers (morbidity rates) and reliability engineers (component and system failure rates). More recently the credit industry has been using hazard regression models to predict the timing of defaults and defection. The interested reader with some background in statistics would find Kalbfleisch and Prentice¹⁰ and Cox and Oakes¹¹ classic references. Those with an interest in sociology should refer to Yamaguchi.¹² Over the last 15 years or so marketing scientists have found significant interest in this area and have published articles on the purchase timing of consumer products.^{13,14}

Note that, since the attrition probabilities are a function of time, the remaining life of a customer depends on how long she has been with the company. This is not unlike the mortality tables used by actuaries, where the average lifespan of a new-born is different to that of an adult of, say, 80 years in age. This same principle holds in assessing the CE of new and veteran customers. Hence, in formulating a decision rule for customer acquisition one uses the CE of a new customer, whereas for customer retention one defines CE as the remaining value of an existing customer.

In estimating CE at the customer level one would need for each customer an estimate of her hazard function. This may be estimated through hazard regression analysis. A discrete hazard regression model can be built using logistic regression with a panel dataset. Here a cohort of customers is tracked over time and each customer contributes multiple observations to the dataset. The observed response variable is binary and equals zero or one, based on whether the customer is active or has lapsed in that period. For instance, a customer who lapsed in the third period contributes three observations with the response variable, equal to one (active) for the first two periods and zero for the third period.

The predictor variables typically include a design matrix of indicator, or dummy, variables of zeros and ones to capture separately the effect of each time period. If the hazard model is to be used to predict the attrition rates of prospects or new customers then the predictor variables must be fixed over time (rather than being time-dependent), since their values need to be known at time zero. In general, however, this formulation based on logistic regression does allow for time-varying predictors. This is particularly useful when hazard regression is used for explanatory purposes. For instance, if the hazard rate of a credit card customer in a given period depends on his credit score in the previous period, then the time-varying credit score (lagged by one period) could be used as a predictor variable. Note, even though the observations associated with each customer will be auto-correlated, the logistic regression model will still produce unbiased standard error estimates.¹⁵ Table 3 shows an algebraic formulation of what is described above.

Model validation

It is always good practice in modelling to validate a model on a hold-out sample. In direct marketing practitioners typically build a model on one sample and validate it on a similar hold-out sample. This is done by comparing the model's discriminatory performance through gains tables

Table 3: Discrete hazard model

Let:
 x_j denote the $(n \times 1)$ column vector of predictor variables for customer j
 B represent the $(1 \times n)$ row vector of partial logistic regression coefficients for x_j
 t_i represent the time variable for period i , such that $t_i = 1$ for period i , or else $t_i = 0$ with $i = 1, \dots, T$ where T is the maximum number of periods in the study period
 C_i is the partial logistic regression coefficient of t_i
 h_{ji} is the hazard for customer j in period i
 A represents the constant term in the logistic regression model.
 Then, h_{ji} may be estimated by the following model:

$$\ln [h_{ji}/(1 - h_{ji})] = A + B x_j + C_i t_i$$

 Here, $\ln$ is the natural logarithm and $h_{ji}/(1 - h_{ji})$ is the odds ratio. A and the row vector B and C_i for $i = 1, \dots, T$ are estimated through maximum likelihood.
 For each customer with profile x_j the model allows one to estimate the customer's hazards for the time periods $1, \dots, T$.

constructed on the model development and hold-out samples. A model is considered to validate well when the two gains tables are nearly identical. In comparing the two tables, the decile (rows of the gains table are generally deciles or demi-deciles of scored customers) score ranges should be similar and so should the observed mean decile values of the modelled characteristic (which would be attrition rates in this case.) The gains table built on the hold-out sample is also typically used as a forecasting tool. Instead of using the model's predicted value directly, practitioners often use the mean of the observed value in the decile in which the scored customer falls. This essentially is equivalent to smoothing, and also adjusts for any systematic bias in the model.

By using the logistic regression formulation for estimating the discrete hazard function, one is assuming that each customer's hazard rate is constant within a given period. This may or may not be true. From a practical standpoint, however, where the hazard function is used to estimate CE this may be reasonable, as long as one defines the duration of the periods carefully. For instance, if the average duration of a customer is, say, one year, using a year as the duration the period would be too coarse to be meaningful. Months, or at most quarters, would make more sense.

Estimating revenues for subscription services

For services like telecommunications, customers' monthly revenues may not vary substantially. In such instances, a mean monthly revenue may be estimated through standard multiple regression. For estimating CE of new customers one would again use only prospect characteristics for the predictors, and the average monthly revenue of active customers as the target or criterion variable.

If the assumption of constant customer revenue per month is not appropriate, the approach could be refined further by using PROC MIXED in SAS for repeated observations. This would allow one to develop a separate revenue estimate for each period. Similar to the discrete hazard model, a design matrix of indicator variables would help estimate the unique additive effect of each time period. Here, each active customer would contribute multiple observations to the dataset and a

covariance matrix would be estimated for the observations over time. A variety of covariance structures can be tested, such as first-order autocorrelation, and the structure that best fits the data is typically selected for the model.¹⁶

Estimating CE for subscription services

Once the hazard rates and revenues are estimated through logistic and multiple regression respectively, each prospect's CE may be estimated by using these two estimates. The revenue estimates must first, however, be used to estimate cash flows. One way to do this may be through regression analysis. For a random sample of customers, compute their cash flow for a randomly selected month (this assumes that the revenues do not vary significantly period to period, or else a repeated-observations approach would again have to be used), based on the recommendations made earlier. Next, build a regression model to predict cash flow as a function of revenue and other prospect/customer characteristics. Test a variety of functional forms for the predictors, and pick the one that produces the best-fitting model. Table 4 shows how the hazard rates and the cash flow estimates can be used to compute CE.

In the example in Table 4, $(1 - h_{i1})$ represents the proportion of the customer that is active at the end of the first period; likewise, $(1 - h_{i1})(1 - h_{i2})$ is the proportion of the customer that is active after two periods and $(1 - h_{i1})(1 - h_{i2}) \dots (1 - h_{i(N-1)})$ is the proportion active after $(N-1)$ periods. There is no hard-and-fast rule regarding how large N should be. Using four or five years as the duration of the study period (N) ensures that one is not using truly ancient data and also allows one to be somewhat cautious. It is recommended that N should exceed the median duration of the relationship; that is, N should be large enough that at least 50 per cent of the customers are deemed inactive after N periods. If the hazard rates have been estimated, they may be used to estimate the median duration of the relationship.

Product purchases

Modelling purchase rates

In computing CE for businesses that sell products the purchase rate is relevant, and although the probability of purchase in any period can be modelled via logistic regression, many marketing scientists have

Defining the hazard rate

Table 4: Calculation of CE based on modelling outputs

<p>Let: h_{it} be the ith customer's estimated hazard in period t, based on a discrete hazard model r_i be the ith customer's estimated revenue per period (assumes customers' revenues do not vary significantly over the time periods) c_i be the ith customer's cash flow per period (assumed to occur at the start of each period) where c_i is estimated from r_i via regression as discussed above d be the discount rate per period. Then, CE_i the customer equity for customer i over N quarters is given by:</p> $CE_i = c_i + c_i(1 - h_{i1})/(1 + d) + c_i(1 - h_{i1})(1 - h_{i2})/(1 + d)^2 + \dots + c_i(1 - h_{i1})(1 - h_{i2}) \dots (1 - h_{i(N-1)})/(1 + d)^{N-1}$
--

modelled the purchase rate as a continuous function of time. In this case, the continuous hazard function, $h(t)$, may be viewed as the instantaneous purchase rate for an individual who did not make a given (first or second or third or n th) purchase prior to time t . For this continuous case, since $h(t)$ is not a probability but a rate, it is not constrained to fall between zero and one and can take on any non-negative value. It is mathematically defined as the ratio of the density function $f(t)$ of the random variable T , the time to occurrence of the event, to its survival function, $S(t)$. That is, $h(t) = f(t)/S(t)$.

Continuous random variables, such as T , have associated with them density functions (like the exponential, or the gamma) which are used to compute the probability of the random variable being less than (or equal to) some value, say t . This probability is typically written as $F(t)$, which is also called the distribution function of T and may be computed by the area under the density function, up to the value t . The survival function $S(t)$ is the probability that the event occurs after time t and hence is given by $1 - F(t)$.

Continuous hazard functions for individuals with a given profile, x (predictor variables), may be estimated using hazard models like the accelerated failure rate (AFR) regression model (which is essentially a censored regression model) by making some assumptions about the distribution of T and then verifying those assumptions statistically or through Cox's proportional hazard (PH) regression, which makes no distributional assumption. In SAS, the AFR regression model may be estimated via PROC LIFEREG and the PH regression model via PROC PHREG.

Though product purchase timing may be modelled as a continuous random variable, for estimating CE it is easier to estimate the probability of purchase in each period and then combine these probabilities with the estimated cash flow given a purchase occurred, to compute an expected cash flow at the customer level. The conditional purchase probabilities can be estimated through logistic regression; that is, through a discrete hazard model as described above. Estimating the cash flows at the customer level has its own challenges. A recommended method is estimating customer revenues, conditional on purchase, and then applying an estimated cash flow margin that varies by customer and purchase occasion to the estimated revenue.

Estimating revenues for product purchases

Revenues generated for each of the purchase occasions may be estimated using panel data regression. Selecting an appropriate sample for the analysis is important. The study design could consist of a single cohort of customers acquired say four or five years ago (the duration of the study period should be determined by the customers' purchase rates and the duration over which one wishes to compute the CE) or multiple cohorts acquired in different months in the last four or five years. Though the latter may be a more representative sample, it increases the modelling complexity. The modelling implications for the multiple cohort samples will be discussed briefly later in this section.

Panel data regression

Once a cohort is selected, the dataset for modelling needs to be assembled. Here, as in the discrete hazard model, each customer contributes multiple observations to the dataset but each observation now corresponds to a purchase occasion. Hence, if a customer purchased, say, seven times in the study period, that customer would contribute seven observations to the dataset. If one is estimating the CE for new customers, only prospect characteristics may be used as predictor variables. These will be fixed and not time-varying. In addition a set of (0/1) indicator predictor variables may be used to represent the effects of each time period. If, for example, there are 16 quarters in the study period, there would be 16 such indicator variables. One can also test if the effects of the predictor variables change with time by using interaction terms of the predictors with the time period indicator variables.

The panel data regression model explicitly accounts for the correlations among the observations contributed by each customer through a covariance matrix. This ensures that the model's standard errors are not misestimated. In SAS this model may be estimated through PROC MIXED for repeated observations. Different covariance structures can be assumed and the procedure estimates the additional covariance terms. Statistical tests (based on information theory) can also help identify the best covariance structure.

If multiple cohorts are used in the modelling sample, additional *random effects* must be assumed for each cohort. This is because there may be something unique about the customers in each cohort that is not captured by the predictor variables (such as the characteristics of the marketing campaign used to acquire the cohort). These additional random effects would contribute additional variance terms to the standard errors.

PROC MIXED is SAS's most recent and advanced procedure for multiple regression. It is aptly named MIXED since it can handle both random effects and repeated observations simultaneously. For the non-technical reader, essentially PROC MIXED takes care of the correlations among observations (the basic linear regression model assumes independence across observations) by estimating additional variance and covariance parameters. This ensures that the variances of the coefficients are not misestimated, which could result in incorrect conclusions regarding the effects of the predictor variables.

The above formulation needs to be adjusted slightly since the revenue estimates would be used in conjunction with the probability of purchase in a period. This requires that there is at most only one purchase per period. If the periods are quarters and if there are only a small number of customers making more than one purchase per quarter, these purchases may be combined by summing their revenues without losing too much accuracy.

The final complexity is that the revenue and the purchase probability may be correlated. If this is not accounted for, the parameters of the regression model (for estimating the revenues, given a purchase occurred) will be biased. Though this complexity may be dealt with by estimating the parameters of both models simultaneously (using the maximum likelihood approach), many researchers use the Heckman correction

Improving the estimates

factor (see Franses and Paap¹⁷). Here the regression model in the second step is expanded by an additional variable, the odds of not making the purchase, often referred to as the *inverse Mill's ratio*, whose values are calculated from the first step's logistic regression model. Adding this additional variable not only allows one to test whether the purchase probability and revenue are independent (by testing the coefficient of the inverse Mill's ratio) but also produces unbiased parameter estimates.

Estimating CE for product purchases

For each period one can now calculate a probability of purchase and a conditional revenue estimate for each customer. As discussed above, one can also estimate the customer's expected cash flow as a function of his revenue and other prospect/customer-based predictors using regression analysis. The product of the purchase likelihood and estimated cash flow would then give an estimate of expected cash flow for that period. Repeating this calculation for all periods in the study period results in an expected cash flow stream for that customer, whose PV is the estimate for CE. Table 5 develops the algebraic expression

Note that since all customers must make an initial purchase, the first term in Table 5 is not adjusted with a probability. Like all model-building efforts, the models built to estimate the hazard rates, purchase likelihood and revenues should be validated on hold-out samples and the discrimination power of the models demonstrated via gains tables, as is standard in direct marketing modelling. For panel-data-based models, a separate gains table should be constructed for each period using only the observations corresponding to that period. As discussed earlier, the gains tables developed on the hold-out samples may be used to estimate smoothed forecasted values for each customer.

Key assumptions

Computing CE is essentially a forecasting problem as it requires one to estimate customer cash flows over future time periods. Hence the fundamental assumption of *stationarity*, used extensively in forecasting, is equally valid here. A stochastic (probabilistic) process is said to be stationary if the joint distribution generating the data is invariant with regard to the displacement in time.¹⁸ This means, for example, in a regression model the joint probability distribution of the response

Defining stationarity

Table 5: CE for product purchases

<p>Let:</p> <p>p_{it} be the ith customer's estimated purchase probability in period t, for $t = 2, \dots, N$ based on a discrete hazard model</p> <p>r_{it} be the ith customer's estimated revenue in period t given a purchase occurred in period t, for $t = 1, \dots, N$</p> <p>c_{it} be the ith customer's cash flow in period t, for $t = 1, \dots, N$ (assumed to occur at the start of each period) where c_{it} is estimated from r_{it} via regression as discussed above</p> <p>d be the discount rate per period.</p> <p>Then, CE_i, the customer equity for customer i over N periods, is given by:</p> $CE_i = c_{i1} + c_{i2} p_{i2} / (1 + d) + c_{i3} p_{i3} / (1 + d)^2 + \dots + c_{iN} p_{iN} / (1 + d)^{N-1}$

variable, y_t , over the time periods $t, t + 1, \dots, t + k$ is equal to the joint distribution at time $t + m, t + 1 + m, \dots, t + k + m$ for all integer values of m and k . This assumption allows one to build a model using data over, say, the 20 quarters for the years 2000–2004 and apply it to any other 20 quarters, say 2001–2005.

Processes are, however, seldom stationary over long periods of time and almost all models are always mis-specified to some extent. This does not mean that the models proposed here are of limited value, but it does mean they need to be periodically validated to ensure that they continue to predict with acceptable accuracy. If the prediction accuracy deteriorates to an unacceptable level, creating a significant negative financial impact, it may be time to recalibrate the models, or in certain instances recreate the entire models, including the investigation and use of new predictors.

Calculating and using CE for decision making is thus not a one-time proposition, but requires the company to set up ongoing tracking and updating mechanisms. The managerial challenges are briefly discussed next.

Organisational requirements

The number one requirement is senior management's acceptance and strong belief that CE can be measured and managed/enhanced through customer-focused initiatives. Once a company accepts this philosophy it can develop the strategies and processes for implementing CE-enhancing strategies. For most companies this implies a radical change in how the marketing function is carried out and requires the organisation to acquire new skills and capabilities. Various authors^{19–22} have discussed the organisational activities, structure, measurement metrics and management's role in creating the right culture and support systems for implementing a CE-based marketing system. Bell *et al.*²³ discuss the barriers to CE management.

A CE-based marketing system can only be established if a specific group of professionals is charged with the ongoing development and refinement of the CE estimates and their appropriate use in decision rules for managing customers. Members of this team need to have backgrounds in statistics, management science and cost accounting, and processes need to be established for:

- building, maintaining, using and updating the models and cost estimates
- producing the analyses to support marketing decisions, particularly with respect to customer acquisition, development and retention
- development of a testing budget that prioritises the customer knowledge acquisition efforts as well as designing tests and analysing test results
- creation of reports that track marketing campaigns and compare them to earlier projections
- producing reports that track the value of new and veteran customers over time as well as key customer metrics for compensating

Senior management's role

employees based on their role and efforts in enhancing CE (see Hansotia^{24,25} for a discussion on customer value scorecards).

Team approach

Marketers' responsibilities include understanding customers' needs and expectations intimately and designing creative value propositions where value is appropriately allocated between the firm and its customers.^{26,27} For CE-enhancing strategies to work, many different types of professionals must be able to work together seamlessly. Sophisticated analytic tools are by no means sufficient if they are not backed up by technology that captures the right data that are accurate, consistent and conveniently accessible. Likewise, unless the company can develop great products and services, the enhancement of customer value will be severely handicapped.

Summary

This paper discussed the applications of CE and then moved on to presenting how CE can be computed for subscription services and businesses that sell products. It separately outlined the challenges in computing cash flows at the customer level as well as the estimation of attrition probabilities or hazard rates. It discussed how with slight adjustments the same types of models can also be used to estimate purchase likelihoods and purchase timing. The models presented here were primarily based on longitudinal data, typically available in customer databases, and are similar in spirit to those discussed in Reinartz and Kumar.²⁸ Such models do not explicitly allow for competitive effects. In a recent article, Rust *et al.*²⁹ used customer-specific brand-switching matrices to accommodate such effects directly. Their work also allowed for heterogeneous interpurchase times. Their approach, however, is based on cross-sectional data collected through surveys. The paper continued with a brief discussion of stationarity, a key underlying assumption in using the model estimates for predicting future values, and concluded with a brief discussion of organisational requirements for implementing a CE-based marketing system.

The approaches discussed here work well in stable markets and the paper briefly touched upon the situation where there is little customer information available when a company is just starting out, or when there is a brand new marketing channel. What was not discussed was the customers' response when industry-changing initiatives are introduced by a company. First, the company would have to estimate the changes in customers' behaviours as a result of such innovations, and then re-estimate their reactions when the competition responds. Needless to say, estimating and updating CE in highly dynamic and new markets increases complexity significantly. The advice is to continue to build the types of models presented, as data become available, but to update the models' estimates judgmentally, or combine them with managerial input, until better data are available. By no means can models and analytical methods eliminate the need for good judgment and an understanding of the marketplace, but they do provide a mechanism for making assumptions

explicit and for evaluating the impact of changes in the inputs (sensitivity analysis). Thus they provide a useful structure for sound decision making.

References

1. Blattberg, R. C. and Deighton, J. (1966) 'Manage marketing by the customer equity test', *Harvard Business Review*, July/August, pp. 136–144.
2. Blattberg, R. C., Getz, G. and Thomas, J. S. (2001) *Customer Equity: Building and Managing Relationships as Valuable Assets*, Harvard Business School Press, Boston, MA.
3. Rust, R., Lemon, K. and Zeithaml, V. (2000) *Driving Customer Equity*, Free Press, New York, NY.
4. Hansotia, B. J. and Wang, P. (1997) 'Analytical challenges in customer acquisition', *Journal of Direct Marketing*, Vol. 11, No. 2, Spring, pp. 7–19.
5. Hansotia, B. J. and Rukstales, B. (2002) 'Incremental value modeling', *Journal of Interactive Marketing*, Vol. 16, No. 3, pp. 35–46.
6. Hansotia, B. J. and Rukstales, B. (2002) 'Direct marketing for multi-channel retailers: Issues, challenges and solutions', *Journal of Database Marketing*, Vol. 9, No. 3, pp. 259–266.
7. Gupta, S., Lehmann, D. R. and Stuart, J. A. (2004) 'Valuing customers', *Journal of Marketing Research*, Vol. XLI, February, pp. 7–18.
8. Cokins, G. (1996) *Activity-Based Cost Management: Making It Work*, Irving, Burr Ridge, IL.
9. Littell, R. C., Milliken, G. A., Stroup, W. W. and Wolfinger, R. D. (1996) *SAS System for Mixed Models*, SAS Institute, Cary, NC.
10. Kalbfleisch, J. D. and Prentice, R. L. (1980) *The Statistical Analysis of Failure Time Data*, Wiley, New York, NY.
11. Cox, D. R. and Oakes, D. (1984) *Analysis of Survival Data*, Chapman and Hall, London, UK.
12. Yamaguchi, K. (1991) *Event History Analysis*, Sage Publications, Newbury Park, CA.
13. Jain, D. C. and Vilcassim, N. J. (1991) 'Investigating household purchase timing decisions: A conditional hazard function approach', *Marketing Science*, Vol. 10, No. 1, pp. 1–23.
14. Helsen, K. and Schmittlein, D. C. (1993) 'Analyzing duration times in marketing: Evidence for the effectiveness of hazard rate models', *Marketing Science*, Vol. 12, No. 4, pp. 395–414.
15. Allison, P. D. (1995) *Survival Analysis Using the SAS System: A Practical Guide*, SAS Institute, Cary, NC.
16. Littell *et al.*, ref. 9 above.
17. Franses, P. H. and Paap, R. (2001) *Quantitative Models in Marketing Research*, Cambridge University Press, Cambridge, UK.
18. Nelson, C. R. (1973) *Applied Time Series Analysis for Managerial Forecasting*, Holden-Day, San Francisco, CA.
19. Hansotia, B. J. (2002) 'Gearing up for CRM: Antecedents to successful implementation', *Journal of Database Marketing*, Vol. 10, No. 2, pp. 121–132.
20. Hansotia, B. J. (2004) 'Company activities for managing customer equity', *Journal of Database Marketing and Customer Strategy Management*, Vol. 11, No. 4, pp. 319–332.
21. Hansotia, B. J. (2004) 'Customer metrics and organizational alignment for maximizing customer equity', *Journal of Database Marketing and Customer Strategy Management*, Vol. 12, No. 1, pp. 9–20.
22. Blattberg *et al.*, ref. 2 above.
23. Bell, D., Deighton, J., Reinartz, W. J., Rust, T. R. and Swartz, G. (2002) 'Seven barriers to customer equity management', *Journal of Services Research*, Vol. 5, No. 1, pp. 77–85.
24. Hansotia, B. J. (2002–2003) 'The customer value scorecard', *Journal of Integrated Communications*, pp. 34–39.
25. Hansotia, ref. 21 above.

26. Bell *et al.*, ref. 23 above.
27. Hansotia, B. J. (2002–2003) ‘The customer value scorecard’, *Journal of Integrated Communications*, Northwestern University, Department of Integrated Marketing Communications, pp. 34–39.
28. Reinartz, W. and Kumar, V. (2000) ‘On the profitability of long lifetime customers: An empirical investigation and implication for marketing’, *Journal of Marketing*, Vol. 64, No. 1, October, pp. 17–35.
29. Rust, T. R., Lemon, K. N. and Zeithaml, V. A. (2004) ‘Return on marketing: Using customer equity to focus on marketing strategy’, *Journal of Marketing*, Vol. 68, No. 1, January, pp. 109–127.